



**ANJANETTE H. RAYMOND:**

Submitted in partial fulfillment of the requirements of the  
Degree of Doctor of Philosophy

**Managing Bias, Partiality, and Dependence  
in Online Justice Environments**

Anjanette H. Raymond

PhD Submission 2019



## STATEMENT OF ORIGINALITY

I, Anjanette H. Raymond confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature: *Anjanette H. Raymond*

Date: 14 August 2019

## Abstract

As Online Dispute Resolution (ODR) begins to enter the wider market, it promises to substantively change the way arbitration works in the justice eco-system. As such, the justice community must determine the fundamental values to be incorporated into the online justice environment so that it can address ethical and regulatory issues in the design of ODR. This thesis is the first attempt of any ODR scholar to identify the fundamental values that must be embedded in the decision making portion of the ODR process.

This thesis advocates for the principle of judicial independence and impartiality: decision makers in the arbitration system must be *both* impartial *and* independent, without exception, even if the adherence to this rule of law principle conflicts with maximizing the efficiencies of the overall system. As it relates to the decision making process that incorporates technology within the model, technology driven bias must be identified and mitigated against. Consequently, the ODR system must be examined through the lens of risk identification and mitigation principles. This thesis is the first, to examine the ODR model within the technology driven aspects of model design.

This thesis examines Rule of Law within the context of decision making in the online justice environment. The thesis argues for adherence to seven rules of action: (1) the rule of law (RoL) must be protected within the online justice environment; (2) an independent and impartial decision maker is an essential RoL principle, such that no departure from the principle should be allowed; (3) it is the method of ensuring the independence and impartiality of the decision maker that is open for greater debate; (4) it is the model design that mitigates dependence and bias; (5) the reduction of negatively impactful bias is the current “best practice” standard; (6) there must be a human within the final step of the technology driven decision making model; (7) accountability is essential as ODR is part of the justice eco-system. The thesis concludes by creating a first of its kind, best practices guidelines with specific examples of risk identification and mitigation principles that must be built into the design.

# Table of Contents

<b>STATEMENT OF ORIGINALITY .....</b>	<b>2</b>
<b>ABSTRACT .....</b>	<b>3</b>
<b>INTRODUCTION.....</b>	<b>7</b>
<b><u>CHAPTER 1: THE RULE OF LAW .....</u></b>	<b>13</b>
A. THE RULE OF LAW .....	13
B. THE RULE OF LAW: MEASURING THE RULE OF LAW .....	26
Table A: WJP Rule of Law Index.....	30
C. THE RULE OF LAW: AN INDEPENDENT JUDICIARY .....	31
D. THE RULE OF LAW: AN IMPARTIAL JUDICIARY.....	34
E. RULE OF LAW: THE LAW AND LEGAL REASONING.....	41
1. <i>Legal Formalism</i> .....	41
2. <i>Legal Realism</i> .....	42
3. <i>Artificial Intelligence and Legal Reasoning</i> .....	43
F. RULE OF LAW CONCLUSIONS .....	45
<b><u>CHAPTER 2: ARBITRATION UNDERPINNINGS .....</u></b>	<b>46</b>
A. ARBITRATION: IMPARTIALITY AND INDEPENDENCE.....	47
Table B: Arbitral Rules of Impartiality/Independence Standards .....	48
B. ARBITRATION: DISCLOSURE AND CHALLENGES.....	52
1. <i>Disclosure</i> .....	52
Table C: Arbitral Disclosure Standards .....	52
2. <i>Institutional Action Concerning the Decision Maker</i> .....	62
3. <i>Transparency of Outcomes</i> .....	64
C. THE ROLE OF BRICK-AND-MORTAR COURTS .....	66
D. A REVIEW OF ENGLISH LANGUAGE CASES .....	68
1. <i>Ties to a Party’s Legal Practice</i> .....	69
2. <i>Prior Arbitration Work</i> .....	69
3. <i>Other Professional Ties</i> .....	70
4. <i>Social Ties</i> .....	70
5. <i>Multiple Arbitrator Parties</i> .....	70
E. ARBITRATION OVERVIEW CONCLUSIONS .....	71
<b><u>CHAPTER 3: ONLINE DISPUTE RESOLUTION .....</u></b>	<b>72</b>
A. FROM ALTERNATIVE DISPUTE RESOLUTION TO ONLINE DISPUTE RESOLUTION.....	73
1. <i>Private ODR Platforms</i> .....	75
Table D: Currently Viable ODR Platforms.....	80
2. <i>Public ODR Platforms</i> .....	87
3. <i>Supra-National/Institutional Efforts</i> .....	95
B. ONLINE DISPUTE RESOLUTION: APPOINTMENT, IMPARTIALITY, INDEPENDENCE, AND DISCLOSURE .....	100
1. <i>Institutional Guidance Concerning ODR Decision Maker</i> .....	100
a. Appointment, functions and independence/impartiality of the decision maker .....	101
b. Governance .....	105
c. Transparency.....	106
2. <i>ODR Providers Identified Manner of Appointment</i> .....	108
Table E: ODR Providers Selection of the Decision Maker .....	109
3. <i>Providers’ Rules of Independence and Impartiality</i> .....	112
Table F: Current ODR Institutional/Platform Attempts to ensure Independence and Impartiality.....	112
C. SMART CONTRACTS DEVELOP A NEW STANDARD .....	117
1. <i>Smart Contracts and ODR</i> .....	121
Table G: Smart Contract Providers Considering ODR.....	121
2. <i>Smart Contract Developers Technology-Driven Solutions</i> .....	124

*Table H: Smart Contract Developers Technology-Driven Attempts to Manage Independence and Impartiality of the Decision Maker* ..... 125

D. ONLINE DISPUTE RESOLUTION SUMMARY CONCLUSION ..... 131

**CHAPTER 4: THE NEXT STEP — A NON-HUMAN ACTOR AS AN ASSISTANT TO THE DECISION MAKER IN THE ONLINE JUSTICE SYSTEM** .....132

A. THE BASICS OF AI AND THE LAW ..... 133

    1. Models..... 136

*Schematic 1: Financial Institution Model* ..... 138

    2. Learning ..... 139

    3. Model Risk ..... 142

        a. Model Risk: Biased Inputs ..... 143

        b. Model Risk: The Challenge of Social Complexity ..... 146

            i. The Framing Trap ..... 147

            ii. The Portability Trap ..... 149

            iii. The Formalism Trap ..... 150

            iv. The Ripple Effect Trap..... 152

            v. The Solutionism Trap ..... 157

B. TECHNOLOGY BASED DECISION MAKERS ..... 158

    1. Expert Systems ..... 158

    2. Computational Models of Legal Reasoning..... 161

    3. Analytics, Feedback and Rankings ..... 162

C. RULE OF LAW CONSIDERATIONS IN LIGHT OF UBIQUITOUS TECHNOLOGY ..... 163

    1. Rule of Law: Due Process and Technology ..... 164

    2. Rule of Law: Consensus and Society-in-the-Loop..... 166

    3. Rule of Law: Code Replaces the Common Law of Precedent ..... 167

    4. Rule of Law: Powerful Quasi-Governmental Actors without Accountability..... 168

    5. Rule of Law: Human-in-the-Decision-Making-Loop ..... 171

D. REGULATIONS OF ARTIFICIAL INTELLIGENCE ..... 173

E. LAW AND AI CHAPTER CONCLUSIONS..... 175

**CHAPTER 5: DESIGN GUIDANCE FOR THE REDUCTION OF PARTIALITY, DEPENDENCE, AND BIAS IN AN ONLINE JUSTICE DECISION MAKER**.....177

A. RULES OF ACTION ..... 180

    1. The Rule of Law (RoL) MUST be protected within the online justice environment..... 181

    2. An Independent And Impartial Decision Maker Is An Essential RoL Principle, Such That No Departure From The Principle Should Be Allowed..... 182

    3. It Is The Method Of Ensuring The Independence And Impartiality Of The Decision Maker That Is Open For Greater Debate ..... 182

    4. It is the Model Design that Mitigates Dependence and Bias..... 184

    5. The Reduction of Negatively Impactful Bias is the Current “Best Practice” Standard..... 185

    6. There Must Be A Human Within The Final Step of the Technology-Driven Decision Making Model..... 187

        a. Rejecting the Application of an Individual Jurisprudential School ..... 188

        b. Rejection of the “Robot” Decision Maker..... 191

        c. The Emotionally Intelligent Decision Maker ..... 194

    7. Accountability is Essential as ODR is part of the Justice Eco-System ..... 196

    8. Best Practices Demand a Wider Conversation about ALL of these Issues..... 200

B. DEFINING THE KEY TERMINOLOGY ..... 201

    1. Decision Maker..... 202

    2. Automated Decision Systems ..... 203

    3. Independence ..... 204

    4. Impartiality..... 204

    5. Bias..... 205

    6. Negatively Impactful Bias/Dependence/Partiality ..... 206

C. ADDRESSING BIAS IN TECHNOLOGY-DRIVEN DECISION MAKING MODELS ..... 206

    1. The Model ..... 207

- a. *The Reality of Legal Models* ..... 207
- b. *Model Selection*..... 209
- b. *Model Learning* ..... 211
- c. *ADR/ODR Decision Model* ..... 213
- 2. *Model Risk in the Decision Making Process*..... 215
  - a. *Model Risk Arising from Errors* ..... 215
  - b. *Incorrect Use Or Lack Of Understanding*..... 217
- 3. *Risk Mitigation In The Decision Making Eco-System*..... 218
  - a. *Managing Model Risk Arising within the Decision Process*..... 220
  - b. *Managing Model Risk in the Decision Maker* ..... 222
    - i. *Managing Risk by the Introduction of a Human in the Decision Making Process* ..... 222
      - (1) *The Role of Institutions and/or Platform Designer* ..... 223
      - (2) *The Roster of Decision Makers* ..... 224
      - (3). *Nomination and Appointment of the Human Decision Maker* ..... 226
        - Schematic 2: *The Suggested full Process*..... 226
        - Schematic 3: *Appointment*..... 227
        - Schematic 4: *Disclosures* ..... 229
        - Schematic 5: *Auditing Outcomes and the Feedback loop*..... 230
      - (4). *Challenges to the Appointment* ..... 231
      - (5). *Transparency of the Appointment Process and the Humans in the Process*..... 232
    - ii. *Risk Arising from the NON Disclosure/Rule Compliance* ..... 234
    - iii. *Managing User Risk*..... 236
    - iv. *Risk not already considered arising from the introduction of discretion* ..... 238
  - 4. *Risk Mitigation: Impact Assessments* ..... 239
- D. *DESIGN GUIDANCE CONCLUSIONS*..... 242
- VI. THESIS CONCLUSIONS**.....**244**
- APPENDIX A: DETAILS OF COLLABORATION AND PUBLICATIONS:**.....**254**
- APPENDIX B: INFLUENCES AND INFLUENCERS** .....**255**
- APPENDIX C: STYLE GUIDE** .....**257**

## Introduction

Technology is becoming ubiquitous in our daily lives. Enormous swathes of sectors such as finance, transport, and banking have all moved online, with speed, convenience and growing trust. The justice environment, while it has lagged behind business in its deployment of virtual communications, is also on the threshold of change. Many are now envisioning a fully online justice system. From kiosks and online platforms that allow citizens to speedily resolve traffic tickets, to e-discovery, to the electronic submissions of pleadings, many justice systems around the world have already begun the long process of technological transformation. Online Dispute Resolution (ODR), a form of arbitration, is one area of the justice ecosystem that is introducing technological transformation.

Online dispute resolution will be unlike any brick-and-mortar justice system before it. It will also be qualitatively different from previous deployments of technology in the arena of justice. For today, reliance upon technology increasingly implies the use of automation and artificial intelligence. And while no one is 100% sure what an online court will look like, many believe that it will require no brick-and-mortar courtrooms. A few believe it may not even require a human decision maker.

Justice systems around the world have already embraced automation to varying degrees. Automation, that is the technological process by which a procedure is performed without human input, is used in everything from paying parking tickets to disputing land boundaries. Until recently automated processes have been strictly limited to what can be accomplished using Yes/No decision trees. Conventional automated processes are rigid and non-deliberative. Most ultimately deliver information to a human decision maker. The introduction of artificial intelligence, however, raises the possibility of courts that operate with no physical space and no human decision maker.

So, now is the time to consider the fundamentally important issues raised by the use of non-human intelligence in the arbitration decision making process. As their full

implementation becomes feasible, we must be conceptually, legally and ethically prepared. This thesis will focus on the role of the decision maker in upholding one key aspect of the principles of rule of law, impartial decision making. Can technological solutions provide that? Will they support impartiality? Will they raise or lower the citizenry's expectation of it?

Before we can respond to these questions, we need to understand what computer-based decision making is. According to Greg Corrado from Google:

Machine learning *systems* are made up of three major parts: (1) Model: the system that makes predictions or identifications; (2) Parameters: the signals or factors used by the model to form its decisions; and (3) Learner: the system that adjusts the parameters — and in turn the model — by looking at differences in predictions versus actual outcome.<sup>1</sup>

Computer models of legal reasoning and computer models of legal argument seek to break down a complex set of human tasks into discrete steps, or an algorithm.<sup>2</sup> Thus, the key to discussing the *design* is to consider the *model*, which includes the incorporated *parameters* and the *learning* process of the system.<sup>3</sup>

In addition, to understand the following discussion it is important to have a brief historical overview of why arbitration is a central part of a justice system. First, regardless of which judicial system we consider, there is little room to debate the fact that many individuals face extreme limitations of access to justice.<sup>4</sup> Second, most judicial systems recognize the value of alternative dispute resolution as a mechanism of reducing these limitations.<sup>5</sup> And the majority consider arbitration as a well-

---

<sup>1</sup> Danny Sullivan, *How Machine Learning Works, As Explained By Google*, MARTECH: MANAGEMENT (Nov. 4, 2015)(describing in detail the process).

<sup>2</sup> See discussion, *infra* Chapter 4, Section A.

<sup>3</sup> See *id.*

<sup>4</sup> See discussion, *infra* Chapter 1, Section B and Table A.

<sup>5</sup> See discussion, *infra* Chapter 2, Section A. It is important to note, on August 7, 2019, the United Nations Commission on International Trade Law, *United Nations Convention on International Settlement Agreements Resulting From Mediation* (March 2019)(hereinafter, the Singapore Convention) opened for signature in Singapore. Over 40 states, including China, India and the U.S., have signed the treaty. See *id.* As such, the Convention is expected to enter into force shortly. See *id.* The Singapore Convention requires member states to enforce settlement agreements and gives effect to settlement agreements such that they may be used to prevent



entrenched alternative to traditional brick-and-mortar justice.<sup>6</sup> Beyond these points of similarity, however, there are major differences. There is a growing split between the E.U. and the U.S. when it comes to the prevailing approach to regulating arbitration. In the U.S., the system is governed primarily by federal law, resulting in a “hands off” regulatory approach, meaning the dispute resolution mechanism is a product of contract, crafted between two parties, with little interference from government.<sup>7</sup> What limited regulation exists is very pro-arbitration. In the E.U., a different approach is taken, resulting in a more protectionist regulatory structure, focusing on fundamental rights, limitations (or prohibitions) on the use of pre-dispute resolution clauses, and restrictions on the use of arbitration, especially as it relates to the use of technology.

In the U.S. context, the ODR system is developing in a private market, designed by private entities who seek to resolve disputes in a low-regulation environment, standing upon and seeking to maintain a long history of Supreme Court permissiveness toward arbitration. For business, the historical shifts introduced by technology present an opportunity to tailor contracts and the dispute system to their advantage, sometimes without observing the fundamental protections afforded to individuals in the traditional brick-and-mortar system. Because of this strategic advantage, we can assume that U.S. based businesses will continue to increase their use of technology as a decision maker in the resolution of disputes. In contrast, ODR in the E.U. is being developed by the existing justice system, thereby creating a stronger presumption that the technology will incorporate existing fundamental protections.<sup>8</sup>

---

parties from litigating the matter. *See id.* Convention Article 1 (3), specifies ““Mediation” means a process, irrespective of the expression used or the basis upon which the process is carried out, whereby parties attempt to reach an amicable settlement of their dispute *with the assistance of a third person or persons. . .*” *Id.* The use of the term ‘person’ may implicate awards issued in an automated process. Because the Convention is legally untested, part of an emerging international mediation community, and *may require* a ‘person’ as a mediator, the Convention will not form a central portion of the discussions within the thesis. However, as the Convention develops in the international commercial community, it should be considered as an important additional protection that could thus, be within the ODR process.

<sup>6</sup> *See discussion, infra* Chapter 2, Section C.

<sup>7</sup> *See discussion, infra* Chapter 2, Introduction.

<sup>8</sup> *See id.*

This thesis does not treat either system as presumptively good or bad. Rather, the differences between them will point us toward many of the key structural decisions that must be considered in the design of ODR. What is the ideal relationship between technological design and fundamental principles of the rule of law?<sup>9</sup> What are the essential features a judicial decision maker must embody in order to uphold the rule of law?<sup>10</sup> As they establish guiding principles and best practices, model designers will need to consider the legally proscribed parameters and protections associated with the legal conceptualizations of the key terms *independence* and *impartiality* as well as the mechanisms needed to ensure a neutral decision maker.<sup>11</sup> It is also important to ask what risk model should be used to evaluate the new systems being designed, and what mitigation can be deployed to reduce identified risks.<sup>12</sup>

On June 4, 2019 the Law Society of England and Wales published a report that concluded “that algorithm-based machine-learning has the potential to improve the criminal justice system — but warns of 'a worrying lack of oversight' over current experiments.”<sup>13</sup> This thesis arose out of the same concerns becoming readily apparent to the author during her years of work with U.N. Working Group III, Online Dispute Resolution. While ODR can improve access to justice and can open the door to greater efficiencies in the justice environment, the community must be careful to consider the essential elements of a justice environment, especially one that serves as an alternative to the traditional brick-and-mortar based system.

Early in the drafting of the thesis it became clear to the author that simple, uninformed arguments based on limited understanding of the full meaning of jurisprudence are no longer acceptable, especially in the face of mounting pressure from various sources to

---

<sup>9</sup> See discussion, *infra* Chapter 1.

<sup>10</sup> See discussion, *infra* Chapter 5, Section A.

<sup>11</sup> See discussion, *infra* Chapter 5, Section B.

<sup>12</sup> See discussion, *infra* Chapter 5, Section C.

<sup>13</sup> Michael Cross, *Justice Algorithms “Need Urgent Oversight” – Society*, THE LAW SOCIETY GAZETTE, (4 June 2019) available at <https://www.lawgazette.co.uk/law/justice-algorithms-need-urgent-oversight-society/5070478.article> (last visited Aug.5, 2019).

value speed and efficiency over Rule of Law. The thesis therefore, starts with an examination of the Rule of Law in one of the most essential areas of the justice system, and the one that stands to be most radically transformed with the advancement of technology, the decision maker. The author intends to demonstrate that the RoL demands adherence, without exception, to the principle of an impartial and independent decision maker, a key actor in a bias-free process. This commitment to impartial and independent decision maker demands that the final step in any dispute decision process resolved in arbitration must involve a human actor. As the thesis explores and explains in subsequent chapters, the current technology limitations, the absence of data to build bias free prediction, and the significant hesitations that arise from the way that humans interpret outcome predictions cannot be overcome via design. The technology captures the systemic bias that exists and while mitigation can lessen the impact, the impact is still- with existing technology and design- far too great to allow technology, via prediction devices, to simply produce an outcome.

Chapter 2 considers the existing alternative justice environments. While a review of existing options finds several similarities, especially in the area of the decision-maker appointment and challenge process, there is also a great deal of variation. The author views this as an opportunity to use the lessons learned in the creation and administration of these environments to inform the development of the emerging ODR eco-system. The justice community should embed well-founded and widely supported traditional approaches to alternative justice into new systems, while also embracing the use of technology as a means to improve the justice eco-system.

Chapter 3 considers the history of ODR and surveys the newest platforms within the area. Drawing from a short history, the thesis moves into guiding a consideration of technology deployments and their growing influence on the justice eco-system. While these are exciting times, technologists lead the delivery of the emerging system, creating an imperative for the justice system to develop an adequate understanding of the technology and craft guidance to ensure the rule of law is embedded within the online justice eco-system. Unfortunately, few guiding documents exist. Chapter 3 sets out the basic features of the existing ODR world, including the emergence of smart

contracts, and sets guideposts for the following chapters consideration of decision making model design.

Chapter 4 seeks to explain the basics of artificial intelligence and the law. Because it became increasingly clear that technologists were the ones designing the ODR system, the thesis works through the issues of risk and risk mitigation in the design/delivery of new technological models. The chapter surveys some more well-known technology deployments that embraced prediction mechanisms, especially those deployed in justice or enforcement environments as these are the most relevant corollaries to the data within the justice environment being used to build predictive outcome mechanisms. The examples are briefly examined to demonstrate to the reader that there are grave concerns about prediction in judicial and criminal enforcement environments. These examples, highlights and briefly explain how the use of data, gathered within a system of systemic inequality can and should not be used to build a predictive tool that is deployed in a justice environment. In response to these concerns, those who design technology have created design enhancements, Chapter 4 explains one of the common design approaches and introduces the risk identification, impact assessments and mitigation concepts and applications. Chapter 4 explains one of the common design approaches and introduces the risk identification, impact assessments and mitigation concepts and applications.

Chapter 5 is the culmination of the thesis in which the system itself is designed. While many imagine robots as judges, the technology is simply not that advanced. That is not to write that technology is not entering the justice environment, it's just not going to be walking into a courtroom any time soon. Design best practices are set out and a schematic of the potential system is introduced. Common risks are identified and potential mitigation actions are described.

Ultimately, this thesis seeks to inform the design of the new technological models that will inevitably change the form of the justice system.

## Chapter 1: The Rule of Law

As a society, we must ask, “How will we ensure that the rule of law is embedded within the technologies of the justice system?” This question is all the more important in light of the growing use of black-box, algorithm-based technology within the justice environment. As Chief Justice McLachlin notes, “[t]he rule of law may be undermined by laws which may require courts to operate in secrecy, away from the cleansing glare of sunlight.”<sup>14</sup>

This chapter seeks to provide an overview of the rule of law in terms of ODR compliance. It examines broad (but complex) historic and modern conceptualizations, then moves into an in-depth examination of the rule-of-law principle of an independent and impartial judiciary. The chapter closes by addressing considerations relating to the fundamental principles necessary within legal reasoning. Chapter 4, Section C. will return to the topic of technology impacts on RoL, and seek to resolve technology-driven challenges to the RoL.

### A. The Rule of Law

According to former U.N. Secretary General Ban Ki-moon,

The rule of law refers to a principle of governance in which all persons, institutions and entities, public and private, including the State itself, are accountable to laws that are publicly promulgated, equally enforced and independently adjudicated, and which are consistent with international human rights norms and standards.<sup>15</sup>

---

<sup>14</sup> Beverley McLachlin, *Reflecting Upon The Relevance of Magna Carta in Today's World*, Magna Carta Lecture: Relevance of Magna Carta in Today's World, BINGHAM CENTRE FOR THE RULE OF LAW (June 18, 2015) available at [https://www.biicl.org/documents/700\\_chief\\_justice\\_mclachlins\\_magna\\_carta\\_lecture\\_slh.pdf?showdocument=1](https://www.biicl.org/documents/700_chief_justice_mclachlins_magna_carta_lecture_slh.pdf?showdocument=1) (last visited Aug. 5, 2019).

<sup>15</sup> United Nations, *Rule of Law, Report of the Secretary-General on the Rule of Law and Transitional Justice in Conflict and Post-Conflict Societies*, available at <http://www.un.org/en/ruleoflaw/index.shtml> (last visited Aug. 5, 2019).

The rule of law encompasses many overlapping, highly complex governance issues.<sup>16</sup> Some argue that the rule of law, as a set of restraints on the state and the members of the ruling elite,<sup>17</sup> is rooted in the political and legal context of ancient Greece.<sup>18</sup> In a more modern understanding, the rule of law is often used to mean maintaining order among people, or as a proxy for “social harmony.”<sup>19</sup> The rule of law is an important component of the founding political documents of the U.S., including the U.S. Constitution. As John Locke proclaimed, “Wherever law ends, tyranny begins.”<sup>20</sup> Today many argue the rule of law embodies two ideals: protecting people from government and protecting people from each other.<sup>21</sup>

Scholars divide the rule of law into two broad categories: formal and substantive.<sup>22</sup> Formal rule of law is considered to include the manner in which the law is formulated, the clarity of the established law, and the prospective nature of the law.<sup>23</sup> Substantive

---

<sup>16</sup> See e.g., Adriaan Bedner, *An Elementary Approach to the Rule of Law*, 2 HAGUE JOURNAL ON THE RULE OF LAW, 50-52 (2010) (discussing the “Rule of law definitions seem bound to vary over time, place, context, and from author to author.” Id at 48.); Jeremy Waldron, *Is the Rule of Law an Essentially Contested Concept*, LAW AND PHILOSOPHY 21 (2):137-164 (2002) (stating: “there are bound to be alarm-bells ringing in analytic circles when a term like “the Rule of Law” is invoked so frequently on so many sides of so many issues in a fraught political debacle.”) Id. at 139.

<sup>17</sup> See Randall Peerenboom, *Human Rights and Rule of Law: What’s the Relationship?*, 36 GEORGETOWN JOURNAL OF INTERNATIONAL LAW, 827, (2004-2005).

<sup>18</sup> See e.g., Brian Tamanaha, ON THE RULE OF LAW: HISTORY, POLITICS THEORY, 7-14, CAMBRIDGE UNIVERSITY PRESS (December 13, 2004).

<sup>19</sup> See e.g., Nicolas Hachez and Jan Wouters, *Promoting the Rule of Law: A Benchmarks Approach*, 4, (April 30, 2013), available at <https://ssrn.com/abstract=2258331> (last visited Aug. 5, 2019).

<sup>20</sup> John Locke, SECOND TREATISE OF GOVERNMENT, Chapter 17, sec. 202, 1690, in online edition Jonathan Bennett (ed.) (2017) available at <http://www.earlymoderntexts.com/assets/pdfs/locke1689a.pdf> (last visited Aug. 5, 2019).

<sup>21</sup> Adriaan Bedner, *An Elementary Approach to the Rule of Law*, supra note 16 at 50-52 (2010); Brian Tamanaha, *A Concise Guide to the Rule of Law*, ST. JOHN’S UNIVERSITY LEGAL STUDIES RESEARCH PAPER SERIES, No. #07-0082, 3-7, (September 2007) available at: <http://www.ruleoflawus.info/The%20Rule/Tamanaha%20Concise%20Guide%20to%20Rule%20of%20Law.pdf> (last visited Aug. 5, 2019).

<sup>22</sup> See e.g., Bedner, *An Elementary Approach to the Rule of Law*, supra note 16, at 55.

<sup>23</sup> See id. at 56.

rule of law scholars acknowledge the same principles, but also tend to go farther and envision the rule of law as including substantive rights that are derived from adherence to the rule of law.<sup>24</sup>

Joseph Raz articulates a set of specific attributes that laws should have to be in compliance with the rule of law. He articulates the following standards:

- Laws should be prospective, not retrospective.
- They should be relatively stable.
- Particular laws should be guided by open, general and clear rules.
- There should be an independent judiciary.
- There should be access to the courts.
- The discretion which law enforcement agencies possess should not be allowed to undermine the purposes of the relevant legal rules.<sup>25</sup>

The rule of law, according to Raz, should function as a protection ensuring that specific laws are passed in the correct manner, and that individuals are enabled to make decisions and guide their own conduct based upon the law.<sup>26</sup>

Building upon the work of formal rule of law scholars, Dworkin adds a “capturing of moral right” component, arguing the rule of law must go beyond the formal. He defines the “moral rights” conceptualization of rule of law as follows:

It assumes that citizens have moral rights and duties with respect to one another, and political rights against the state as a whole. It insists that these moral and political rights be recognised in positive law, so that they may be enforced upon the demand of individual citizens through courts or other judicial institutions of the familiar type, so far as this is practicable. The rule of law conception is the ideal of rule by an accurate public conception of individual rights. It does not distinguish, as the rule book conception does, between the rule of law and substantive justice;

---

<sup>24</sup> *See id.* at 63.

<sup>25</sup> Joseph Raz, *The Rule of Law and Its Virtue*, in Joseph Raz, *THE AUTHORITY OF LAW: ESSAYS ON LAW AND MORALITY* (Oxford Univ. Press)(1979), Published to Oxford Scholarship Online, (March 2012) pp. 216-17 [hereinafter Raz, *THE AUTHORITY OF LAW*] available at <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198253457.001.001/acprof-9780198253457-chapter-11> (last visited Aug. 5, 2019).

<sup>26</sup> *See* Raz, *THE AUTHORITY OF LAW*, *supra* note 25, at 212.

on the contrary, it requires, as part of the ideal of law, that the rules in the book capture and enforce moral rights.<sup>27</sup>

Tom Bingham has advanced a definition of rule of law in terms of eight principles that encompass both procedural (sometimes called “thin”)<sup>28</sup> and substantive (sometimes called “thick”)<sup>29</sup> principles.<sup>30</sup> Thick and thin principles of the conceptualization of the rule of law is not without supporters and detractors, like the majority of topics within the broad discussion of the rule of law.<sup>31</sup> For example, Brian Tamanaha is often cited

---

<sup>27</sup> Ronald Dworkin, *A MATTER OF PRINCIPLE*, 11-12 HARVARD UNIVERSITY PRESS; Reprint edition (October 1985).

<sup>28</sup> See Bedner, *An Elementary Approach to the Rule of Law*, *supra* note 16, at 55.

<sup>29</sup> See *id.* at 55.

<sup>30</sup> See Tom Bingham, *THE RULE OF LAW*, 37 (Penguin; Reprint edition (24 Feb. 2011) [hereinafter Bingham, *RULE OF LAW*].

<sup>31</sup> The definition of the rule of law [summarizing, “these disagreements strongly indicate that the rule of law is an essentially contested concept” see Jørgen, Møller and Svend-Erik Skaaning, *Systematizing Thin and Thick Conceptions of the Rule of Law*, *JUSTICE SYSTEM JOURNAL*, Volume 33, Issue 2: Rule of Law (2012) at 136-7 (examining the variance amongst definitions and concepts captured within the rule of law) ], and what the rule of law entails [see Raz’s assertion that the rule of law “is not to be confused with democracy, justice, equality (before the law or otherwise), human rights of any kinds or respect for persons or for the dignity of man.” Raz, *supra* note xx at 211)] notion that a hierarchy of rule of law conceptualizations exist (or not) [See Brian Tamanaha, *ON THE RULE OF LAW; HISTORY, POLITICS AND THEORY*, Cambridge University Press: New York (2004); Lauth and Sehring (2009) and] There is almost no limit the vastness of debates within the rule of law. The author is not the first, to attempt to “elucidate the dominant understandings of the concept and to order these in a unified typology.” [see Svend-Erik Skaaning, above in note, at 137] Despite prior authors, and wide discussion in various and broad topics, there is, however, a clear concession to be made, the author has focused on very specific aspects of rule of law, to focus attention into a very narrow category- independence and impartiality of the decision maker. The narrow, often binary focused examination contained in the thesis, is a potentially, and likely is, a massive bias. The author support this limitations as not being enough to subject the scholarship to wider



for his book entitled On the Rule of Law; History, Politics and Theory<sup>32</sup> in which he, like others,<sup>33</sup> distinguish systematically between different definitions of the rule of law, focusing on the distinguishing characteristics between formal and substantive definitions of the rule of law.<sup>34</sup> He then creates, within each category continuum between ‘thin’ and ‘thick’ definitions.<sup>35</sup> The question arises, can the hierarchy and/or continuums be collapsed into a single dimension? While this approach has been viewed with scrutiny- and some disagree with this approach<sup>36</sup>- the outcomes had led some scholars to support the basic premise of ‘thin’ as formal and ‘thick’ as “substantive” and to be able to consider these divergent aspects in a single consideration, especially in some areas of consideration, like consent.<sup>37</sup> For example, as was explored earlier in the thesis as a mechanism of setting key attributes,<sup>38</sup> the World Justice Project has a definition of rule of law,<sup>39</sup> and then breaks this broad definitional set principles into a

---

scrutiny based in the widespread agreement amongst scholars, nations, and the global legal community that arbitration is a voluntary process of dispute resolution where the parties are empowered to select from various options and the nations are capable of reviewing arbitration through the process of recognition and enforcement, via the New York Convention. (see Chapter 2, Introduction) As such, the thesis has adopted the common approach to examining the argument based in crafting of law into code- the division between objective driven rules and procedures and subjective aspects of decision making, often occurring in the substantive portion of deliberation and decision making. [For a more full explanation of this distinction, *see* Kevin Ashley, *infra* Chapter 4(A) note 161 at 4 and corresponding footnotes.

<sup>32</sup> Brian Z. Tamanaha, *ON THE RULE OF LAW: HISTORY, POLITICS, THEORY*, Cambridge University Press (2004).

<sup>33</sup> *See* Møller, Systematizing Thin and Thick Conceptions, *supra* note 31 at 142.

<sup>34</sup> *See* Møller, Systematizing Thin and Thick Conceptions, *supra* note 31 at 4 (summarizing the academic field).

<sup>35</sup> For example, Tamanaha, *On Rule of Law*, *supra* note 31 at 91.

<sup>36</sup> *See* Møller, Systematizing Thin and Thick Conceptions, *supra* note 31 at 149.

<sup>37</sup> *See id.* at 149.

<sup>38</sup> *See infra*, Chapter 1(B)(Table A) and corresponding discussion.

<sup>39</sup> *See id.*

granular level of specific attributes and then creates measures in particular areas.<sup>40</sup> Within the project, conceptualizations of thick and thin are combined in each granular metric.<sup>41</sup> Similar to Tamanaha<sup>42</sup> and Jorgen Moller/Sven-Erik Skaaning<sup>43</sup> the World Justice projects creates a list of attributes to measure rule of law<sup>44</sup> and although those attributes are granular, they fit into a larger analysis to determine rule of law definitions and measures.<sup>45</sup> Similarly, both Tamanaha<sup>46</sup> and Bingham<sup>47</sup> argue for a hierarchy of attributes to be expected within the concept of the rule of law. The hierarchical approach is present in the thesis, yet, much has been eliminated or not directly addressed as the narrowing of the discussion to arbitration as an institution, builds in some of the necessary attributes, which are briefly explored but then not returned to as the focus of the thesis narrows, substantially.

For example, while many argument can be advanced and undoubtedly a much larger conversation should be undertaken about the importance of the rule of law, democratic process, the authority empowered to make such important rules, and the concept of the rule of law is a universal human good, as Chapter 2 and Chapter 3 explore in specific detail, there is no arbitration system, institution, or nation that fails to recognize the importance of the independence and impartiality of the ultimate decision maker, this is especially true in arbitration proceedings.<sup>48</sup> As delineated in Chapter 2, the global community has widely

---

<sup>40</sup> *See id.* World Justice Project “Access to justice is provided by competent, independent, and ethical adjudicators, attorneys or representatives, and judicial officers who are of sufficient number, have adequate resources, and reflect the makeup of the communities they serve.” *See id.*

<sup>41</sup> *See infra*, Chapter 1(B)(Table A) and corresponding discussion.

<sup>42</sup> *See* Tamanaha, *On Rule of Law*, *supra* note X at 91.

<sup>43</sup> *See* Moeller, *supra* note 31.

<sup>44</sup> *See* Chapter 1, Table B.

<sup>45</sup> *See* Moeller, *supra* note 31

<sup>46</sup> *See* Tamanaha, *On Rule of Law*, *supra* note 31 at 91.

<sup>47</sup> *See infra*, at pages 24-27 (examining Bingham, *supra* note 30 ) where the thesis author breaks down the attributes Bingham relies upon.

<sup>48</sup> *See* Chapter 2 (A); Chapter 3(B)(1).

supported arbitration and arbitration institutions as appropriate rule makers, so long as there is a series of checks and balances to ensure notions of law and public policy.<sup>49</sup> As arbitration institutions have been empowered with authority to craft such rules<sup>50</sup>- and arbitration and institutional selection occurs through a process of voluntary consent,<sup>51</sup> Chapter 2 specifically details the parameters of independence and impartiality of the decision maker in arbitration.<sup>52</sup> Chapter 3 uses the same analysis for the emerging Online Dispute Resolution (ODR) providers,<sup>53</sup> but criticizes some providers as departing from the establish expectations of arbitration as an alternative to traditional justice environment.<sup>54</sup> Thus, the question with the rule of law context, in light of the specific, narrow topic on the thesis, is what exactly do these two broad concepts mean- and how do they break down (or operationalize) in the context of a technology driven final decision maker located in a justice environment.<sup>55</sup>

The importance of this clarifying the topic, the crafting of definitions, the identification of key aspects of a topology and the objective measures to ensure a system is compliant with the rule of law in this narrow category is an essential aspect of furthering conversations about the use of technology as a decision maker in online justice environment. As Moeller and Svend-Erik Skaaning write when considering the importance of seeking definitions and fleshing out concepts prior to broader discussion:

Why is this exercise relevant? Seen from the higher ground, the problem pestering the rule-of-law research agenda is exactly that the term is often employed without justifying or even spelling out the definition and, *a fortiori*, without selecting empirical measures that match the intended definition. Such nonchalance is problematical because the establishment of a technical language based on sound logical premises is important for rigorous and cumulative research. In this connection, it is clearly worthwhile to remember Sartori's dictum that "concept formation stands

---

<sup>49</sup> *See infra*, page 42, including footnote.

<sup>50</sup> *See id.*

<sup>51</sup> *See id.*

<sup>52</sup> *See* Chapter 1, Table B at page 44.

<sup>53</sup> *See* Chapter 1, Table D at page 76 and Table F page 108; Chapter 3, Section B

<sup>54</sup> *See* Chapter 1, Table H at page 121 and text immediately following at pages 125-27.

<sup>55</sup> Specifically done within Chapter 5, Section 3.

prior to quantification." This, in turn, demands that the competing definitions are clarified and ordered...<sup>56</sup>

Although narrowed to a specific aspect of potentially a much larger topic, and undoubtedly a debate, the exercise of identifying the issue, finding various potential approaches, identifying commonality and specifying a definition, and then operationalizing of those aspects into the design process of the decision making portion of the ODR world, is essential first step in advancing broader conversations in the area of online justice.

For the purposes of this thesis the approach taken by Lord Bingham, in which he does in fact draw together the thick and thin aspects of rule of law, especially as it relates to the primary focus of the thesis, impartiality and independence of the decision maker, will be used.

According to Lord Bingham,

(1) The law must be accessible and, so far as possible, intelligible, clear and predictable.<sup>57</sup>

(2) Questions of legal right and liability should ordinarily be resolved by application of the law and not discretion.<sup>58</sup>

(3) The laws of the land should apply equally to all, save to the extent that differences justify differentiation.<sup>59</sup>

(4) Ministers and public officials at all levels must exercise the powers conferred to them in good faith, fairly, for the purpose for which the powers were conferred, without exceeding the limits of such powers and not unreasonably.<sup>60</sup>

(5) The law must offer adequate protection of human rights.<sup>61</sup>

---

<sup>56</sup> Møller, *Systematizing Thin and Thick Conceptions*, *supra* note 31 at 137.

<sup>57</sup> *Id.* at 37.

<sup>58</sup> *Id.* at 48.

<sup>59</sup> Bingham, *RULE OF LAW*, *supra* note 37, at 55.

<sup>60</sup> *Id.* at 60.

<sup>61</sup> *Id.* at 66.

(6) Means must be provided for resolving, without prohibitive cost or inordinate delay, bona fide civil disputes which the parties themselves are unable to resolve.<sup>62</sup>

(7) Adjudicated procedures provided by the state should be fair.<sup>63</sup>

(8) The law requires compliance by the state with its obligations in international law as in national law.<sup>64</sup>

To begin analyze these framing principles, Lord Bingham argues that:

all persons and authorities within the state, whether public or private, should be bound by and entitled to the benefits of laws publicly made, taking effect (generally) in the future and publicly administered in the courts.<sup>65</sup>

Lord Bingham demands that “any departure from the rule I have stated calls for close consideration and clear justification.”<sup>66</sup> Lord Bingham asks policymakers to understand that “belief in the rule of law does not import unqualified admiration of the law, or the legal profession, or the courts, or the judges.”<sup>67</sup> His principles do not a call for rigid and blind adherence, but rather should prompt *robust consideration of any departures*.

Note that an online justice environment appears to stand at direct odds with the principle that justice must be *publicly* administered by the *courts*.<sup>68</sup> In regard to the courts aspect, ODR’s proponents have benefited from years of support for arbitration as an alternative to the traditional brick-and-mortar justice environment, as will be explored in detail in Chapter 2. Arbitration in general is already established as an acceptable *alternative* to the court system. In other words it presents a departure from Lord Bingham’s rules deserving of scrutiny,<sup>69</sup> but one that has clear justification and years of well-founded support. When considering the “public” expectation, we must

---

<sup>62</sup> *Id.* at 85.

<sup>63</sup> *Id.* at 90.

<sup>64</sup> Bingham, RULE OF LAW, *supra* note 37, at 110.

<sup>65</sup> *Id.* at 8.

<sup>66</sup> *Id.*

<sup>67</sup> *Id.* at 9.

<sup>68</sup> See Paul Magrath, *Online Courts And Cyber Judges*, InfoLaw Newsletter, (March 2017) (calling into question the ability of online courts to meet the standard of public) available at <https://www.infolaw.co.uk/newsletter/2017/03/online-courts-cyber-judges/> (last visited Aug. 5, 2019).

<sup>69</sup> See Bingham, RULE OF LAW, *supra* note 37, at 9.

refine the definition. There is growing support for a view that successful RoL requires “the cooperation of state and society, and is an *outcome* of complex and deeply rooted social processes.”<sup>70</sup> If the purpose of the expectation that dispute resolution be public is to support an informed citizenry able to engage in debates about the social construction of justice and the parameters of such conversation, then we can reasonably unpack the definition of “public” to imply an expectation of transparency, of both process and outcome. If *public* necessarily implies that an individual can physically be present in a court room, then online justice faces a fundamental problem. Assuming, however, that the more substantive definition is correct, transparency *is* possible in the online justice eco-system.<sup>71</sup>

Continuing his argument, Lord Bingham, building upon the work of Sir Matthew Hale, addresses the importance of judges and of the manner in which they conduct themselves.<sup>72</sup> For his part, Sir Hale addresses the responsibility of the decision maker to ensure that their own conduct is in compliance with the rule of law, laying down the expectation: “[t]hat in the execution of justice, I [the judge] carefully lay aside my own passions, and not give way to them however provoked.”<sup>73</sup> We call this today the expectation of judicial impartiality. Hale also specifies, “[t]hat I never engage myself in the beginning of any cause, but reserve myself unprejudiced till the whole be heard.”<sup>74</sup> Thus, the conduct expected of the decision maker is to be impartial, independent, and to act “irrespective of public opinion.”<sup>75</sup> Lord Bingham emphasizes that “the judges’ role in maintaining the rule of law is crucial . . .”<sup>76</sup>

---

<sup>70</sup> Michael Johnston, *Good Governance: Rule of Law, Transparency, and Accountability*, UNITED NATIONS PUBLIC ADMINISTRATION NETWORK, available at <http://unpan1.un.org/intradoc/groups/public/documents/un/unpan010193.pdf> (last visited Aug. 5, 2019).

<sup>71</sup> See discussion, *infra* Chapter 4, Section C.

<sup>72</sup> See Bingham, RULE OF LAW, *supra* note 37, at 21

<sup>73</sup> Bingham, RULE OF LAW, *supra* note 37, at 20 *citing* Sir Hale, THINGS NECESSARY TO BE CONTINUALLY HAD IN REMEMBRANCE.

<sup>74</sup> Bingham, RULE OF LAW, *supra* note 37, at 21.

<sup>75</sup> *Id.* at 21.

<sup>76</sup> *Id.* at 22.

These issues form the backbone of this thesis: the role of the decision maker, the influence that is to occur in making the decisions, and the accountability of the decision maker. They will be returned to continually throughout the thesis. As a short preview, it is important to note that technology can, and perhaps should, be allowed to enhance the trustworthiness of the justice eco-system through careful, well-governed, deployments. As will be explored more in Chapter 4, this is type of governance is possible, and can be achieved as outlined in Chapter 5.<sup>77</sup>

Third, Lord Bingham believes that:

Judges do have a role in developing the law, and the common law has grown up as a result of their doing just this. But, and this is an all-important condition, there are limits.<sup>78</sup>

He goes on to explore the line between judicial discretion and judicial activism. He warns against exercising discretion in “radically innovative or adventurous ways”<sup>79</sup> to the point that the law becomes “uncertain and unpredictable.”<sup>80</sup>

[T]he rule of law does not require that official or judicial decision makers should be deprived of all discretion, but it does require that no discretion should be unconstrained so as to be potentially arbitrary.<sup>81</sup>

To prevent chaos, “rules have grown up to direct the exercise of this discretion.”<sup>82</sup> Thus, decision makers are guided by the law, bounded by the rules, and are directed in the application of rules that govern appropriate moments for discretion. This crucial balance of stricture and discretion is all the more important to achieve when undertaking departures from the ordinary rule of law.

As will be discussed in much greater detail in Chapters 3 and 4, the ODR environment will be based upon a carefully crafted set of legal rules. However, human-based discretion is almost impossible to capture in this format, and currently the judiciary relies heavily upon discretion to address situations that fail to be fully captured in the legal rules. Thus, *when* discretion will be allowed and *how* it will operate are two

---

<sup>77</sup> See discussion, *infra* Chapter 4, Section B and Chapter 5, Section C.

<sup>78</sup> Bingham, RULE OF LAW, *supra* note 37, at 45.

<sup>79</sup> *Id.* at 46.

<sup>80</sup> *Id.* at 46.

<sup>81</sup> *Id.* at 54.

<sup>82</sup> Bingham, RULE OF LAW, *supra* note 37, at 52.

important questions for the designers of ODR. Chapter 5 will outline a potential approach to developing a form of judicial discretion in ODR.

Finally, Lord Bingham argues:

Any departure from the general rule of equal treatment should be scrutinized to ensure that the differential treatment is based on real differences.<sup>83</sup>

Thus, differential treatment is allowed, so long as it is scrutinized to ensure that real differences exist, and are relevant in the context of the question being asked. This aspect of rule of law, appropriately differentiated treatment, has proven incredibly controversial in the ODR community. Why? Much of the ODR community subscribes to the belief that efficiency is the primary value of ODR, and this type of deliberative reasoning is simply not as streamlined as the simpler decision-trees that make no allowance for scrutiny.<sup>84</sup> Some argue that efficiency is such an essential consideration that other aspects of rule of law can and should be adjusted in order to prioritize it.<sup>85</sup> Sophie Nappert demands ODR policy makers consider if “technological development entails downscaling arbitral due process in favor of time and cost-saving measures.”<sup>86</sup> This has emerged as one of the most difficult, and important, questions within the ODR community. Currently, many argue that low-cost disputes can and should be decided through an automated process,<sup>87</sup> one that includes no human decision maker, while maintaining that those with high-value disputes should remain entitled to a human decision maker. The question facing ODR policy makers in this instance thus becomes:

---

<sup>83</sup> *Id.* at 56.

<sup>84</sup> See Anjanette H Raymond and Scott Shackelford, *Technology, Ethics And Access To Justice: Should An Algorithm Be Deciding Your Case?* 35 MICHIGAN JOURNAL OF INTERNATIONAL LAW 485, 516-7 (2014) (arguing for that a Balance Must be Struck between Efficiency and Justice).

<sup>85</sup> *See id.*

<sup>86</sup> Sophie Nappert, *Game changers: Defining and Defending Values in International Arbitration*, 2019 International Arbitration Dispute Resolution Symposium, Washington University, Keynote Presentation, Video, (March 1, 2019) available at <https://lecturecapture.wustl.edu/mediasite/Play/43893c65920f46b990595d877f55487d1d> (last visited Aug. 14, 2019).

<sup>87</sup> For example, in the case of blind bidding negotiation, such as CyberSettle. *See discussion, infra* Chapter 2, at Table D.



should the total value of the recovery sought determine which type of justice the plaintiff has access to?

Advocates of efficiency in the ODR community are emboldened in their argument by a significant and immediate need to improve access to justice around the world.<sup>88</sup> Nonetheless, this thesis joins others in the ODR community<sup>89</sup> in arguing the position that the several protections, responsibilities, and commitments embedded in the rule of law are crucial to the relationship between government and citizens, and that efficiency cannot be gained at the expense of damage to these aspects of rule of law. As the Chief Justice Beverley McLachlin remarked as she reflected upon the *Relevance of Magna Carta in Today's World*:<sup>90</sup>

But the rule of law embedded in the Magna Carta is more than a mere demand for legal primacy. The document supplements it with an insistence that the law must be just, available and free from corruption. . . . The twin notions of the primacy of law and the effectiveness of the law — an insistence that the law be just, impartial and available to all — combine to create what we call the rule of law. Both are essential to the rule of law. If the law is subordinated to other forms of power, there is no rule of law. Nor can the rule of law survive if the law is partial, corrupt or inaccessible to citizens. The Magna Carta contains the germs of both

---

<sup>88</sup> See e.g., Daniel Rainey, *Access to Justice and ODR*, National Center for Technology and Dispute Resolution, (March 14, 2019)(asserting “If we really want to increase access to justice, we need to target not just easy access to the courts for traditional parties, but apps and education for those who have traditionally been voluntarily disenfranchised by their own fear or reluctance to approach a court, online or offline.”); Daniel Becker, Andrea Maia, *ODR as an Effective Method to Ensure Access to Justice: The Worrying, But Promising Brazilian Case*, MEDIATE.COM, (Sept. 2019) (arguing “even though “access to courts” was granted by the Brazilian justice system, it failed to grant “access to justice” to Brazilian citizens. . . . “ODR methods will undoubtedly serve to improve access to justice since these methods are capable of reducing the judicialization of ordinary conflicts of a simpler nature.”) *Id.* available at <https://www.mediate.com/articles/becker-odr-effective.cfm> (last visited Aug. 14, 2019). See also, World Justice Project, *WJP Rule of Law Index*, (2019), Table A, *infra* Section 1, Subsection B.

<sup>89</sup> See e.g., Julia Hornle: *CROSS BORDER INTERNET DISPUTE RESOLUTION*, 256, Cambridge University Press (2009) (asserting “the main concern of this book is the fair resolution of Internet disputes, and it has been argued that, in order to achieve this, minimum due process standards, set at a higher level than those pertaining to traditional arbitration, should be incorporated into all online arbitration procedures.”) *Id.* at 256.

<sup>90</sup> McLachlin, *Relevance of Magna Carta in Today's World*, *supra* note 3.

these ideas — ideas that were developed in later centuries to produce the modern concept of the rule of law.<sup>91</sup>

Tracing its origins to the Magna Carta, today's modern concept of the rule of law may be broken into two important sub-concepts: the primacy of the law and the effectiveness of the law. Both are essential to our discussion, but it is the second notion, effectiveness, understood as an “insistence that the law be just, impartial, and available to all”<sup>92</sup> which most concerns us. As noted by United Nations in its *Declaration Of The High-Level Meeting Of The General Assembly On The Rule Of Law At The National And International Levels*:

We are convinced that the independence of the judicial system, together with its impartiality and integrity, is an essential prerequisite for upholding the rule of law and ensuring that there is no discrimination in the administration of justice.<sup>93</sup>

This thesis seeks to ensure the principles of impartiality and independence are embedded into the design of any decision maker within the online justice environment.

## B. The Rule of Law: Measuring the Rule of Law

As explained by Christopher Stephens,

Over the course of the 800 years since the Magna Carta, organisations . . . institutes, universities, think tanks, NGOs, and law societies have, through a relentless focus on the rule of law, raised the concept to a normative principle that is almost universally recognised as an aspirational premise of civil society.<sup>94</sup>

However, determining by what metric to measure the success of rule of law is no less challenging than distilling a precise definition of the rule of law. These are incredibly important challenges, as “the rule of law has become a globally recognised normative

---

<sup>91</sup> *Id.* at 12.

<sup>92</sup> *Id.*

<sup>93</sup> United Nations General Assembly, *Declaration Of The High-Level Meeting Of The General Assembly On The Rule Of Law At The National And International Levels*, A/Res 67/1 (Nov. 24, 2012) at para 13 available at: [http://www.un.org/en/ga/search/view\\_doc.asp?symbol=A/RES/67/1](http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/67/1) (last visited Aug. 14, 2019).

<sup>94</sup> Christopher Stephens, *The Rule Of Law In Development*, at 29, in Jeffrey Jowell, J Christopher Thomas, Jan van Zyl Smit, (eds), *RULE OF LAW SYMPOSIUM (2014): THE IMPORTANCE OF THE RULE OF LAW IN PROMOTING DEVELOPMENT* (2014).

value among nations.”<sup>95</sup> The issue is even more complex; however, than it may first appear, as explained by Martin Krygier:

[R]ule of law is not something you either have or not . . . rather, like wealth, one has more or less of it. Whether one has enough of it is a judgment to be made along continua — multiple continua — not a choice between binary alternatives.<sup>96</sup>

Overly simple measures, especially those that focus merely on institutions and not on outcomes and processes,<sup>97</sup> are inadequate to measure meaningful compliance with the rule of law. Several institutions have attempted to capture the complexity in indices designed to evaluate individual nations’ justice systems and note areas of needed improvement. There are three widely regarded rule of law indices promulgated, respectively, by: the United Nations, the World Bank, and the World Justice Project.

First, the United Nations attempts to set the aspirational definition of the rule of law as a:

[p]rinciple of governance in which all persons, institutions and entities, public and private, including the State itself, are accountable to laws that are publicly promulgated, equally enforced and independently adjudicated, and which are consistent with international human rights norms and standards. It requires, as well, measures to ensure adherence to the principles of supremacy of law, equality before the law, accountability to the law, fairness in the application of the law, separation of powers, participation in decision-making, legal certainty, avoidance of arbitrariness and procedural and legal transparency.<sup>98</sup>

This definition is further developed into a detailed rubric. To measure a given state’s compliance with the rule of law, the United Nations:

...assesses the functioning of the police (41 indicators), judiciary (51 indicators), prisons (43 indicators) by looking at the following elements:

---

<sup>95</sup> *Id.* at 31.

<sup>96</sup> Martin Krygier, *False Dichotomies, Real Perplexities and the Rule of Law*, at 261, in: András Sajó, (ed.) *HUMAN RIGHTS WITH MODESTY: THE PROBLEM OF UNIVERSALISM*, (Leiden/Boston: Martinus Nijhoff, 2004).

<sup>97</sup> See Rachel Kleinfeld Belton, *Competing Definitions of the Rule of Law: Implications for Practitioners*, 27 (Carnegie Papers no. 55, January 2005), available at: <http://carnegieendowment.org/files/CP55.Belton.FINAL.pdf> (last visited Aug. 14, 2019).

<sup>98</sup> United Nations, *Rule of Law Indicators: Implementations Guide and Project Tools* (2011) at pages v. and vi. available at: [http://www.un.org/en/events/peacekeepersday/2011/publications/un\\_rule\\_of\\_law\\_indicators.pdf](http://www.un.org/en/events/peacekeepersday/2011/publications/un_rule_of_law_indicators.pdf) (last visited Aug. 5, 2019).

(i) performance, (ii) integrity, transparency and accountability, (iii) treatment of vulnerable groups and (iv) capacity.<sup>99</sup>

However, the United Nations measures are narrowly focused on justice in criminal cases. As such, they fail to be fully useful in the instance of a justice system designed to handle a wider array of cases. As our focus is primarily on civil dispute resolution, it is not an adequate tool for the purposes of this thesis.

A second widely regarded mechanism designed to measure compliance with the rule of law has been established by the World Bank. It sets its aspirational definition of the rule of law as:

Rule of law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular [captures] the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence.<sup>100</sup>

The World Bank mechanism for measuring compliance is positioned within a larger set of benchmarks for governance, the World Governance Indicators Project.<sup>101</sup> The limitation of the World Bank's measure for rule of law is that it is focused on external perceptions of the trustworthiness of a given justice system. This captures indicators that may be relevant to many considerations, for example, corporate decisions to enter foreign business environments, but for the purposes of this thesis it again too narrowly focused.

Finally, the World Justice Project (WJP) is an independent nonprofit organization whose mission is to “develop communities of opportunity and equity by advancing the rule of law worldwide.”<sup>102</sup> One of the major initiatives of the WJP is the Rule of Law

---

<sup>99</sup> United Nations, *Rule of Law Indicators*, *supra* note 60, at 4.

<sup>100</sup> See World Bank, *World Governance Indicators, Rule of Law*, available at <http://info.worldbank.org/governance/wgi/index.aspx#doc> (last visited Aug. 14, 2019) discussed in detail in Daniel Kaufmann Aart Kraay Massimo Mastruzzi, *The Worldwide Governance Indicators Methodology and Analytical Issues*, Policy Research Working Paper 5430 at 4, (Sept. 2010).

<sup>101</sup> Kaufmann, *The Worldwide Governance Indicators*, *supra* note 62.

<sup>102</sup> See World Justice Project, *WJP Rule of Law Index*, (2019) available at <http://data.worldjusticeproject.org/> (last visited Aug. 5, 2019). Our focus here is on the indicators that measure: accessibility and cost; the absence of unreasonable delays; and ADR, because these factors demonstrate a global need to begin to consider a greater use of ODR.

Index (2019), which is a quantitative assessment tool that offers a detailed and comprehensive picture of the extent to which countries adhere to the rule of law in practice.<sup>103</sup> It contains a larger data set, more nuance in the various categories than other indices, and specifically addresses civil justice. As such, the index provides the most suitable framework for our discussion.<sup>104</sup>

The civil justice portion of the WJP index measures seven key factors:

- people can access and afford civil justice
- civil justice is free of discrimination
- civil justice is free of corruption
- civil justice is free of improper government influence
- civil justice is not subject to unreasonable delays
- civil justice is effectively enforced
- alternative dispute resolution mechanisms (ADRs) are accessible, impartial, and effective.<sup>105</sup>

The WJP asserts that access to civil justice “requires that the system be accessible, affordable, effective, impartial, and culturally competent.”<sup>106</sup>

---

<sup>103</sup> See World Justice Project, *WJP Rule of Law Index, basic information page*, (2019) available at <https://worldjusticeproject.org/our-work/wjp-rule-law-index> (last visited Aug. 5, 2019).

<sup>104</sup> See Mark David Agrast, Juan Carlos Botero, and Alejandro Ponce, *The World Justice Project, World Justice Project Rule of Law Index 2011*, 117 (2011), available at [http://worldjusticeproject.org/sites/default/files/WJP\\_Rule\\_of\\_Law\\_Index\\_2011\\_Report.pdf](http://worldjusticeproject.org/sites/default/files/WJP_Rule_of_Law_Index_2011_Report.pdf) (last visited Aug. 5, 2019).

<sup>105</sup> See *WJP Rule of Law Index*, (2019) *supra* note 61, available at <http://data.worldjusticeproject.org/#table> (last visited Aug. 5, 2019). The table includes the WJP indicators that measure: accessibility and cost; the absence of unreasonable delays; and ADR, because these factors demonstrate a global need to begin to consider a greater use of ODR. *See id.*

<sup>106</sup> *Id.* According to Professor Mauro Cappelletti, there are three main obstacles that make civil and political liberties inaccessible in many parts of the world. First, due to economic reasons, individuals are unable to access information or adequate representation. Second, due to organizational obstacles, the isolated individual lacks sufficient motivation, power, and information to initiate and pursue litigation. Third, sometimes procedural processes are inadequate, that is, traditional contentious litigation in court might not be the best possible way to provide effective vindication rights for many individuals. *See* Mauro Cappelletti, *Alternative Dispute Resolution Processes with the Framework of the Worldwide Access to Justice Movement*, 56 MOD. L. REV. 282, 283 (1993).

I have produced an edited summary of WJP scores for several countries, selecting those measures most relevant to the points discussed in this thesis. Note that these data represent only a portion of the data publicly available on the WJP website.

*Table A: WJP Rule of Law Index*<sup>107</sup>

	Civil Justice Overall	7.1 People can access and afford	7.2 Free of discrimination	7.3 Free of corruption	7.4 Free of improper government influence	7.5 Not subject to unreasonable delays	7.6 Effectively enforced	7.7 ADRs are accessible, impartial, and effective
<b>Australia</b>	0.76	0.62	0.70	0.87	0.90	0.68	0.76	0.83
<b>Canada</b>	0.70	0.57	0.58	0.90	0.89	0.47	0.73	0.78
<b>China</b>	0.54	0.66	0.49	0.43	0.22	0.76	0.57	0.67
<b>Germany</b>	0.86	0.78	0.85	0.88	0.91	0.83	0.90	0.86
<b>India</b>	0.45	0.40	0.44	0.50	0.64	0.18	0.39	0.58
<b>Italy</b>	0.56	0.62	0.57	0.67	0.69	0.31	0.37	0.69
<b>Japan</b>	0.79	0.70	0.80	0.94	0.77	0.71	0.75	0.88
<b>Mexico</b>	0.40	0.43	0.33	0.34	0.47	0.28	0.40	0.56
<b>Norway</b>	0.85	0.71	0.68	0.96	0.93	0.83	0.91	0.96
<b>Russia</b>	0.52	0.58	0.52	0.54	0.32	0.70	0.42	0.55
<b>South Africa</b>	0.59	0.49	0.46	0.66	0.65	0.51	0.60	0.75

<sup>107</sup> World Justice Project, *WJP Rule of Law Index*, (2019) *supra* note 51, available at, <http://data.worldjusticeproject.org/#table> (last visited Aug. 5, 2019). A prior (2011) version of this chart appeared in my publication: Anjanette H Raymond and Scott Shackelford, *Technology, Ethics And Access To Justice*, *supra* note 58 at 516-7 (Table at 489).

<b>United Kingdom</b>	0.73	0.53	0.64	0.91	0.84	0.67	0.72	0.78
<b>United States</b>	0.64	0.46	0.42	0.82	0.71	0.62	0.68	0.77

Source: World Justice Project (2019)

The WJP measures offer insight into the existing justice environment around the world. The data points highlight the extent to which the United States, and many other nations, are lacking justice environments that are widely accessible. Also notable is the absence of ADR in many of these countries.

This coupling of low levels of ADR with high rate of access-to-justice issues points toward the nuanced importance of carefully developing new dispute resolution mechanisms. On the one hand, these scores suggest that an opportunity exists for justice systems to use ADR more robustly to supplement brick-and-mortar courtrooms with alternate avenues of access to justice. It also indicates that countries which currently score low on indices of rule of law have an incentive to portray ADR as constituting adequate access, in order to improve their ranking. Finally, overall, the WJP table illustrates the importance of continued justice system improvement. No system has a claim to civil justice excellence based on its score. The table reveals no system should be comfortable with its rating and all systems would benefit from changes designed to improve both compliance with the rule of law and access to justice.

### C. The Rule of Law: An Independent Judiciary

When considering the creation of an online justice environment, especially when considering potential issues of bias, dependency, and partiality, great consideration must be given to protecting the basic principles clearly embedded within the rule of law. As articulated by Raz, the rule of law requires that:

The independence of the judiciary must be guaranteed. . . The rules concerning the independence of the judiciary—the method of appointing judges, their security of tenure, the way of fixing their salaries, and other conditions of service—are designed to guarantee that they will be free from extraneous pressures and independent of all



authority save that of the law. They are, therefore, essential for the preservation of the rule of law.<sup>108</sup>

From this we can see that ensuring independence of the judiciary is a complex aspect of ensuring the rule of law.

Few international institutions proscribe the method of judicial appointment, but most international bodies consider appointment to be an essential consideration when discussing independence of the judiciary. Under the Basic Principles on the Independence of the Judiciary,<sup>109</sup>

Persons selected for judicial office shall be individuals of integrity and ability with appropriate training or qualifications in law. Any method of judicial selection shall safeguard against judicial appointments for improper motives. In the selection of judges, there shall be no discrimination against a person on the grounds of race, colour, sex, religion, political or other opinion, national or social origin, property, birth or status, except that a requirement, that a candidate for judicial office must be a national of the country concerned, shall not be considered discriminatory.<sup>110</sup>

This passage draws attention to both the manner of selection and the individual qualifications to be expected in a decision maker.

The European Convention on Human Rights prescribes an: “independent and impartial tribunal established by law.”<sup>111</sup> Interpreting this phrase (and others) the European Court of Human Rights has consistently held that:

[I]n order to establish whether a tribunal can be considered “independent” for the purposes of article 6 § 1, regard must be had to the manner of appointment of its members and their term of office, the

---

<sup>108</sup> Raz, *THE AUTHORITY OF LAW*, *supra* note 25, at 217.

<sup>109</sup> United Nations Human Rights, Office of the High Commissioner, United Nations Congress on the Prevention of Crime and the Treatment of Offenders, *Basic Principles on the Independence of the Judiciary*, General Assembly resolutions 40/32 of 29 November 1985 and 40/146 of 13 December 1985, *available at* <http://www.ohchr.org/EN/ProfessionalInterest/Pages/IndependenceJudiciary.aspx> (last visited Aug. 14, 2019).

<sup>110</sup> United Nations Human Rights, *Basic Principles on the Independence of the Judiciary*, *supra* note 57 at Art. 10.

<sup>111</sup> Council of Europe, CONVENTION FOR THE PROTECTION OF HUMAN RIGHTS AND FUNDAMENTAL FREEDOMS, Rome, 4.XI.1950, Art. 6(1).



existence of safeguards against outside pressures and the question whether it presents an appearance of independence.<sup>112</sup>

This interpretation adds to manner of selection and candidate qualifications the additional criteria of safeguards against “outside pressure” and, interestingly, the “appearance of independence”. The latter consideration implies that transparency and public trust are important aspects.

When it comes to “outside pressure,” a variety of documents have defined the term expansively, and outlined several important safeguards against it. For example the Inter-American Commission on Human Rights recommends that Member States

...take the steps necessary to protect the integrity and independence of members of the Judiciary in the performance of their judicial functions, and specifically in relation to the processing of human rights violations; in particular, judges must be free to decide matters before them without any influence, inducements, pressures, threats or interferences, direct or indirect, for any reason or from any quarter.<sup>113</sup>

Terms of judicial employment may be considered important to preventing outside influence. The Basic Principles on the Independence of the Judiciary include the provision:

The term of office of judges, their independence, security, adequate remuneration, conditions of service, pensions and the age of retirement shall be adequately secured by law.<sup>114</sup>

Thus not only the initial selection but the ongoing terms of judicial employment, for example security of tenure, are considered paramount to protecting judicial independence.

And finally, many systems also prevent outside influence being placed upon the judiciary by requiring any discipline or other intervention taken against a member of the judiciary to occur through an independent panel that stands outside the sphere of

---

<sup>112</sup> European Court of Human Rights, *Case of Incal v. Turkey*, (9 June 1998), Reports 1998-IV, p. 1571, para. 65.

<sup>113</sup> Organization of American States, *Annual Report of the Inter-American Commission on Human Rights*, OAS doc. OEA/Ser.L/V/II.95, doc. 7 rev., (1996) at 761.

<sup>114</sup> United Nations Human Rights, *Basic Principles on the Independence of the Judiciary*, *supra* note 57 at Art. 12.

the potential influence of the judge. For example, the Council of Europe Recommendations specify that “all decisions concerning the professional career of judges should be based on objective criteria”<sup>115</sup> and “should be based on merit, having regard to qualifications, integrity, ability and efficiency.”<sup>116</sup> Moreover, “In the decision-making process, judges should be independent and be able to act without any restriction, improper influence, inducements, pressures, threats or interferences, direct or indirect, from any quarter or for any reason.”<sup>117</sup>

Collectively, these sources tell us that in order to fulfill the rule of law requirement of an independent judiciary, the online justice environment should be designed to ensure: (1) qualified decision makers; (2) whose appointment is free of influence; (3) whose tenure and salary are fixed by external agents, without reward or consequence based on outcomes; (3) further, that steps are built into the system to monitor potential outside influence and protect decision makers, while also protecting the overall integrity of the system; and (4) that the manner of selection, the rules that govern conduct and conflict of interest, and remuneration, are all publicly available.

#### D. The Rule of Law: An Impartial Judiciary

For rule of law to be sound, the judiciary must be not only independent but impartial. In general, impartiality “requires that an arbitrator neither favors one party nor is predisposed as to the question in dispute.”<sup>118</sup> As such, it is a state of mind or attitude of the tribunal in relation to the issues and the parties in a particular case.”<sup>119</sup> As Lord Bingham explains:

---

<sup>115</sup> Council of Europe, Committee of Ministers, Recommendation No. R (94) 12 Of The Committee Of Ministers To Member States on *The Independence, Efficiency And Role Of Judges*, Principle I - General Principles On The Independence Of Judge, at c.

<sup>116</sup> Council of Europe, *Role of Judges*, *supra* note 88, General Principles On The Independence Of Judge, at c.

<sup>117</sup> Council of Europe, *Role of Judges*, *supra* note 88, General Principles On The Independence Of Judge, at d.

<sup>118</sup> Alan Redfern and Martin Hunter, with N. Blackaby and C. Partasides, *LAW AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION*, para 4-51, (4th Ed.), (2004).

<sup>119</sup> *Id.* at 2.

[R]ules of natural justice have traditionally been held to demand, first, that the mind of the decision maker should not be tainted by bias or personal interest. . . .<sup>120</sup>

Raz expands upon this expectation, placing it within the context of correct application of the law, declaring:

The principles of natural justice must be observed. Open and fair hearing, absence of bias, and the like are obviously essential for the correct application of the law.<sup>121</sup>

In the Declaration of the High-Level Meeting on the Rule of Law,<sup>122</sup> Member States highlighted the independence of the judicial system, alongside impartiality and integrity, as an essential prerequisite for upholding the rule of law and ensuring that there is no discrimination in the administration of justice.<sup>123</sup>

Linda G. Mills suggests that impartial decision making in practice has a profound relationship with bias, arguing:

The doctrinal conception of bias is based on an opposition between bias and impartiality, a rejection of one for the other. Yet my findings and the studies that preceded them indicate that in practice, bias is a consistent dimension of what is considered “impartial” decision making.<sup>124</sup>

Within the context of this thesis, impartiality and bias will be interpreted in a contextual manner, rejecting the treatment of the terms as interchangeable. Although Lord Bingham and others use the term *bias* as almost interchangeable with *partiality*,<sup>125</sup> the term is being reviewed by these authors in the context of (necessarily human) decision makers in the brick-and-mortar judiciary, whereas this thesis also considers them in the context of technology-driven decision making. Hence, in the context of this thesis,

---

<sup>120</sup> Bingham, RULE OF LAW, *supra* note 37, at 62.

<sup>121</sup> Raz, THE AUTHORITY OF LAW, *supra* note 25 at 217.

<sup>122</sup> United Nations General Assembly, *Declaration of the High-level Meeting of the 67th Session of the General Assembly On The Rule Of Law At The National And International Levels*, A/RES/67/1, (Nov. 2012).

<sup>123</sup> *See id.* at para. 13.

<sup>124</sup> Linda G. Mills, *A Penchant for Prejudice: Unraveling Bias in Judicial Decision-Making*, 13 UNIVERSITY OF MICHIGAN PRESS, (1999).

<sup>125</sup> *See* Bingham, RULE OF LAW, *supra* note 37 at 93 (discussing impartiality and demanding decision makers “alert themselves’ to “any extraneous circumstance which might bias their judgement.”) *Id.* at 93.

bias and partiality are not interchangeable. They are two separate considerations within the technology-driven environment.

In the case of impartiality, this thesis will adopt the widest possible definition of impartiality. Philosophers often frame discussion of impartiality by writing about an impartial *point of view*, but this thesis takes the view that within the area of judicial decision making impartiality is better examined in terms of the impartial *agent* or *observer*. An impartial observer is a person who makes moral judgments without being influenced by the sort of contaminating biases or prejudices that tend to arise from the occupation of a particular point of view.<sup>126</sup> Some authors, such as Firth, suggest that the ideal observer is both “disinterested,” meaning that the decision maker is “entirely lacking in particular interests,”<sup>127</sup> and “dispassionate,” in that she is “incapable of experiencing any emotions at all.”<sup>128</sup> Some commentators demand that the ideal impartial observer must embody practical wisdom, in the Aristotelian sense.<sup>129</sup>

It is possible that John Stuart Mills captures the relevant features of impartiality best:

Impartiality, in short, as an obligation of justice, may be said to mean, being exclusively influenced by the considerations which it is supposed ought to influence the particular case in hand; and resisting the solicitation of any motives which prompt to conduct different from what those considerations would dictate.<sup>130</sup>

---

<sup>126</sup> See e.g., Adam Smith, *Theory of the Moral Sentiments*, Oxford: Oxford University Press.(1976); David Hume, A TREATISE OF HUMAN NATURE, Second edition, ed. L.A. Selby-Bigge & P.H. Nidditch, Oxford University Press (1978); Roderick Firth, *Ethical Absolutism and the Ideal Observer*, PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH, 12(3): 317–345 (1952); Richard Brandt, *The Definition of an ‘Ideal Observer’ in Ethics*, PHILOSOPHY AND PHENOMENOLOGICAL RESEARCH, 15, 407-13 (1954).

<sup>127</sup> Firth, *Ethical Absolutism and the Ideal Observer*, *supra* note 98 at 335.

<sup>128</sup> *Id.* at 340-41.

<sup>129</sup> Practical wisdom “is the ability to do the right thing, at the right time, for the right reason.” John Bradshaw, RECLAIMING VIRTUE: HOW WE CAN DEVELOP THE MORAL INTELLIGENCE TO DO THE RIGHT THING AT THE RIGHT TIME FOR THE RIGHT REASON, Bantam (April 28, 2009).

<sup>130</sup> J.S. Mill, *Utilitarianism*, In *On Liberty and Utilitarianism*, (1992) in Knopf: EVERYMAN'S LIBRARY, Volume 81. p. 154. See also Firth, *Ethical Absolutism and the Ideal Observer*, *supra* note at 336.

In the United States, judicial impartiality is most often characterized as the absence of judicial bias. Within this context, judicial bias involves positively or negatively prejudiced “feelings or spirit”<sup>131</sup> toward the claimants in the cases being heard. According to prevailing U.S. judicial doctrinal conceptions, to apply the standard of judicial impartiality, judges must eliminate both any “hostile feeling or spirit of ill will”<sup>132</sup> and “undue friendship or favoritism”<sup>133</sup> toward any litigant that appears before them.<sup>134</sup> It is an expectation, in short, about specific feelings that judges are obligated by the ideal of judicial impartiality to expunge from their reasoning and decisions.<sup>135</sup>

Judicial bias jeopardizes the individual’s right to a fair trial under rule of law. The U.S. Supreme Court has specified: “The alleged bias and prejudice to be disqualifying must stem from an extrajudicial source and result in an opinion on the merits on some basis other than what the judge learned from his participation in the case.”<sup>136</sup> It further defines fairness (within the context of a fair tribunal) as an “absence of actual bias.”<sup>137</sup> Consequently, the presence of actual bias, that is external events that are taken into consideration and that impact the outcome, is the standard measure which gives rise to questions implicating a fair trial in the United States.

The European Court of Human Rights has articulated a two-pronged measure of impartiality, dividing the considerations into: *subjective* and *objective*. “First, the

---

<sup>131</sup> 46 AMERICAN JURISPRUDENCE 2d § 147.

<sup>132</sup> *Id.* at § 167.

<sup>133</sup> *Id.*

<sup>134</sup> *See id.*

<sup>135</sup> *See id.* § 147.

<sup>136</sup> *United States v. Grinnell Corp.*, 384 U.S. 563, 583 (1966). *See also Berger v. U.S.* 255 U.S. 22, 31 (1921) (“the reasons and facts for the belief the litigant entertains are an essential part of the affidavit, and must give fair support to the charge of a bent of mind that may prevent or impede impartiality of judgment.”) *Id.* at 255; *Liteky et al. v. U.S.* 510 U.S. 540 (1994)(determining recusal of a judge in a federal criminal case is subject to the limitation that has come to be known as the “extrajudicial source” doctrine which is measured by the judge: (1) relied upon knowledge acquired outside such proceedings, and/or (2) displayed deep-seated and unequivocal antagonism that would render fair judgment impossible.) *Id.* at 556.

<sup>137</sup> *In re Murchison*, 349 U.S. 133, 136 (1955)( “A fair trial .in a fair tribunal is a basic requirement of due process. Fairness, of course, requires an absence of actual bias in the trial of cases.”) *Id.* at 136.

tribunal must be subjectively free of personal prejudice or bias.”<sup>138</sup> In the Case of *Daktaras* the ECHR articulated, “Personal impartiality is presumed unless there is evidence to the contrary.”<sup>139</sup> In addition, “[the tribunal] must also be impartial from an objective viewpoint, that is, it must offer sufficient guarantees to exclude any legitimate doubt in this respect.”<sup>140</sup> Under the objective test, “it must be determined whether . . . there are *ascertainable facts* which may raise *doubts* as to their impartiality.”<sup>141</sup> As summarized in *Kleyn*, “the question whether, in the circumstances of the case, [the Court] had the requisite “appearance” of independence, or the requisite “objective” impartiality.”<sup>142</sup> As the ECHR, highlights

In this respect even appearances may be of a certain importance. What is at stake is the confidence which the courts in a democratic society must inspire in the public and above all, as far as criminal proceedings are concerned, in the accused (...).<sup>143</sup>

In deciding whether there is a legitimate reason to fear that a particular court lacks independence or impartiality, the standpoint of the accused is important without being decisive. What is decisive is whether his doubts can be held to be objectively justified (...).<sup>144</sup>

Despite broad consensus on the importance of an impartial judiciary and decision maker,<sup>145</sup> as can be discerned from the above descriptions, the best measure of impartiality is open for debate. For example, some argue that the incredibly high standard of “actual” and/or “justifiable doubts” based upon verifiable facts is too high

---

<sup>138</sup> *Case of Kleyn and others v. The Netherlands*, [2003] ECHR, at para 191.

<sup>139</sup> *Case of Daktaras v. Lithuania*, [2000] ECHR, at para. 30.

<sup>140</sup> *Kleyn*, [2003] ECHR at para 191.

<sup>141</sup> *Daktaras*, [2000] ECHR, at para. 32

<sup>142</sup> *Kleyn*, [2003] ECHR at para 193.

<sup>143</sup> *Kleyn*, [2003] ECHR at para 191 *citing Morris v. the United Kingdom*, no. 38784/97, § 58, ECHR 2002-I.

<sup>144</sup> *Kleyn*, [2003] ECHR at para 194 *citing Hauschildt v. Denmark*, judgment of 24 May 1989, Series A no. 154, p. 21, § 48.

<sup>145</sup> For example, in addition to the texts described above the concept is enumerated within the United Nations Human Rights, *International Covenant on Civil and Political Rights*, (1976), Article 14, (“All persons shall be equal before the courts and tribunals. In the determination of any criminal charge against him, or of his rights and obligations in a suit at law, everyone shall be entitled to a fair and public hearing by a competent, independent and impartial tribunal established by law.” *Id.* at Art. 14(1)); Organization of African Unity, *African Charter of Human and Peoples’ Rights*, (2012), Art. 26, (States parties to the present Charter shall have the duty to guarantee the independence of the Courts).

a bar, instead arguing that the “appearance of impartiality” is in fact the appropriate standard. Many scholars argue that “appearance,” or the existence of *reasonable suspicion* of a lack of impartiality on behalf of a judge should be the measure. For example, Article 5 of the Universal Charter of the Judge reads:

In the performance of the judicial duties the judge must be impartial and must so be seen. The judge must perform his or her duties with restraint and attention to the dignity of the court and of all persons involved.<sup>146</sup>

While the independence of the decision maker is ensured primarily through the processes of appointment and discipline, which occurs outside the court itself, impartiality is protected largely in specific cases through the actions of judges and attorneys themselves via recusals and challenges.<sup>147</sup> As noted by Lord Bingham,

Of course, since judges and other decision makers are human beings and not robots, they are inevitably, to some extent, the product of their own upbringing, experience and background. The mind they bring to the decision of the issues cannot be a blank canvas. But they should seek to alert themselves to, and so neutralize, any extraneous considerations which might bias their judgement, and if they are conscious of bias, or of matter which might give rise to an appearance of bias, they must decline to make the decision in question.<sup>148</sup>

Council of Europe, when enumerating judicial responsibilities recommends that,

[J]udges should withdraw from a case or decline to act where there are valid reasons that should be defined by law and may, for instance, relate to serious health problems, conflicts of interest, or interests of justice.<sup>149</sup>

The inclusion of the phrase *conflict of interest* in this recommendation further expands the set of potential measures of bias, partiality and dependency issues. In this particular

---

<sup>146</sup> International Association of Judges, *The Universal Charter of the Judges*, Article 5, (1999) (“In the performance of the judicial duties the judge must be impartial and must so be seen..”) *Id.* at Art. 5 available at <http://www.iaj-uim.org/universal-charter-of-the-judges/> (last visited Aug. 14, 2019).

<sup>147</sup> For example, both *Kleyn*, [2003] ECHR and *Daktaras*, [2000] ECHR were brought before the ECHR based upon prior significant involvement in the cases where the parties sought to challenge the impartiality of the tribunal.

<sup>148</sup> Bingham, *RULE OF LAW*, *supra* note 37 at 93. This point is articulated by Justice Sonia Sotomayor in 2001, this will be explored in further detail in Chapter 5 Section A, 6. (b) Rejection of the “Robot” Decision Maker. *See infra* Chapter 5, Section A(6)(b).

<sup>149</sup> Council of Europe, *Role of Judges*, *supra* note 77, Principle V- Judicial Responsibilities, subsection (c).

instance, it is argued that “conflict of interest” includes both issues of material conflict, as, for example when the judge stands to financially benefit from the outcome of the case,<sup>150</sup> and of moral conflict, such as when a judge has close personal ties to the issue, a party, or some other aspect of the case.<sup>151</sup>

In these situations, and many others, judges are expected to refuse to hear the case and/or to stand down if the issue arises or comes to light during the course of the proceeding. In situations where the judge declines to recuse him- or herself, on the basis of potential bias or conflict of interest, the justice system typically introduces a process in which the parties themselves can challenge the judge’s participation. It is through this combination of self-monitoring and external appeal that the impartiality of the decision maker is ensured.

It is easy to see the overlap between dependence and partiality issues. Despite the distinctions, in many situations dependence and partiality are closely linked.<sup>152</sup> As noted by Lord Bingham: “Closely allied to the requirement of independence is the requirement that the decision maker be impartial.”<sup>153</sup>

In the creation of an online justice environment, the designers of the various platforms must ensure that the rule of law is followed, including the right to an impartial decision maker, understood as a decision maker free from external bias and from internal feelings of partiality. The online justice environment must: (1) define impartiality; (2) determine if actual bias, appearance of bias, or a hybrid contextual determination will be the governing standard; (3) determine a mechanism for the disclosure of potential impartiality and/or bias of the decision maker; (4) determine if an external third party

---

<sup>150</sup> See e.g., *AT&T Corporation v Saudi Cable Co* [2000] EWCA Civ. 154 (chairman of the tribunal held shares in corporation, however, the benefit would be de minimis and of minimal benefit to the arbitrator, thus chairman was not to be removed.) See *id.* at para. 55.

<sup>151</sup> See *Case of Pescador Valero v. Spain*, [ECHR] [4<sup>th</sup> section] (2003) (discussing the impartiality debate arising when the plaintiff had been laid off by the university with which the trial judge had close professional connections).

<sup>152</sup> *Kleyn*, [2003] ECHR at para. 192 (stating “The concepts of independence and objective impartiality are closely linked. . .”) *Id.* at 192.

<sup>153</sup> Bingham, *RULE OF LAW*, *supra* note 37 at 93.



will adjudicate the determination of bias/impartiality or if the decision will be left to the parties; (5) create a process to allow the decision maker to withdraw from making decisions when alerted to bias; and (6) create a system to replace the decision maker in such a case. Solutions to these six problems will be presented within the final chapter which specifically suggests a model designed to address the issues presented throughout the thesis.

## E. Rule of Law: The Law and Legal Reasoning

Because this thesis considers the impartiality and independence of the decision maker within the justice environment, one final consideration must be undertaken: the manner of decision making. In general, theories of adjudication, or how judges should decide cases, can be broadly divided into two general types, “formalist” theories and “realist” theories.<sup>154</sup> Because the manner in which decisions are made is an essential component of considering the design of a decision making system, the following section will briefly explore these two broad categories of theory and their relationship to the questions of this thesis.

### 1. Legal Formalism

According to Brian Leiter, “Formalist” theorists claim that:

(1) the law is “rationally” determinate, i.e., the class of legitimate legal reasons available for a judge to offer in support of his decision justifies one and only one outcome either in all cases or in some significant and contested range of cases; and (2) adjudication is thus “autonomous” from other kinds of reasoning, that is, the judge can reach the required decision without recourse to non-legal normative considerations of morality or political philosophy.<sup>155</sup>

It can be asserted that legal formalism is directly linked to the theory of law as objective. That is, if the law is objective then the outcome of any case to which the same set of legal rules is applied should be consistent regardless of the decision maker. Formalism is an ideal that posits that bias can be entirely eliminated. As such, technology-driven

---

<sup>154</sup> See, Brian Leiter, *Legal Formalism and Legal Realism: What Is the Issue?*, UNIVERSITY OF CHICAGO PUBLIC LAW & LEGAL THEORY WORKING PAPER, No. 320 (2010).

<sup>155</sup> *Id.* at 2.

decision-making systems that adopt the legal formalism approach should be designed with robust rules, ensuring predictable and consistent outcomes. Legal formalism embraces the idea that well-defined rules leave little room for alternative considerations, and outcomes are determined solely by rules, as in a closed system, without reference to the parameters of the external society or its commitments to morality. The rules guide the process, fully, without exception, to a predictable conclusion.

## 2. Legal Realism

According to Brian Leiter “realist” theorists purport to give “an unsentimental and honest account of what judges really do.”<sup>156</sup> That is, realists advance a descriptive theory of adjudication according to which:

- (1) legal reasoning is indeterminate (i.e., fails to justify a unique outcome) in those cases that reach the stage of appellate review;
- (2) appellate judges, in deciding cases, are responsive to the “situation-types”—recurring factual patterns (e.g., “seller of a business promises not to compete with the buyer, and then tries to break the promise”)—that elicit predictable normative responses (“this is unfair” or “this is economically foolish”) from most jurists, responses that are not, however, predictable based on existing “paper” rules and doctrine; and
- (3) in the commercial law context judges look to the “normal” practices in the existing business culture in deciding what is the right outcome.<sup>157</sup>

Realists argue,

[T]he law is filled with gaps and contradictions, that the law is indeterminate, that there are exceptions for almost every legal rule or principle, and that legal principles and precedents can support different results.<sup>158</sup>

---

<sup>156</sup> *Id.* at 3 citing Richard Posner, *HOW JUDGES THINK* (Harvard University Press, 2008).

<sup>157</sup> For detailed discussion and evidence, see Brian Leiter, *NATURALIZING JURISPRUDENCE: ESSAYS ON AMERICAN LEGAL REALISM AND NATURALISM IN LEGAL PHILOSOPHY* (Oxford: Oxford University Press, 2007), esp. Chs. 1-3. For a more concise account, see Brian Leiter, *American Legal Realism*, in *THE BLACKWELL GUIDE TO PHILOSOPHY OF LAW AND LEGAL THEORY* (M. Golding & W. Edmundson eds. 2005).

<sup>158</sup> Brian Tamanaha, *BEYOND THE FORMALIST-REALIST DIVIDE: THE ROLE OF POLITICS IN JUDGING*, 1, Princeton University Press (2009).

Realism seeks to establish a new theory that can track the logic of the day-to-day realities of judicial decision making.

Of course, the binary between formalism and realism fails to consider a hybrid model in which objective law exists, yet legal realism is the decision making model employed. Extreme adherence to either formalism or realism could lead a judge to run afoul of the rule of law in three key areas: judicial activism, discretion, and equality before the law.<sup>159</sup>

For example, judicial activism is a clear violation of the rule of law as it damages the predictability of the law as well as the expectation that the standards of the law are knowable at the time of the activity. Simply put, rule of law demands that citizens have the ability to know what the laws are, so they may inform their course of behavior and help them predict consequences and measure risk.<sup>160</sup> Taking a legal realist perspective too far runs the risk of allowing judicial activism to “recast the law”<sup>161</sup> and undermine this principle that the law should be stable enough that citizens can be expected to know what it is.

Legal formalism, when taken to the extreme, can be just as problematic, as it does not allow for judicial judgement and/or discretion. However, court systems are necessarily full of instances of judicial judgement that do not rise to the level of discretion. As will be discussed later in Chapter 4, decisions are based on everything from witness and expert testimony to evidence and document production: while these issues are not considered discretionary, nonetheless they are all within the purview of the decision maker. Overly formalistic approaches to such discussions unnecessarily limit the function of the decision maker and negatively impact the rule of law.

### 3. Artificial Intelligence and Legal Reasoning

---

<sup>159</sup> See Bingham, RULE OF LAW, *supra* note 37, at 45, 48-9, and 55.

<sup>160</sup> See *id.* at 37.

<sup>161</sup> See *id.* at 45.

Which of the widely divergent theories of judicial reasoning is used will have substantial implications for the design of the decision maker in the online environment. In fact, exploration of this problem is already well underway. As will be discussed more fully in Chapter 4, formalism is the backbone of current design, but it is already apparent that it will prove inadequate as we attempt to design more complex systems that address higher-complexity issues beyond mere compliance with statute.

One might assume legal analytics is in its infancy as a discipline, but this would be incorrect. As Kevin Ashley explains,

[T]he field of AI & Law has long studied how to design computer programs that can reason logically with legal rules from statutes and regulations. It has made strides, and demonstrates some success, but it has also developed an appreciation of just how difficult the problem is.<sup>162</sup>

It should not surprise anyone that the law, as it is written, is difficult to decipher and even more difficult to teach a computer to apply in context. As described by Kevin Ashley, when commenting on statutory text,

The problems of resolving syntactic ambiguity, reformulation, negation, counterfactual conditions, and semantic ambiguities are problems of interpreting natural language text.<sup>163</sup>

This comment is made in relation to statutory text, and does not even scratch the surface of the challenges associated with developing judicial decision making.

While a legal formalism approach can be captured in code, legal realism, with its acknowledgement of discretion and the necessity of adjustment to particular circumstances, is difficult to capture in an algorithmic environment. Nevertheless, the designers of the online justice system must understand how to balance both approaches. As articulated in prior sections of the chapter, stable rules are important, but bounded discretion is also essential.<sup>164</sup>

---

<sup>162</sup> Kevin D. Ashley, *ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS, NEW TOOLS FOR LAW PRACTICE IN THE DIGITAL AGE*, Cambridge University Press, at 39 (2017).

<sup>163</sup> *Id.* at 52.

<sup>164</sup> Bingham, *RULE OF LAW*, *supra* note 37, at 48-55.

Online justice model designers must resolve the following issues: (1) Will the system reflect the written law in a purely formal way, will it reproduce the reality of current judicial reasoning processes and outcomes, or will it be a hybrid system the uses both approaches? (2) When will formalist, realist, and hybrid approaches, respectively, most effectively maintain the rule of law? (3) Assuming formalism is used in some measure, what limitations must be set on it? (4) Assuming realism is used in some measure, what limitations must be set on it? (5) What are the appropriate points at which to introduce judicial discretion into the ODR process? And, (6) what RoL considerations and what potential impacts on the RoL will need to be reviewed once technology becomes a central part of the justice eco-system? Recommended answers to these questions will be presented in Chapter 5.

#### F. Rule of Law Conclusions

The importance of the decision maker's role in maintaining rule of law sets a high bar for the design of decision making in ODR. However, it is important to note at this stage that it is not necessarily the case that ODR will have no human decision makers — as we shall continue to explore in more depth later. In the meantime, we should also recall that departure from specific rules may be found allowable on Bingham's account after, and so long as, robust consideration is first given.<sup>165</sup> Rather than eliminating the human decision maker, ODR technology may be used to support them in achieving the independence, impartiality and freedom from the influence of public opinion called for by RoL principles. It might do this in several ways. First, technology can identify and thus facilitate the elimination of problematic dependency relationships prior to appointment of the decision maker. Moreover, supervising authorities and commentators will be able to monitor the outcomes of the ODR system for partiality, and both decision makers themselves and outside monitors will be able to receive feedback when patterns of partiality are discovered. As will be explored more in Chapters 4 and 5, technology deployed in such a manner can in fact reduce partiality and dependence issues. This is one of the positive arguments for expanding the use of technology in justice systems.

---

<sup>165</sup> Bingham, *RULE OF LAW*, *supra* note 37, at 8.

## CHAPTER 2: Arbitration Underpinnings

Online Dispute Resolution draws much of its anticipated structure from the widespread use of arbitration as an alternative dispute resolution mechanism. Although arbitration is not the only alternative to the justice system, nor is it the only option within the world of ODR,<sup>166</sup> it has a rich and robust history that provides much of the context for the development of new solutions. There is an extensive literature on fair hearing issues that arise outside the traditional justice system in the context of arbitration. Thus, arbitration will be used to frame much of the present discussion of impartiality and independence in ODR.

Let us first establish the basic tenets of arbitration. The arbitration process is juridical in nature; consequently, applicable national law places limits upon the parties' freedom of contract.<sup>167</sup> In general, this is because the tribunal stands in the place of the judiciary and its awards have the same final and binding effect as a judgement issued from a national court.<sup>168</sup> Arbitral awards are enforced at a local level via the authority of the

---

<sup>166</sup> See discussion, *infra* Chapter 3, Section C. As explained in Footnote No. 5. It is important to note, on August 7, 2019, the UNITED NATIONS CONVENTION ON INTERNATIONAL SETTLEMENT AGREEMENTS RESULTING FROM MEDIATION [United Nations: United Nations Commission on International Trade Law, (March 2019)(hereinafter, the Singapore Convention)] opened for signature in Singapore. Over 40 states, including China, India and the U.S., have signed the treaty. See *id.* As such, the Convention is expected to enter into force shortly. See *id.* The Singapore Convention requires member states to enforce settlement agreements and gives effect to settlement agreements such that they may be used to prevent parties from litigating the matter. See *id.* Convention Article 1 (3), specifies ““Mediation” means a process, irrespective of the expression used or the basis upon which the process is carried out, whereby parties attempt to reach an amicable settlement of their dispute *with the assistance of a third person or persons. . .*” *Id.* The use of the term ‘person’ may implicate awards issued in an automated process. Because the Convention is legally untested, part of an emerging international mediation community, and *may require* a ‘person’ as a mediator, the Convention will not form a central portion of the discussions within the thesis. However, as the Convention develops in the international commercial community, it should be considered as an important additional protection that could thus, be within the ODR process.

<sup>167</sup> Julian D. M. Lew, Loukas A. Mistelis, Stefan M Kroll, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, Ch. 5, 1.1 and 2.1, Kluwer Law (2003)[hereinafter Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*]

<sup>168</sup> See *id.* at Ch. 26, 3.1.

state;<sup>169</sup> thus national law requires the arbitral proceedings and the composition of the tribunal to meet minimum standards.<sup>170</sup> These standards provide a benchmark for what are usually considered to be essential elements within a justice environment.<sup>171</sup>

This chapter will first explore the various institutional rules relating to impartiality and independence that have been developed for arbitration environments. Second, the chapter will analyze areas of convergence and divergence among these rules. Third, the chapter will examine existing rules relating to disclosure and challenges to the decision maker within arbitration. To further explore the nuances surrounding determinations of independence and impartiality, the chapter will consider the role of the traditional brick-and-mortar court in the determination of impartiality and independence. And finally, the chapter will engage in a limited review of cases that set out parameters and standards when courts are called on to determine issues of decision maker dependence and partiality.

#### A. Arbitration: Impartiality and Independence

As emphasized by Lew, Mistelis, and Kroll:

in line with his special adjudicative function, the arbitrator has to be and remain independent and impartial and disclose all facts which may be relevant.<sup>172</sup>

One can thus isolate two timeframes during which impartiality and independence are at issue: (1) at the time of appointment, and (2) throughout the decision making process. The noted authorities assert that arbitrators must disclose all relevant facts, and one can infer they intend this to impose upon arbitrators a requirement of disclosure prior to appointment as well as an ongoing obligation to disclose all relevant facts that arise during the decision making process. The proscribed duty placed upon the arbitrator is a general, yet fundamental expectation.

---

<sup>169</sup> *See id.*

<sup>170</sup> *See id.*

<sup>171</sup> *See id.*

<sup>172</sup> Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140 at 255.

If international arbitration is to have a bedrock of commonly recognized core principles, it is essential for fundamental concepts such as independence and impartiality to have generally accepted meanings and standards in the international arbitration community.<sup>173</sup>

The institutions of arbitration have developed more specific and detailed explanations of this duty. The following table is a compilation of the rules set forth by the major arbitration institutions in the area of independence and impartiality, gathered from the institutions’ websites and widely available online and in publications.

*Table B: Arbitral Rules of Impartiality/Independence Standards*<sup>174</sup>

Rules	Disclosure standard: Objective/Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
UNCITRAL 2013	Objective	Impartiality and Independence	Article 11 “When a person is approached in connection with his or her possible appointment as an arbitrator, <b>he or she shall disclose any circumstances likely to give rise to justifiable doubts</b> as to his or her impartiality or independence.” [No individual qualification clause]

<sup>173</sup> Nick Gray and Deborah Crosbie, *Winds of change? The Pending Publication Of LCIA Reasoned Decisions On Arbitral Independence*, page 7, Slaughter and May Online Publication, (April 2009).

<sup>174</sup> The chart was compiled in August 2017 from the institutions, via their institutional websites listed in the first column.



Rules	Disclosure standard: Objective/ Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
LCIA 2014	Viewpoint of Parties	Impartiality and Independence	<p>5.3 “All arbitrators shall be and remain at all times impartial and independent of the parties.”</p> <p>5.4 “Before appointment by the LCIA Court...the candidate shall sign a written declaration stating: (i) whether there are any circumstances currently known to the candidate which are likely to give rise <b>in the mind of any party to any justifiable doubts</b> as to his or her impartiality or independence and, if so, specifying in full such circumstances in the declaration;”</p>
ICC 2012	Viewpoint of Parties	Only Impartiality (Note: The rationale is that the omission of a statement of independence enables states to appoint state nationals as arbitrators.)	<p>Article 11.1 “Every arbitrator must be and remain impartial and independent of the parties involved in the arbitration”</p> <p>Article 11.2 “Before appointment or confirmation, a prospective arbitrator shall sign a statement of...impartiality and independence. The prospective arbitrator shall disclose in writing to the Secretariat any facts or circumstances which might be of such a nature as to call into question the arbitrator’s independence <b>in the eyes of the parties</b>, as well as any circumstances that could give rise to reasonable doubts as to the <b>arbitrator’s impartiality.</b>”</p>
SCC arbitration 2010	Objective	Impartiality and Independence	<p>Article 14(1) “Every arbitrator must be impartial and independent.”</p> <p>Article 14(2) “Before being appointed as arbitrator, a person shall disclose any circumstances which may give rise to justifiable doubts as to his/her impartiality or independence. If the person is appointed as arbitrator, he/she shall submit to the Secretariat a signed statement of impartiality and independence <b>disclosing any</b></p>

Rules	Disclosure standard: Objective/Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
			<b>circumstances which may give rise to justifiable doubts as to that person’s impartiality or independence...</b>
ICDR 2014	Objective	Impartiality and Independence	Article 13.1 “Arbitrators acting under these Rules shall be impartial and independent and shall act in accordance with the terms of the Notice of Appointment provided by the Administrator.” Article 13.2 “The arbitrator shall <b>disclose any circumstances that may give rise to justifiable doubts as to the arbitrator’s impartiality or independence and any other relevant facts the arbitrator wishes</b> to bring to the attention of the parties.”
CIETAC 2015	Objective	Impartiality and Independence	Article 24 “An arbitrator shall not represent either party, and shall be and remain independent of the parties and treat them equally.” Article 31.1 “An arbitrator nominated by the parties or appointed by the Chairman of CIETAC shall sign a Declaration and <b>disclose any facts or circumstances likely to give rise to justifiable doubts as to his/her impartiality or independence.</b> ”
SIAC 2016	Objective	Impartiality and Independence	13.1 “Any arbitrator appointed in an arbitration under these Rules, whether or not nominated by the parties, shall be and remain at all times independent and impartial.” 13.4 “A nominated arbitrator shall <b>disclose to the parties and to the Registrar any circumstances that may give rise to justifiable doubts as to his impartiality or</b>

Rules	Disclosure standard: Objective/Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
			<b>independence</b> as soon as reasonably practicable and in any event before his appointment.”

Author compilation of online website information (Up-to-date April 2019)

The table illustrates the difficulty of synthesizing a single standard. Institutions have used a variety of terminology to make rules about impartiality and independence. There is no single institutional rule or industry standard that encompasses all of the necessary determinations that go into the examination of arbitrator impartiality and independence. In fact, some rules do not explicitly prescribe the standards, but, by providing grounds on which an arbitrator’s participation may be challenged, imply them. Adding to the confusion, many institutional rules adopt familiar legal phrasing, but do so without definition or proscriptions for interpretation. As a result, terms such as “justifiable doubts” and “reasonable doubt” need to be interpreted on a case-by-case basis based on the applicable rule and local law. In addition, arbitration institutions, codes of ethics, and local law all use differing, yet similar, terminology. For example, the *International Bar Association Rules of Ethics for International Arbitrators* consider impartiality and independence within the context of the elements that can bias the arbitrator.<sup>175</sup> Finally, while institutional rules play an important part in the interpretation of the parties’ arbitration agreement, in the end much of the interpretation will be done in a traditional brick-and-mortar courthouse, leaving the interpretation subject to local law.

The differences between the various institutions and corresponding institutional rules may be seen as providing a market opportunity, that is, an opportunity for the parties to an arbitration to select the institutional rules and jurisprudence that best fits their needs under the circumstances. This thesis, however, takes the position that there must be limits placed on this variability: the parties should not be able to simply contract around the basic protections of the rule of law. Yet, the ability of parties to select between

---

<sup>175</sup> See, International Bar Association, *Rules of Ethics for International Arbitrators*, Canon 3 (1987) [Hereinafter IBA, Ethics].

institutions and the voluntary, consent driven aspect of this choice is an essential consideration within the context of rule of law considerations, as was described in Chapter 1.<sup>176</sup>

## B. Arbitration: Disclosure and Challenges

Arbitral institutions generally use two mechanisms to reduce the impact of dependence and partiality in the decision making process: (1) disclosure of information by the arbitrator, and (2) challenges to the appointment of the arbitrator. Each of these issues will be explored, in turn.

### 1. Disclosure

According to Lew, Mistelis and Kroll:

To ensure compliance with the requirements of independence and impartiality, arbitrators are generally under a duty to disclose to the parties all facts which may be relevant in the respect.<sup>177</sup>

As can be seen, arbitrators must disclose *all relevant* facts. The duty of disclosure leads to a series of questions: (1) When must this disclosure occur? (2) How should “relevant” be defined? (3) What must be disclosed to satisfy the duty of disclosing “all” relevant facts? (4) What point of view is to be considered when making the determination of the facts to disclose?

The following table is a compilation of the rules laid out by the major arbitration institutions in the area of disclosure standards, as gathered from their respective websites and widely available online and in publications:

#### *Table C: Arbitral Disclosure Standards*

---

<sup>176</sup> See Chapter 1, discussion at *supra* footnote 31.

<sup>177</sup> Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140 at 265.

Rules	Disclosure standard: Objective/ Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
UNCITRAL 2013	Objective	Impartiality and Independence	Article 11 “When a person is approached in connection with his or her possible appointment as an arbitrator, <b>he or she shall disclose any circumstances likely to give rise to justifiable doubts</b> as to his or her impartiality or independence.” [No individual qualification clause]
LCIA 2014	Viewpoint of Parties	Impartiality and Independence	5.3 “All arbitrators shall be and remain at all times impartial and independent of the parties.” 5.4 “Before appointment by the LCIA Court...the candidate shall sign a written declaration stating: (i) whether there are any circumstances currently known to the candidate which are likely to give rise <b>in the mind of any party to any justifiable doubts</b> as to his or her impartiality or independence and, if so, specifying in full such circumstances in the declaration;”
ICC 2012	Viewpoint of Parties	Only Impartiality (Note: The rationale is that the omission of a statement of independence enables states to appoint state nationals as arbitrators.)	Article 11.1 “Every arbitrator must be and remain impartial and independent of the parties involved in the arbitration” Article 11.2 “Before appointment or confirmation, a prospective arbitrator shall sign a statement of...impartiality and independence. The prospective arbitrator shall disclose in writing to the Secretariat any facts or circumstances which might be of such a nature as to call into question the arbitrator’s independence <b>in the eyes of the parties</b> , as well as any circumstances that could give rise to

Rules	Disclosure standard: Objective/Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
			reasonable doubts as to the <b>arbitrator's impartiality.</b> "
SCC arbitration 2010	Objective	Impartiality and Independence	Article 14(1) "Every arbitrator must be impartial and independent." Article 14(2) "Before being appointed as arbitrator, a person shall disclose any circumstances which may give rise to justifiable doubts as to his/her impartiality or independence. If the person is appointed as arbitrator, he/she shall submit to the Secretariat a signed statement of impartiality and independence <b>disclosing any circumstances which may give rise to justifiable doubts as to that person's impartiality or independence...</b> "
ICDR 2014	Objective	Impartiality and Independence	Article 13.1 "Arbitrators acting under these Rules shall be impartial and independent and shall act in accordance with the terms of the Notice of Appointment provided by the Administrator." Article 13.2 "The arbitrator shall <b>disclose any circumstances that may give rise to justifiable doubts as to the arbitrator's impartiality or independence and any other relevant facts the arbitrator wishes</b> to bring to the attention of the parties."
CIETAC 2015	Objective	Impartiality and Independence	Article 24 "An arbitrator shall not represent either party, and shall be and remain independent of the parties and treat them equally." Article 31.1 "An arbitrator nominated by the parties or appointed by the Chairman of CIETAC shall sign a Declaration and <b>disclose any facts or circumstances likely to give rise to justifiable doubts as to his/her impartiality or independence.</b> "

Rules	Disclosure standard: Objective/Viewpoint of Parties?	Impartiality Disclosure only?	Relevant Wordings of Qualification and/or Disclosure
SIAC 2016	Objective	Impartiality and Independence	13.1 “Any arbitrator appointed in an arbitration under these Rules, whether or not nominated by the parties, shall be and remain at all times independent and impartial.” 13.4 “A nominated arbitrator shall <b>disclose to the parties and to the Registrar any circumstances that may give rise to justifiable doubts as to his impartiality or independence</b> as soon as reasonably practicable and in any event before his appointment.”

Author compilation of online website information (Up-to-date April 2019)

Unlike the varied standards for impartiality and independence themselves, the disclosure standards of the major arbitration institutions are incredibly similar. These standards can be collectively summarized as:

- (1) Disclosure must occur prior to the appointment, and must be of facts known to the arbitrator at the time of appointment.
- (2) The disclosure duty is ongoing. Consequently, if new facts should come to the attention of the arbitrator the facts must be disclosed, immediately.
- (3) “Relevant” is to be measured in the eye of the parties.
- (4) Yet, the terms “all” and “relevant” are limited to facts that give rise to “justifiable doubts” implicating the arbitrator’s impartiality and independence.

While the requirements of multiple arbitration bodies can be easily condensed into this simple list, interpretation of the phrases, especially of the term “justifiable doubts,” has required courts to closely examine the true intention of these arbitration rules. The current trend articulated by some commentators is to “disclose anything and everything.”<sup>178</sup> The “disclose everything” standard is reflected in the 2009 case of *SA*

---

<sup>178</sup> For example, the London Court of International Arbitration (LCIA) recommends “Any doubt as to whether a relationship should be disclosed must be resolved in favour of disclosure.” London Court of International Arbitration, *Notes for Arbitrators*, Section 2.6, (2014).

*J&P Avax SA v. Société Tecnimont SPA*,<sup>179</sup> in which the Paris Court of Appeal annulled an arbitral award because the chairman of the arbitral tribunal had failed to disclose certain circumstances, even though the circumstances did not exist at the time of his appointment, nor did he become aware of the circumstances during the arbitral process.<sup>180</sup>

Reliance upon court interpretation, however, leaves the ecosystem of international commercial arbitration dependent upon varying judicial readings of these standards. International commercial arbitration has always been resistant to court involvement, which introduces variable local interpretations of key features, as this variability tends to undermine the international nature of arbitration. Arbitration institutions seek to develop guidance to mitigate this problem, especially in essential areas of importance within the arbitration community, such as disclosure, independence and impartiality. In response to such concern, the International Bar Association developed *Guidelines on Conflicts of Interest in International Arbitration*. Within the *Guidelines*, the Association seeks to protect the parties using the alternative justice system, to this end embracing a party-centric view:

[I]t is important that more information be made available to the parties, so as to protect awards against challenges based upon alleged failures to disclose, and to promote a level playing field among parties and among counsel engaged in international arbitration.<sup>181</sup>

The party-centric view of “disclose everything” had led to its own difficulties,<sup>182</sup> however, as arbitrators worry that they are being asked to disclose information that is otherwise not relevant, or that their appointments are likely to be blocked for trivial

---

<sup>179</sup> *SA J&P Avax SA v. Société Tecnimont SPA*, Court of Appeal of Paris, 12 February 2009, Rev. Arb. 2009 186, discussed in Alexis Mourre, *Conflicts of Interest: Towards Greater Transparency and Uniform Standards of Disclosure?*, Castaldi Mourre & Partners, (May 19, 2009) available at <http://kluwerarbitrationblog.com/2009/05/19/conflicts-of-interest-towards-greater-transparency-and-uniform-standards-of-disclosure/> (last visited Aug. 14, 2019).

<sup>180</sup> See *id.*

<sup>181</sup> See International Bar Association, *Guidelines on Conflicts of Interest in International Arbitration*, Introduction Comment 1, (2014).

<sup>182</sup> See discussion, *infra* Chapter 5, Section 3(b)(i).



reasons.<sup>183</sup> For example, since the implementation of the 2014 Rules the LCIA has published *Notes for Arbitrators*,<sup>184</sup> designed to assist decision making in key areas of interest for various parties in the arbitration process. Included in the *Notes* is specific guidance on independence and impartiality:

Parties to arbitrations are entitled to expect of the process a just, well-reasoned and enforceable award. To that end, they are entitled to expect arbitrators: to disclose possible conflicts of interest at the outset; to avoid putting themselves in a position where conflicts will arise during the course of the proceedings; to conduct the arbitration fairly, in a timely manner and with careful regard to due process; to maintain the confidentiality of the arbitration; and to reach their decision in an impartial manner.<sup>185</sup>

The LCIA Section goes on to note:

In completing their statements of **independence**, arbitrators should take into account, amongst other things, the existence and nature of any past or present **relationships**, direct or indirect, with any of the parties or their counsel. **Any doubt as to whether a relationship should be disclosed must be resolved in favour of disclosure.**<sup>186</sup>

The LCIA and the IBA both seek to protect parties' trust in the arbitration process as an alternative to the brick-and-mortar courtroom. Parties must trust the system and the requisite trust arises from adherence to rule of law principles, especially those articulated by Lord Bingham in relation to the neutrality of the decision maker.<sup>187</sup> Parties simply must have information such that they feel confident that the decision maker is free of dependency and partiality.

The duty to disclose raises three issues from the point of view of the person bearing it. First, is it even possible for anyone to every disclose "everything," and more importantly, should that be the expectation? Surely, as an arbitrator I have a right not to disclose certain aspects of my life that have no bearing on the case. Of course, that

---

<sup>183</sup> See IBA, *Conflicts*, *supra* note 153 at Comment 1.

<sup>184</sup> LCIA, *Notes for Arbitrators*, *supra* note 150.

<sup>185</sup> *Id.*

<sup>186</sup> *Id.* at Section 2.8 (2014) (highlights added).

<sup>187</sup> See Bingham, *RULE OF LAW*, *supra* note 37 at 91.

raises the issue of who should get to decide what information has a potential bearing on the case.<sup>188</sup> Second, assuming arbitrators are willing to disclose “everything,” the level of complexity and time commitment needed to evaluate lengthy disclosures seems prohibitive in today’s world.<sup>189</sup> Conversely, if the duty is interpreted narrowly, for example by being limited only to disclosing interpersonal relationships, disclosure may not effectively assuage concerns about partiality and dependence. Thus, I suggest a possible fifth standard component of the duty of the disclosure:

(5) the expectation that the disclosure will provide sufficient information so the party to the arbitration can make an informed decision about the participation of the particular arbitrator in this particular proceeding.

In other words, decision makers should be required to disclose enough information to allow a party to make an informed decision,<sup>190</sup> but should not be required to disclose “everything.” Adding this fifth expectation to the general duty to disclose embeds the desired outcome, adequate information for all parties, into the duty to disclose, and should also serve to address the concerns of decision makers who are attempting to draw the line between failure to comply with disclosure expectations and over-disclosure of personal information.

To further limit, and provide guidance, on the “disclose everything” standard, the IBA included limits within the *Guidelines*,

In order to promote greater consistency and to avoid unnecessary challenges and arbitrator withdrawals and removals, the [IBA] Guidelines list specific situations indicating whether they warrant disclosure or disqualification of an arbitrator.<sup>191</sup>

---

<sup>188</sup> See discussion, *infra* Chapter 5, Section 3(B)(i)(1).

<sup>189</sup> See *id.*

<sup>190</sup> Informed consent is an elevated expectation in conversation concerning consent, it is the standard to be used in the thesis. The importance of consent, within the discussion of the rule of law is explored more fully in Chapter 1, discussion at *supra* footnote 31.

<sup>191</sup> See IBA, *Conflicts*, *supra* note 153, at comment 3. However, note: “Although there may strictly speaking be a logical inconsistency between a subjective disclosure test and the Green List’s objective non-disclosure situations as noted by the ICC, it is suggested that this is not such a fundamental incompatibility but rather that the Green List is a sensible and workable compromise. The Green List has the potential to bring clarity and consistency to disclosure decisions. It is not likely unduly to restrict proper disclosure, but rather to encourage subjective disclosure where appropriate.” See Nick Gray, *Winds of change?*, *supra* note 146 page 3.

Recognizing the limited value of a principle that presented without guidance on its applicability, the *Guidelines* also add level of detail and specificity to guide decision makers in their disclosure choices. For example, the *Guidelines* establish the “Non-Waivable Red List”:

1.1 There is an identity between a party and the arbitrator, or the arbitrator is a legal representative or employee of an entity that is a party in the arbitration.

1.2 The arbitrator is a manager, director or member of the supervisory board, or has a controlling influence on one of the parties or an entity that has a direct economic interest in the award to be rendered in the arbitration.

1.3 The arbitrator has a significant financial or personal interest in one of the parties, or the outcome of the case.

1.4 The arbitrator or his or her firm regularly advises the party, or an affiliate of the party, and the arbitrator or his or her firm derives significant financial income therefrom.<sup>192</sup>

Conflicts specified on the Non-Waivable Red List *must* be disclosed, and when they are discovered the parties are not allowed to waive the conflict; consequently, after such a disclosure the decision maker may in no circumstance continue to serve as arbitrator. Non-Waivable Red List conflicts are so fundamentally at odds with the expectations of the rule of law principle that there are no protections that can be placed within the process to offset them.

Not every conflict rises to this level. As described in Chapter 1, when exploring Lord Bingham’s rule of law principles: decision makers are guided by the law, bounded by the rules, and are directed in the application of rules that govern appropriate moments for discretion.<sup>193</sup> The *Guidelines* also establish a Waivable Red List of conflicts; these are circumstances that do allow the introduction of discretion, in this case, by the parties.

---

<sup>192</sup> IBA, *Conflicts*, *supra* note 153, Non- Waivable Red List at Section 1, page 20, (excerpt).

<sup>193</sup> Bingham, *RULE OF LAW*, *supra* note 37; *discussion*, *supra* Chapter 1, Section A.

When confronted with conflicts of interest presented within the Waivable Red List, the parties to arbitration have the discretionary power to determine if the hearing will progress despite the disclosure. The following is a short excerpt from the list:

2.1 Relationship of the arbitrator to the dispute

2.1.1 The arbitrator has given legal advice, or provided an expert opinion, on the dispute to a party or an affiliate of one of the parties.

2.1.2 The arbitrator had a prior involvement in the dispute.

2.2 Arbitrator's direct or indirect interest in the dispute.

2.2.1 The arbitrator holds shares, either directly or indirectly, in one of the parties, or an affiliate of one of the parties, this party or an affiliate being privately held.

2.2.2 A close family member of the arbitrator has a significant financial interest in the outcome of the dispute. . . .

2.2.3 The arbitrator, or a close family member of the arbitrator, has a close relationship with a non-party who may be liable to recourse on the part of the unsuccessful party in the dispute.<sup>194</sup>

2.3 Arbitrator's relationship with the parties or counsel

2.3.1 The arbitrator currently represents or advises one of the parties, or an affiliate of one of the parties.

2.3.2 The arbitrator currently represents or advises the lawyer or law firm acting as counsel for one of the parties.

2.3.3 The arbitrator is a lawyer in the same law firm as the counsel to one of the parties.<sup>195</sup>

The Waivable Red List contains conflicts of interest, based on potential dependency or potential partiality, that are better evaluated on a case-by-case basis and may lead to a change of decision maker or may not. Dependency issues arising from prior representation are common, especially in global law firms. Hence, greater detail is

---

<sup>194</sup> IBA, *Conflicts*, *supra* note 153, Waivable Red List at Section 2, pages 20-22, (excerpt).

<sup>195</sup> *Id.* at Section 2.3, pages 21-22, (excerpt).

demanded and case-by-case consideration is essential before it can be determined whether to waive the conflict or not; a single rule will not be satisfactory in the majority of instances.

In total, the *Guidelines* provide three lists, each with a different level of guidance for the decision maker. The third is the Green List, which sets out examples of relationships that should not cause any hesitation on the appointment of a decision maker. For example, the Green List contains situations such as “teaching in the same faculty or school as another arbitrator or counsel”<sup>196</sup> and “arbitrator was a speaker, moderator or organizer in one or more conferences.”<sup>197</sup> The presence of the Green List informs all that some relationships are allowed; parties are not to expect a complete detachment or absence of relationships in a decision makers’ environment. Second, in 2014 the *Guidelines* were updated to reflect the growing connectivity of the global community and law firms. For example, the 2014 *Guidelines* add:

3.3.6 A close personal friendship exists between an arbitrator and a counsel of a party.<sup>198</sup>

...

3.3.9 The arbitrator and another arbitrator, or counsel for one of the parties in the arbitration, currently act or have acted together within the past three years as co-counsel.<sup>199</sup>

These are situations where the circumstances may lead parties to the dispute to doubt the decision maker’s impartiality and independence. Thus, the parties should be informed and can elect whether or not to allow the process to move ahead with the particular decision maker in question.

The IBA Guidelines are one of the best sources available to explore the issue of conflicts of interest, for two main reasons. First, the guidelines are created by an institution with years of experience and a wealth of members knowledgeable in arbitration. This point is incredibly important, as the arbitration community has always

---

<sup>196</sup> *Id.* at Section 4.3.3.

<sup>197</sup> *Id.*

<sup>198</sup> *Id.* at Section 3.3.6.

<sup>199</sup> *Id.* at Section 3.3.9.

functioned based on appointed decision makers. The courts have a system of appointment, governing law and local rules, and mechanisms to remove judges. As it became an increasingly popular alternative route to justice, arbitration needed to establish similar processes and rules. The international commercial arbitration community stepped up to this task, and in doing so, has already effectively considered and resolved some of the issues pertinent to the design of the online justice environment. Second, arbitration as an essential stage within online justice can benefit from existing arbitration norms, such as those embodied within the *Guidelines*. Moreover, the arbitration community has long embraced a commitment to the rule of law, especially adherence to impartial and independent process and decision makers. The ODR community's willingness to embrace and comply with this commitment to the rule of law should not be underappreciated.

It is important to note:

[T]he Guidelines were only ever intended to be (and could realistically only ever have been) a starting point in a process that would, in order to meet its objectives, require the co-operation of the wider international arbitration community. As with any other code, statute, or general principle, it is therefore primarily and most effectively through reasoned decisions as to its application in any given situation that its meaning and application becomes increasingly clear and consistent.<sup>200</sup>

Hence, the Guidelines serve as a crucially useful resource for developing the dependence and partiality limits within the ODR environment, as we will unpack in Chapter 5. However, a fully informed approach to designing ODR would not be complete without also understanding the importance of institutions and the courts in creating the full body of rules.

## 2. Institutional Action Concerning the Decision Maker

We have seen that decision maker disclosure and recusal are not always sufficient protections; parties must therefore have within their power a right to challenge the

---

<sup>200</sup> Gray, *Winds of change? supra* note 146 at page 4.

appointment of the arbitrator.<sup>201</sup> We have also seen that several divergent standards have emerged to define grounds for such a challenge.<sup>202</sup> The standard most prevalent amongst the various institutions is one of “justifiable doubts” about the decision maker’s impartiality and independence.<sup>203</sup> However, even this standard is supported by less than a clear and overwhelming majority.

Moreover, even when tests of impartiality and independence exist,<sup>204</sup> the application of the rule and the applicable test is still something that must occur on a case-by-case basis within the authority of the arbitral institution<sup>205</sup> or at the time of enforcement of the final award by the enforcing jurisdiction.<sup>206</sup> For example, the LCIA has the authority, as established within LCIA Article 10, to “revoke any arbitrator’s appointment upon its own initiative, at the written request of all other members of the Arbitral Tribunal or upon a written challenge by any party . . .”<sup>207</sup> The LCIA is thereby empowered as the selected arbitration institution to make the determination concerning the partiality and dependence of the arbitrator. When considering such an issue, the institution applies its rules to the case at hand and determines the ability of the decision maker to resolve the case.<sup>208</sup> In other words, each institution is able to create its own test of impartiality and independence, the LCIA test described above being only one example. The parties to arbitration authorize that power in their choice of arbitration institution and its corresponding rules.<sup>209</sup> Yet, since each arbitration institution determines the parameters

---

<sup>201</sup> See Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140 at 301 (“... it should be possible for the parties to remedy the situation and have the arbitrator removed.”) *Id.*

<sup>202</sup> See *supra* Chapter 2, Table C.

<sup>203</sup> See *id.*

<sup>204</sup> See *supra* Chapter 2, Table B.

<sup>205</sup> See *discussion*, Chapter 2, Section B(1).

<sup>206</sup> Please note: partiality and dependence of the decision maker can arise when a party seeks to remove an arbitrator and asks the assistance of the local judicial system or when one of the parties seek to enforce the final arbitration award. See *infra* Chapter 2, Section C: The Role of Brick and Mortar Courts.

<sup>207</sup> London Court of International Arbitration, *Arbitration Rules*, Article 10.1, (2014).

<sup>208</sup> For example, the LCIA standard is enumerated in Article 10 which states: “The LCIA Court may determine that an arbitrator is unfit to act . . .” LCIA, *Arbitration Rules*, Article 10.2.

<sup>209</sup> For example, the LCIA rules state: “the parties thereto shall be taken to have agreed in writing that any arbitration between them shall be conducted in accordance with the LCIA Rules. . .” LCIA, *Arbitration Rules*, Preamble.

and application of the test, defining the exact measure of impartiality and independence falls to them.

Some institutions have published the determinations they have made at the institutional level; this is especially true as it relates to determinations of independence and impartiality. In 2011, the LCIA published 28 abstracts of decisions relating to arbitrator challenges that had occurred between 1996 and 2010.<sup>210</sup> The database lists a further 32 LCIA Court decisions relating to dependence and partiality covering the following period of 2010 to 2017. Since there are over 1,600 arbitrations registered with the LCIA,<sup>211</sup> we can quickly see that challenges based on partiality and dependence are rare. The total number of successful challenges is even lower: of the “32 challenges detailed in the database, only seven were upheld in whole or part — the remainder were rejected by the LCIA Court.”<sup>212</sup> Of course, the low number of challenges could occur for various reasons, from lack of trust in the system to enforce outcomes to a simple lack of awareness of the process. Such negative reasons seem relatively unlikely, however, as the process is well established in the international arbitration community. Thus it seems more likely that the low number of challenges is an indication that other safeguards are working well, leaving the parties without the need to formally challenge at the institutional level.

### 3. Transparency of Outcomes

Although transparency may not immediately be seen as an adjacent concern to the discussions surrounding the impartiality and independence of the decision maker, in the context of arbitration it is actually an important consideration. Unfortunately, transparency is not part of the commercial arbitration eco-system. Rather,

---

<sup>210</sup> See London Court of International Arbitration, *Arbitration International*, vol. 27, no. 3,(2011)(noting: “Of the 28 published abstracts, six cases were considered to be sufficiently problematic to satisfy the LCIA challenge standard, . . . about an arbitrator’s impartiality or independence.”) *Id.*

<sup>211</sup> LCIA, *Facts and Figures*, 2017 Casework Report, 4, (2018).

<sup>212</sup> *Id.* at 17.



confidentiality is a bedrock tenet.<sup>213</sup> Arbitration is a process of dispute resolution that champions confidentiality as one of its selling points.<sup>214</sup> The proceedings are confidential,<sup>215</sup> the papers within the proceedings are confidential,<sup>216</sup> the outcomes are confidential,<sup>217</sup> most parties consider<sup>218</sup> — and courts systems expect<sup>219</sup> — the process to be confidential. Unfortunately, this has become a barrier preventing essential terminology from taking shape, and causing confusion and uncertainty for all involved.

The lack of transparency came to the forefront of debate in 2004 when the First Edition of the International Bar Association *Guidelines on Conflicts of Interest in International Arbitration* commented: “existing standards lack sufficient clarity and uniformity in their application.”<sup>220</sup> The IBA was not the only entity commenting on the uncertainty. For example, Yves Derains and Eric A. Schwarz point out in their *Guide to the ICC Rules of Arbitration* that the failure to define “independence” or to refer to “impartiality” has resulted in “confusion and controversy” about the nature of the requirement set out in [ICC] Article 7(1).<sup>221</sup> Relying upon these commentaries (and others) Geoff Nicholas and Constantine Partasides argued in 2006 that the LCIA should publish the Court's decisions as they relate to challenges to the arbitrator. The authors conclude:

There is something troubling about institutions choosing to withhold the guidance upon which they rely in making challenge decisions from the parties that are making or defending those challenges. Such withholding

---

<sup>213</sup> See Anjanette H. Raymond *Confidentiality, In a Forum of Last Resort? Is the Use of Confidential Arbitration a Good Idea for Business and Society?*, 16 AMERICAN REVIEW OF INTERNATIONAL ARBITRATION, 479 (2005).

<sup>214</sup> See Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140 at Chapter 24, Section 24-99.

<sup>215</sup> See *id.*

<sup>216</sup> See *id.*

<sup>217</sup> See *id.* at Section 24-101.

<sup>218</sup> See, White and Case/Queen Mary, School of International Arbitration, 2018 *International Arbitration Survey, entitled “The Evolution of International Arbitration,”* White and Case Publication (2018)(“87% of respondents believe that confidentiality in international commercial arbitration is of importance. Most respondents think that confidentiality should be an opt-out, rather than an opt-in, feature.”) *Id.* at 3.

<sup>219</sup> See Raymond, *Confidentiality*, *supra* note 184 at 479.

<sup>220</sup> IBA, *Conflicts*, *supra* note 153, at Comment 3.

<sup>221</sup> Yves Derains and Eric A. Schwarz, *GUIDE TO THE ICC RULES OF ARBITRATION*, 116 (Second Edition), Kluwer Law International, (2005).

is unnecessary, and increasingly difficult to justify to users (parties, counsel and arbitrators alike) in ever-greater need of such guidance. The LCIA can make a unique contribution in this regard by making publicly available the wealth of learning that it is accumulating in its reasoned challenge decisions.<sup>222</sup>

As a result of these arguments, the LCIA announced its intent to publish its reasoned decisions on challenges to arbitrators, becoming the first arbitration body to do so.<sup>223</sup>

As Nick Gray and Deborah Crosbie argue:

... increased disclosure increases the probity and integrity of the whole system. Accordingly, reasoned decisions ought to show the distinction between disclosure and disqualification, thereby removing the common misconception that disclosure leads to disqualification, and at the same time increasing the confidence of all concerned to make and to receive disclosures appropriately.<sup>224</sup>

Recognizing the value of publication, especially as it relates to challenges to the central decision maker, in 2015 the International Chamber of Commerce announced that the ICC International Court of Arbitration will start explaining its reasons for many of the decisions it takes under the 2012 ICC Rules of Arbitration.<sup>225</sup> This includes but is not limited to “challenges to an arbitrator under Article 14 of the ICC Rules decisions to initiate replacement proceedings and subsequently to replace an arbitrator on the Court’s own motion under Article 15(2) of the ICC Rules.”<sup>226</sup>

### C. The Role of Brick-and-Mortar Courts

The traditional brick-and-mortar legal system performs an important function within the arbitration eco-system, as it is the national court system that both protects the fundamental right of individuals<sup>227</sup> and stands as the main enforcement mechanism of

---

<sup>222</sup> Gray, *Winds of change?*, *supra* note 146 at 6.

<sup>223</sup> However, the first decision was not published until 2011. See Thomas W. Walsh, Ruth Teitelbaum, *The LCIA Court Decisions on Challenges to Arbitrators: An Introduction*, LCIA, ARBITRATION INTERNATIONAL (2011).

<sup>224</sup> Gray, *Winds of change?*, *supra* note 146 at 7.

<sup>225</sup> See Silvia Farre, *ICC Court to Provide Reasoned Decisions*, DLA Piper, INTERNATIONAL ARBITRATION NEWSLETTER, (17 Dec, 2015).

<sup>226</sup> See Farre, *ICC Court To Provide Reasoned Decisions*, INTERNATIONAL ARBITRATION NEWSLETTER, 17 Dec, 2015.

<sup>227</sup> See Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140, Chapter 15, Section 15-3 (stating: “courts may also intervene to guarantee that

arbitration clauses and awards.<sup>228</sup> It is important to note, however, that within the context of the arbitration eco-system, the court system most often addresses issues of impartiality and independence only after the completion of the arbitration proceedings.<sup>229</sup> The timing of this review usually occurs in the context of a party seeking to prevent the court system from *enforcing* an award.<sup>230</sup> Within the court system, parties may challenge the award on various — although limited — grounds,<sup>231</sup> the most relevant to this discussion being that an arbitrator lacked impartiality and independence.<sup>232</sup> The challenge or resistance to the enforcement of *the award*, however, occurs under applicable national law — as such, the local courts are involved and national law is implicated in the discussion.<sup>233</sup> Fortunately, the international nature of the arbitration eco-system has prompted a convergence of approaches in some key areas, even when national laws are implicated. The convergence is especially in effect in the protection of fundamental rights, such as the parties’ right to an independent and impartial decision maker.

---

the minimum requirements of procedural fairness are fulfilled and exercise a supervisory function.”) *Id.*

<sup>228</sup> See Lew, COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION, *supra* note 140, Chapter 15, Section 15-2.

<sup>229</sup> For example, Section 10(a)(2) of the U.S. Federal Arbitration Act provides that “evident partiality” by the arbitrator is one basis upon which an arbitral award may be set aside. *See generally*, Gary Born and Claudio Salas, *The Different Meanings of an Arbitrator’s “Evident Partiality” Under U.S. Law*, KLUWER ARBITRATION BLOG, (March 20, 2013).

<sup>230</sup> This is because parties can challenge the appointment of an arbitrator under institutional rules, these challenges are, however, usually heard by the institution appointed. *See discussion, supra* Chapter 2, Section B(2).

<sup>231</sup> Both the New York Convention (United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards (1958) ) and domestic law, which tends to follow the UNCITRAL Model Law (UNCITRAL Model Law on International Commercial Arbitration 1985, amended in 2006) both provide very limited grounds. *See Lew, COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION, supra* note 140 at Chapter 25, 25-1, 25-15 (discussing the application of the Model Law to challenge awards) and Chapter 26, 26-3 (discussing the limited grounds for resisting enforcement) and 25-65 (discussing the limited grounds to refuse enforcement under the New York Convention).

<sup>232</sup> For example, when a party asserts the arbitrator lacked impartiality and as such, the New York Convention, Art. V(d): Irregularity in the composition of the arbitral tribunal or arbitral procedure must be considered.

<sup>233</sup> *See Lew, COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION, supra* note 140, Chapter 15, Section 2.6.

There is another area of convergence that is important to note up front, which is that in the context of contracts, most national laws recognize party autonomy as an essential contract principle<sup>234</sup> that should be limited in only the most extreme situations. National courts respect party autonomy, which in the case of arbitration<sup>235</sup> includes the parties' right to select the arbitration institution,<sup>236</sup> the applicable arbitration rules,<sup>237</sup> and the choice of law.<sup>238</sup> Consequently, national courts asked to enforce an award, but faced with questions regarding the impartiality and independence of the arbitrator, look to the parties' agreement, and the incorporated institutional rules and applicable national law, as the primary guide.<sup>239</sup> Assuming the agreement comports with the national law, the parties' agreement will be honored, and this agreement will include whatever rules concerning impartiality and independence are held by the arbitration institution the parties selected.<sup>240</sup> In instances where interpretations or standards are not specified within the arbitration rules or by the parties in their agreement, the court is left to look to national law.<sup>241</sup>

#### D. A Review of English Language Cases

In order to better assess the most frequently occurring cases challenging the impartiality and independence of the arbitrator (or arbitrators) a basic survey of international arbitration cases was conducted via Westlaw in February of 2019. A total of sixty-six

---

<sup>234</sup> See Alan Redfern and Martin Hunter, with N. Blackaby and C. Partasides, *LAW AND PRACTICE OF INTERNATIONAL COMMERCIAL ARBITRATION*, (4th Ed.), (2004) at p. 265 (“Party autonomy is the guiding principle in determining the procedure to be followed in an international commercial arbitration. It is a principle that has been endorsed not only in national laws, but by international arbitration institutions and organizations. The legislative history of the Model Law shows that the principle was adopted without opposition...” ) *Id.*

<sup>235</sup> See *id.* at 265.

<sup>236</sup> See *id.*

<sup>237</sup> See *id.*

<sup>238</sup> See *id.*

<sup>239</sup> See Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140, Chapter 8, Section 2.1.

<sup>240</sup> See Lew, *COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION*, *supra* note 140, Chapter 8, Section 2.6.

<sup>241</sup> See *id.*

cases were discovered, with fifty-two of the cases available in English or translated into English. Despite the small sample size, (which, as mentioned above, may in some sense be viewed as an encouraging sign that systems are working to prevent problems with judicial bias) several patterns may be discerned.

### 1. Ties to a Party's Legal Practice

The most common challenge to an arbitration award is an allegation that an arbitrator had ties to one of the law firms or is representing or engaged in legal practice for the benefit of one of the parties. From this broad category, several more specific issues can be pulled. Notably, practicing within the same U.S. Circuit: this has been found *not* to have compromised the impartiality of an arbitrator. Judges in these cases have focused even further and found that, generally, belonging to the same law firm as the counsel of one of the parties is not sufficient to indicate partiality or dependence. For example, an arbitrator was found unlikely to have been influenced by a junior associate from the same firm who acted as assistant legal counsel for one of the parties.

While tenuous connections to counsel's legal practice were found insufficient to establish compromised independence or impartiality, direct involvement with the parties could be enough to cast doubt upon an arbitrator's independence. For example, an award was thrown out after it was discovered that the arbitrator had also served as counsel to one of the parties in a different arbitration that had similar factual and legal issues. In other cases, simply acting as counsel at any point in time (the cases pulled indicated a window of six months to eight years) was considered enough to have an award thrown out.

### 2. Prior Arbitration Work

Another common issue was the frequency with which the same arbitrators were appointed to cases. When determining whether a decision maker is partial or dependent, the courts examined the frequency and nature of the disputes that an arbitrator had handled in the past. For example, upwards of fifty appointments by the same party to similar cases was determined to be corruptive to an arbitrator's impartiality. The court

also dismissed an arbitrator for being partial after he was appointed eleven times by the same party for similar arbitration matters. However, when an arbitrator was appointed by *different* parties on the recommendation of a single law firm, the court found no issues with that arbitrator's independence. Though the body of cases to support this trend is small, it appears that courts need both frequency and similarity to find major issues with an arbitrator's independence when it comes to prior arbitration work.

### 3. Other Professional Ties

In addition to ties to legal practice, other challenges were raised regarding the professional ties of the arbitrator to a member of one of the parties. The more ties an arbitrator has to a party, the more likely a court is to question his or her impartiality. Distant ties were usually not an issue. For example, an arbitrator who was a non-executive director of a competitor company of a party was not found to be partial and the award was enforced. Similarly, participation in the same arbitration conferences or meetings was not found to be an indicator of bias.

The presence of financial ties also heightens the probability that an arbitrator will be found dependent or partial. For example, failing to disclose that a party provided funding to an institution that employed the arbitrator as a director was considered to be an indication of dependency.

### 4. Social Ties

Several interesting themes emerged in the cases where an arbitrator had social ties to one of the parties. Familial connections from the first to the third degree were generally considered to be an indication of dependence and partiality. Non-familial social connections were more difficult to challenge, and connections like membership in different Rotary Clubs or belonging to the same nationality were considered insufficient to challenge an arbitrator's independence or impartiality.

### 5. Multiple Arbitrator Parties

For the sample of cases of available, most tribunals with more than one arbitrator had a three-arbitrator panel to render decisions. Finding the impartiality and independence compromised in one of the three arbitrators was not enough to throw out an award, however, reasonable doubt about the impartiality and independence of two out of the three arbitrators was enough to render the award unenforceable. Lastly, a common trend in arbitration panels was that the independence and impartiality of the group was not at issue if the award had previously been thrown out with the order to reconsider the award given.

#### E. Arbitration Overview Conclusions

As a creature of contract, with oversight of agreement and awards by the judiciary, arbitration has struggled with many of the questions that are now being considered in the deployment of technology in the online justice eco-system. This chapter explored arbitration as a means to provide lessons for the creation of the online justice environment. For the purposes of this thesis, the primary considerations are ones of dependence and partiality of the arbitrator and the mechanism and procedures put in place to mitigate the possibility of a dependent or partial arbitrator participant in a case. In general, institutions are in control of the arbitrators who are allowed to hear cases within their system. These institutions are thus in a position to ensure independence and impartiality through the creation of rules and processes that ensure compliance with the applicable laws and more generally with the RoL. The embrace by arbitration institutions of the importance of compliance with the RoL and their willingness to take steps to ensure that few awards are susceptible to challenge due to a dependent or partial decision maker, has led to arbitration being widely regarded as an appropriate and respected alternative to the justice system.

The chapter sought to explore if common terminology, such as dependence and partiality, could be distilled to a simple, clear, specific rule. Wide variance was discovered. While many argue this is a positive occurrence as it allows parties entering into an arbitration agreement to select from several institutions and to find the one that best matches their expectations, it does mean that the determination of partiality and dependence is handled on a case by case basis under several different frameworks.

Moreover, the increasing expectation that conflicts of interest can be eliminated by embracing the principle of full disclosure is a troubling one as many decision makers are being asked to over-disclose personal information. A balance can be struck, but it will need to ensure that the *reason* for asking for disclosure is not lost: that reason being that disclosure is the mechanism that allows parties to have adequate information when making a choice of the individual that is their decision maker. Arbitration is a creature of contract, and as such, in many ways parties create their own justice environment. However, parties should not be able to subvert or contract around key aspects of the RoL, such as an independent and impartial decision makers.

### CHAPTER 3: Online Dispute Resolution

Online Dispute Resolution (ODR) suffers from lack of definition. Originally, ODR was simply Alternative Dispute Resolution (ADR) in a digital environment: technology was used to support or replicate certain aspects of existing systems. Today, however, ODR is emerging as an eco-system of technology deployments that constitutes an area of the justice system in its own right. This chapter will briefly trace the shift from traditional ADR facilitated through the use of basic telecommunication technology, to justice provided in an online environment. First, the chapter will explore the basics of traditional ADR and briefly review the differences between private and public ADR systems. The chapter will then turn to the current issues related to judicial impartiality and independence within the current ODR eco-system. The chapter will finish with an overview of the newest conversations, occurring on an almost daily basis, about how smart contracts developers are embracing ODR and envisioning a new way to ensure the independence and impartiality of decision makers in the justice system.



## A. From Alternative Dispute Resolution to Online Dispute Resolution<sup>242</sup>

Some commentators have used ODR as a term referring to nothing more than the use of technology in an already existing judicial system.<sup>243</sup> This usage encompasses instances in which technology facilitates document receipt, search and storage, and sometimes simple digital communication. Other authorities have argued that the definition of ODR must include ADR *plus artificial intelligence* (a term the meaning of which is contentious in itself). While the literature has not come to a definitive consensus, for the purposes of this thesis the author maintains the following: an ODR system must involve the *resolution of a dispute* through the *use of technology* that is more than a mere communication platform *and must* present the *option* of a neutral facilitator (mediation) and/or a neutral decision maker (arbitration). Briefly, Online Dispute Resolution *must* refer to more than an online complaint system, even if the complaint is ultimately resolved inside the system.<sup>244</sup>

In general, the earliest versions of ODR were crafted around one of four processes: (1) consumer complaint resolution (sometimes known as negotiation),<sup>245</sup> (2) mediation,<sup>246</sup>

---

<sup>242</sup> The Overview Section is an updated research stream that builds upon some of my prior writings, which are listed in the ‘Details of Collaboration and Publication’ of the thesis. In each instance, I was the author writing the ODR specific sections of the paper.

<sup>243</sup> See e.g., Joseph W. Goodman, *The Pros And Cons Of Online Dispute Resolution: An Assessment Of Cyber-Mediation Websites*, 2 DUKE L. & TECH. REV. 1, 1-16 (2003).

<sup>244</sup> Of course, technology can and should be used to reduce the number of disputes and should be used to improve access to justice- these technology uses may be included in the ODR eco-system, but are important additions surrounding and supporting the *dispute* resolution system.

<sup>245</sup> eNegotiation is a process that uses “a negotiation support system including computers or other forms of electronic communication that enable parties to negotiate their own agreement.” The most important aspect is that the parties are in full control of accepting or rejecting the outcome. See Ernest Thiessen, Paul Miniato, and Bruce Hiebert, *ODR and eNegotiation*, in ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION, 329 (eds Wahab, Katsh, and Rainey (2012)).

<sup>246</sup> The use and definition of this term is controversial. Generally the term is thought to be defined as ““A voluntary process in which an impartial mediator actively assists disputants in identifying and clarifying issues of concern and in designing and agreeing to solutions for those issues.” DICTIONARY OF CONFLICT RESOLUTION, 278 (Douglas H. Yarn ed., 1999). The key is that a neutral third party is involved, parties’

(3) arbitration,<sup>247</sup> and/or (4) a combination of all three. In the current domestic-based online dispute resolution (ODR) environment, ODR service providers are generally involved in providing platforms that can be used for any or all of these processes, with various amounts of technology supporting or fully instantiating the system.<sup>248</sup> The digital transformation of justice that is occurring now is resulting in new uses of technology within the dispute resolution process.<sup>249</sup> I would like to suggest three broad categories to describe the technology being deployed: (1) Traditional ODR or “Online ADR” in which technology provides storage and communication platforms that are utilized in a traditional ADR process. (2) Advanced ODR, which uses the traditional model but with a higher level of technology integration — for example, when the negotiation is fully automated. (3) Adventurous ODR, which deploys the newest technology, such as smart contracts and blockchain, in such a manner that the technology becomes a unique and full participant in the process.

Dispute resolution platforms can be broadly divided into public and private categories.<sup>250</sup> Public platforms are part of — or closely adjacent to — an established court system. Such a platform replaces or supports existing court processes and is therefore held to the same standards and complies with the protections that exist within a brick-and-mortar court system. In contrast, private platforms and processes may be external to the court system, and thus may or may not rely upon the court system for enforcement of outcomes. Legislation may exist placing limitations upon private platforms — generally these will be narrow and specific, aimed at protecting

---

interests are considered and the outcome is often thought of as a win-win- or the product of compromise. *See generally* Noam Ebner, *E-Mediation*, in *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012).

<sup>247</sup> *See* Mohamed S. Abdel Wahab, *ODR and E-Arbitration*, in *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012).

<sup>248</sup> *See infra* Chapter 3, Section B.

<sup>249</sup> *See infra* Chapter 3, Section C.

<sup>250</sup> This is an update of an explanation and examination of these two system contained in my article. *See* Anjanette H. Raymond and Scott Shackelford, *Technology, Ethics And Access To Justice*, *supra* note 58, at 501-6 (2014) (explaining the earliest ODR platforms in these two areas).

fundamental rights. With these basic parameters in mind, let us examine each category in detail.

### 1. Private ODR Platforms

Private ODR providers can generally be divided into two types: self-contained and full-service.<sup>251</sup> A self-contained provider resolves disputes within a community, and as such the members of that community are controlled by and agree to the terms of service and associated agreements that regulate the community and the use of the platform. In contrast, a private full-service platform provides any and all parties access to an ODR mechanism.<sup>252</sup> This distinction is fundamental and highly important to the discussion as private providers are creatures of contract, and are—at the current time—largely unregulated.

When considering private platforms, no discussion can begin without examination of the eBay marketplace members' platform.<sup>253</sup> When eBay launched in 1995 as “AuctionWeb,” it was rather small and known to relatively few users, so buyers and sellers mostly interacted like friends in a community.<sup>254</sup> The creator, Pierre Omidyar, established a “Feedback Forum,” telling users to “[g]ive praise where it is due; make complaints where appropriate . . . .”<sup>255</sup> Since then, eBay has developed an eBay Resolution Center that facilitates communication between the buyer and the seller in the event that something goes wrong with an eBay marketplace transaction.<sup>256</sup> Buyers

---

<sup>251</sup> See *id.* 516-7 (describing early versions of public ODR platforms).

<sup>252</sup> See *id.*

<sup>253</sup> Of course, others also offer compliant services. For example, Amazon.com allows customers of third party sellers to file an A-to-z Guarantee claim if you purchased physical goods or eligible services on the Amazon.com website. See Amazon.com Author, *Help: File a Claim*, available at [http://www.amazon.com/gp/help/customer/display.html/ref=hp\\_left\\_sib?ie=UTF8&nodeId=200783750](http://www.amazon.com/gp/help/customer/display.html/ref=hp_left_sib?ie=UTF8&nodeId=200783750) (last visited Aug 8, 2019). A similar service is offered by Etsy. See Etsy author, *Escalate Your Case to Etsy's Trust & Safety*, available at <http://www.etsy.com/help/article/35> (last visited Aug. 8, 2019).

<sup>254</sup> See JACK GOLDSMITH & TIM WU, *WHO CONTROLS THE INTERNET?: ILLUSIONS OF A BORDERLESS WORLD*, 130 (2006).

<sup>255</sup> *Id.* at 131.

<sup>256</sup> See eBay Author, *eBay Resolving Buying Problems*, available at <http://pages.ebay.com/help/buy/protecting.html> (last visited Aug. 8, 2019).

and sellers can “open a case” within the eBay marketplace platform which allows the parties to communicate and attempt to resolve the issue amongst themselves.<sup>257</sup> After three days, if no resolution has been reached, a party can “escalate” the claim to an eBay representative.<sup>258</sup> eBay will then resolve the issue within 24 hours.<sup>259</sup> In some limited circumstances, eBay Buyer Protection may cover the purchase price plus original shipping.<sup>260</sup>

Of course, the customer can also post negative feedback on the seller and can contact the eBay Trust and Safety team to investigate issues within the marketplace. For example, eBay India has used online negotiation and mediation to resolve reputational disputes. Known as the eBay Community Court, this eBay India system draws on trusted eBay community members to resolve reputation-reporting disputes among the buyers and sellers within the eBay platform.<sup>261</sup> As platform designer and noted ODR authority Colin Rule highlights: “once each side has made their case (uploading material, etc.), the matter is put in front of a jury of twenty-one randomly selected eBay community members.”<sup>262</sup> While the “jury” members do have to meet stringent eligibility criteria, they are uncompensated volunteers.<sup>263</sup> The decisions are reviewed, when needed, for patterns and problems with outcomes and (in very rare cases) issues with the decisions of a particular member.<sup>264</sup> Although some tinkering was initially

---

<sup>257</sup> *See id.*

<sup>258</sup> *See id.*

<sup>259</sup> *See id.*

<sup>260</sup> The sellers redress options are different, mainly that the seller may both file a claim and seek to sell the original item still in its possession. See eBay Author, *e-Bay Help, Sellers Questions*, available at <http://pages.ebay.com/help/sell/questions/no-payment.html> (last visited Aug. 8, 2019).

<sup>261</sup> *See* Zhao Yun, Timothy Sze, Tommy Li and Chittu Nagarajan, *Online Dispute Resolution in Asia*, 510-11, in *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012)

<sup>262</sup> Colin Rule & Harpreet Singh, *ODR and Online Reputational Systems*, 193, in *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012).

<sup>263</sup> *See id.*

<sup>264</sup> *See id.*

needed,<sup>265</sup> the feedback on the Community Court has been very positive among all of the users.<sup>266</sup>

Self-contained dispute resolution platforms have advantages over other private systems in that the marketplace can respond to parties that fail to comply with dispute outcomes. For example, within the eBay platform, eBay can take action against parties who fail to comply by suspending accounts and/or allowing the winning party to post negative feedback about the non-compliant party.<sup>267</sup> Moreover, in many of these settings, the payment mechanism is internal to the marketplace. In these situations the payment portion of the marketplace can institute a delay in payment and/or can even reverse charges in the event that issues arise relating to the transaction. Systems that allow for delayed payments or that incorporate a “savings account” portion of the payment system allow for funds to be returned to the customer, without the need for merchant compliance with the mediated outcome.<sup>268</sup> The use of such an internal payment mechanism, especially one that incorporates a chargeback facility,<sup>269</sup> is an important means of compliance and is often heralded as one of the essential features of a successful private ODR platform.<sup>270</sup> These systems, coupled with the internal “trust mark,”<sup>271</sup> allow for a fully internal system of dispute resolution, relying on limited external support for enforcement of decisions.

---

<sup>265</sup> Such as the terms ‘jury’ and court’ which was changed to community review panels. *See id* at 181.

<sup>266</sup> *See id.*

<sup>267</sup> *See id.*

<sup>268</sup> For example, eBay. *See* Richelle Monfort, *The eBay Seller’s Guide to PayPal Claims, Disputes, and Chargebacks*, SaleHoo eCommerce Blog, (2019) available at <https://www.salehoo.com/blog/the-ebay-sellers-guide-to-paypal-claims-disputes-and-chargebacks> (last visited Aug. 8, 2019).

<sup>269</sup> “A chargeback is a charge that is returned to a payment card after a customer successfully disputes an item on their account transactions report.” Julia Kagan (rev.) *Chargeback*, INVESTOPEDIA, (2018).

<sup>270</sup> *See* Vikki Rogers, *Knitting the Security Blanket For New Market Opportunities*, 114-15, *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012).

<sup>271</sup> Trustmark is, for example, a top rated seller designation based on customer feedback. *See generally* Pablo Cortes, *Developing Online Dispute Resolution for Consumers in The EU: A Proposal for the Regulation of Accredited Providers* 19 *INT’L J. L. & INFO. TECH* 1 (2010)(describing the potential use of trustmarks in the E.U.).

Unfortunately, not all platforms offer the option to incorporate the chargeback process into the system. For example, some online environments lack financial integration,<sup>272</sup> while others exist in countries where chargebacks are impossible or limited under existing law.<sup>273</sup> Moreover, it is often the case, regardless of the existence of a chargeback facility, that some disputes arise too late in the payment process for this option to apply.<sup>274</sup>

The great benefit of offering arbitration through ODR is that it ties in higher-level enforcement mechanisms: Enforcement of the result of any arbitration process demands compliance with either the New York Convention<sup>275</sup> or — in the U.S. and many other jurisdictions — the UNCITRAL Model Law.<sup>276</sup> Both of these legal texts place restrictions upon the arbitration process<sup>277</sup> and require adherence to basic notions of due process,<sup>278</sup> including an impartial decision maker.<sup>279</sup> Although many disputes never reach the arbitration stage, the use of arbitration as the final recourse has the beneficial effect of bringing private dispute resolution systems under the umbrella of the rule of law. Thus, private full-service platforms that fail to use arbitration in the final stage are left with limited options in the area of enforcement, relying primarily upon the good will of the losing party, or their incentive to protect their reputation.

Self-contained dispute resolution platforms specific to a given company have been around for decades now.<sup>280</sup> The new, private full-service platforms that are now

---

<sup>272</sup> See *infra*, Table D.

<sup>273</sup> See *id.*

<sup>274</sup> See *id.*

<sup>275</sup> The New York Convention grants parties the right to resist enforcement in situations where “The composition of the arbitral authority or the arbitral procedure was not in accordance with the agreement of the parties, or, failing such agreement, was not in accordance with the law of the country where the arbitration took place” United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards, Art. V(d), (1958). See *discussion, supra*, Chapter 2, Section C.

<sup>276</sup> UNCITRAL Model Law, Art 12. See *discussion, supra*, Chapter 2, Section C.

<sup>277</sup> See *discussion, supra*, Chapter 2, Section C.

<sup>278</sup> See *id.*

<sup>279</sup> See *id.*

<sup>280</sup> See *e.g.*, Ethan Katsh and Orna Rabinovich-Einy, DIGITAL JUSTICE, TECHNOLOGY AND THE INTERNET OF DISPUTES, 1-11, Oxford University Press (2017); Robert

emerging are taking things to a new level by offering dispute resolution services that any party can use to settle any dispute, greatly expanding the sense in which these platforms can serve as an alternative to the traditional justice system. One of the significant players in the private ODR platform marketplace is Modria.<sup>281</sup> The creator of the Modria platform, Colin Rule, was also the creator of the original eBay and PayPal dispute resolution systems.<sup>282</sup> Modria allows parties to file any type of dispute.<sup>283</sup> The Modria software helps diagnose the problem by collecting and organizing information about the issue and suggesting solutions.<sup>284</sup> The software also enables the parties to discuss the matter online.<sup>285</sup> If the parties fail to communicate online, the software guides them to mediation and arbitration, provided by members of the Modria team.<sup>286</sup>

Impressively, the Modria platform has gained traction by securing several large businesses as clients.<sup>287</sup> In fact, the *Wall Street Journal* has called the Modria platform “[t]he small-claims court for the 21st century.”<sup>288</sup> More recently, the platform was purchased by Tyler Technologies and is working to expand its influence to a wider audience and even broader variety of disputes,<sup>289</sup> focusing particularly on providing online dispute resolution (ODR) for government and commercial entities.<sup>290</sup>

---

Ambrogi, *Is There a Future for Online Dispute Resolution for Lawyers?*, LawSites Blog, (April 11, 2016) available at <https://www.lawsitesblog.com/2016/04/future-online-dispute-resolution.html> (last visited Aug. 5, 2019).

<sup>281</sup> The original Modria platform no longer is a stand-alone project. For descriptions of the platform, see Katsh Digital Justice, *supra* note 251, at 35, 38, 144, 151.

<sup>282</sup> *See id.*

<sup>283</sup> *See id.*

<sup>284</sup> *See id.*

<sup>285</sup> *See id.*

<sup>286</sup> *See id.*

<sup>287</sup> *See id.*

<sup>288</sup> Deborah Gage, *VC Dispatch, Modria An Online ‘Small-Claims Court for 21st Century’*, WALL ST. J. ONLINE (Nov 2012).

<sup>289</sup> The Modria platform was acquired by Tyler Technologies, and has greatly expanded its influence and reach. See Gabrielle Orum Hernández, *The Modria platform was acquired by Tyler Technologies*, Law.Com, (June 2, 2017) and has greatly expanded its influence and reach. More information Tyler Technologies website available at <https://www.tylertech.com/products/Modria> (last visited Aug. 5, 2019).

<sup>290</sup> As such, it now straddles both the private platform provider and the public provider. More information Tyler Technologies Website available at <https://www.tylertech.com/solutions-products/modria> (last visited Aug. 5, 2019).

Modria/Tyler Technologies is not the only ODR platform providing private ODR services; however, most of its competitors deploy more technology while providing fewer choices in terms of resolution mechanisms. Table E summarizes some of the notable players in this marketplace.

Table D: Currently Viable ODR Platforms

Name	Description	Key Features
<a href="#">Cybersettle</a>	Cybersettle's patented system helps parties to a dispute, such as a claim against an insurance company, reach agreement much faster and at a much lower cost than they could using traditional negotiation methods.	<ul style="list-style-type: none"> <li>● Patented “double-blind” technology allows parties to submit confidential offers and demands which are never disclosed to the opposing party</li> <li>● Data Security &amp; Encryption</li> <li>● Bulk submission of cases through proprietary API</li> </ul>
<a href="#">CPR Institute for Dispute Resolution</a>	Mediation and arbitration services through secure email.	<ul style="list-style-type: none"> <li>● CPR Dispute Resolution Services utilizes a secure email solution (Hewlett Packard Enterprise Secure Cloud)</li> </ul>
<a href="#">Cibertribunal peruano</a>	Arbitration and mediation via video calls, email, and online chat.	No longer available
<a href="#">Camera Arbitrale di Milano - Risolvionline.com</a>	Italian Online mediation, primarily for retail purposes. Consumers initiate complaints; vendors are invited by the platform to participate; an expert provides a non-binding evaluation.	<ul style="list-style-type: none"> <li>● Provided by the Milan Chamber of Arbitration</li> </ul>
<a href="#">Netneutrals.com</a>	NetNeutrals.com helps consumers and businesses quickly resolve disputes. It offers the choice of a free direct negotiation forum, or the option to request a skilled Neutral with technical expertise to join the discussion and help the parties reach resolution.	<ul style="list-style-type: none"> <li>● Working with Ebay</li> <li>● Run by DeMars &amp; Associates, Ltd. (DMA), national alternative dispute resolution consulting firm</li> </ul>



Name	Description	Key Features
<a href="#">Better Business Bureau Online Complaint System</a>	<p>Consumers initiate complaints; vendors are invited by the platform to participate. Complaints are usually closed within 30 calendar days. In some cases, BBB mediation or arbitration may be offered to assist in resolution.</p>	<p>Will only handle a narrow set of disputes, generally governed by the BBB</p>
<a href="#">Arbitranet</a>	<p>Brazil Online Arbitration Provider that promises a decision in 100 days. The online system chooses an arbitrator based on the indications of the parties</p>	<p>Arbitration Clause available</p>
<a href="#">ADNDRC</a>	<p>ADNDRC was jointly established by the China International Economic and Trade Arbitration Commission (CIETAC) and the Hong Kong International Arbitration Centre (HKIAC) in 2002.</p>	<p>Downloadable Domain Name Complaint Forms</p>
	<p>As a dispute resolution service provider accredited by ICANN, the Centre administers domain name dispute resolution proceedings through its four offices not only under the UDRP, but also under the following ICANN policies:</p>	<ul style="list-style-type: none"> <li>• Fees based on number of domain names involved in the complaint and how many panelists decide the case</li> </ul>
	<p>Uniform Rapid Suspension System (URS)</p>	
	<p>Trademark Post Delegation Dispute Resolution Procedure (TM-PDDRP)</p>	
	<p>Registrar Transfer Dispute Resolution Policy (TDRP)</p>	
<p>Sunrise Dispute Resolution Policy (SDRP)</p>		

Name	Description	Key Features
	<p>Charter Eligibility Dispute Resolution Policy (CEDRP)</p> <p>ADNDRC is also a Dispute-Resolution Service Provider approved by the Organization for the Promotion, Protection and Progress of Frogans Technology. They are a non-profit standards developing organization whose purpose is to hold, promote, protect and ensure the progress of the Frogans technology in the form of an open standard for the Internet, available to all, free of charge.</p>	
<p><a href="#">SmartResolution</a></p>	<p>SmartResolution is a platform designed for professionals to offer their own customized ODR services.</p>	<p>Free, Open-source</p> <p>Specific modules available for special dispute types (ex. maritime)</p>
<p><a href="#">Modria</a></p>	<p>Diagnose the issue by gathering relevant information</p> <p>Enable an online negotiation between the parties</p> <p>Provide access to mediation if needed</p> <p>Refer the case for an evaluation</p>	<p>Built-in chat</p> <p>Document and email templates</p> <p>Pre-built resolution flows</p> <p>Neutral parties can facilitate or decide an issue</p> <p>Real-time dashboards, pre-configured reports, and data export options</p>
<p><a href="#">OnlineMediators.com</a></p>	<p>This directory is for searching and contacting dispute resolution professionals</p>	<p>Online Directory for ODR providers</p>

Name	Description	Key Features
	for the purpose of acquiring dispute resolution services.	Does not provide its own services

Author compilation of online website information (Up-to-date August 2019)

As can be seen from the chart, technology is a large part of many of the existing ODR systems.<sup>291</sup> And the technology embedded within the system can be used for a host of activities. For example, some negotiation support systems, such as the family mediation support platform known as DEUS,<sup>292</sup> allow the parties to communicate within the platform, to exchange offers, and to view graphs and other interactive tools that compare the current offer to initial offers and the expectations that lead to the offer.<sup>293</sup> A few private systems exist that allow for technology-based metrics to be employed in online mediation as intelligent negotiation support.<sup>294</sup> In these systems, parties rank and value each issue within the dispute by allocating a sum amongst all issues.<sup>295</sup> The platform then uses the numbers to optimize each other's identified interests and to make

<sup>291</sup> See Arno R. Lodder & John Zeleznikow, *Artificial Intelligence and Online Dispute Resolution*, 65, in *ONLINE DISPUTE RESOLUTION: THEORY AND PRACTICE, A TREATISE ON TECHNOLOGY AND DISPUTE RESOLUTION*, (eds Wahab, Katsh, and Rainey, 2012). "Modria's goal is to resolve about 90% of cases through software, without humans." Deborah Gage, VC Dispatch, *Modria An Online 'Small-Claims Court for 21st Century'*, WALL ST. J. ONLINE (Nov 2012).

<sup>292</sup> See John Zeleznikow and Emilia Bellucci, *Building Negotiation Decision Support Systems by Integrating Game Theory and Heuristics*, *DECISION SUPPORT IN AN UNCERTAIN AND COMPLEX WORLD: THE IFIP TC8/WG8.3 INTERNATIONAL CONFERENCE 873-74* (2004). DEUS is a template-based system that helps mediators understand the extent of the issues in dispute. The goal of DEUS was to separate the people from the property at issue. DEUS helps mediators understand what issues are in dispute and the extent of the dispute over these issues. The system, however, was not concerned with the principles of justice. Its goal is to indicate the range and significance of issues in dispute. *See id.* 873-74. The choice of the developers in naming the platform DEUS highlights the power that programmers' perhaps hope that such systems may wield, which bears on the question of whether this is in the best interests of the justice system.

<sup>293</sup> See Lodder, *Artificial Intelligence and Online Dispute Resolution*, *supra* note 262 at 65.

<sup>294</sup> These platforms, such as Smartsettle, assist the parties in clarifying interests, identifying trade-offs, and generating optimal solutions. *See* Lodder, *Artificial Intelligence and Online Dispute Resolution*, *supra* note 262, at 65.

<sup>295</sup> *Id.*

suggestions of a fair outcome.<sup>296</sup> These platforms, such as Smartsettle,<sup>297</sup> assist the parties in clarifying interests, identifying trade-offs, and generating optimal solutions.<sup>298</sup> On the other hand, platforms such as Cybersettle and Clicknsettle assist parties in the entire negotiation process with no human intervention whatsoever.<sup>299</sup>

Private platforms do have one immediate drawback: the outcomes can be difficult to enforce. As creatures of contract, ODR platforms rely upon two common types of enforcement: community enforcement rules (such as chargebacks or removal from the site) that are agreed upon as part of the terms of service, and/or the use of local legal doctrine. In the latter instance, as we discussed above, it is the use of arbitration that is the key to the ability to enforce outcomes. In general, mediation outcomes are essentially new contracts created by the parties during the mediation.<sup>300</sup> Hence, mediation outcomes are enforced like any other contract<sup>301</sup> — through a lengthy court proceeding.<sup>302</sup> However, as previously discussed, arbitration is different.<sup>303</sup> Arbitration is an alternative to the court that provides court enforcement of outcomes, so long as the arbitration complies with certain fundamental due process rights (such as impartiality of the decision maker).<sup>304</sup> In fact, arbitration awards are treated like court judgments, so long as a simple legal process of filing the award is followed.<sup>305</sup> Many traditional private ODR systems rely upon arbitration and thus interact closely with the justice system.<sup>306</sup> But, some private full service platforms are more narrowly

---

<sup>296</sup> *See id.*

<sup>297</sup> *See* Smartsettle, <http://www.smartsettle.com/> (last visited Aug. 8, 2019).

<sup>298</sup> *See* Lodder, *Artificial Intelligence and Online Dispute Resolution*, *supra* note 262, at 65.

<sup>299</sup> *See* Cybersettle, <http://www.cybersettle.com> (last visited Aug. 8, 2019).

<sup>300</sup> *See* Ellen E. Deason, *Procedural Rules for Complimentary Systems of Litigation and Mediation- Worldwide*, 80 NOTRE DAME L. REV. 553, 580 (2005) (discussing the manner of mediation enforcement). Do note, the *United Nations Convention on International Settlement Agreements Resulting From Mediation* may change the ability to enforce mediation outcomes, globally. *See supra* Chapter 1, footnote 5.

<sup>301</sup> Deason, *Procedural Rules for Complimentary Systems of Litigation*, *supra* note 271 at 580.

<sup>302</sup> *See id.*

<sup>303</sup> *See id.* at 577-579; *supra* Chapter 2, Section Introduction.

<sup>304</sup> *See, supra* Chapter 2, Section Introduction.

<sup>305</sup> *See, supra* Chapter 2, Section Introduction.

<sup>306</sup> *See e.g., supra*, Chapter 2, Section A(1) (e-Bay market place discussion).

constructed, relying instead upon only negotiation and/or mediation.<sup>307</sup> Because these platforms operate without the option of final and binding arbitration as the final dispute resolution step, these platforms are without a judicial enforcement mechanism. In these instances, communities — as discussed above — may enforce outcomes through payment or reputation mechanisms that are part of the system.

If arbitration is not included in an ODR process, the parties are vulnerable to potentially biased dispute resolution that is designed in contract. Without the courts' influence and insistence upon adherence to the rule of law when requesting the enforcement of an outcome, platform designers are fully within their rights to design a process that suits them, and there are few other safeguards to guarantee that this will be compliant with the rule of law. To date, not many platform designers have taken advantage of this situation, but there are examples where basic RoL protections have been rejected.<sup>308</sup> For example, PeopleClaim, a company whose tagline is "Online Dispute Resolution for Any Problem."<sup>309</sup> PeopleClaim allows individuals to file claims and to enter a system of negotiation with the other party in an online environment.<sup>310</sup> In essence, this stage is an online communications portal for the parties to begin discussing and resolving their dispute. From here however, the system becomes less RoL-focused. Outcomes depend on individuals agreeing to a suggested solution. Suggestions can be submitted by multiple parties: as a member of the justice community, you are encouraged to "Find a case with the RHU (RHU Coin) icon. Suggest your idea for fair resolution. Receive a 'Yes' vote from the community. Win five free RHU coins.\*"<sup>311</sup> It's crucial feature is

---

<sup>307</sup> See e.g., eNegotiation, *supra* footnote 216.

<sup>308</sup> At the time of writing, People Claim comes to mind. See PeopleClaim website available at <https://www.peopleclaim.com/> (last visited Aug. 9, 2019).

People Claim allows parties to submit claims without verification of the claim, allows public shaming or businesses and individuals, allows an online jury- each of whom volunteer to participate- without screening or conflict disclosures- to decide the outcome based on a popularity of the party type process. See *id.*

<sup>309</sup> PeopleClaim website available at <https://www.peopleclaim.com/> (last visited Aug. 9, 2019).

<sup>310</sup> See *id.*

<sup>311</sup> *Id.* It is important to note, the RHU have no immediately transferable value to recognized currency. Instead, like many cryptocurrencies the coin are valued based within the community and those that extend value to the coin. It is too early in the development of these crypto based systems to determine the success or failure of such a system. However, former chairman of the U.S. Federal Reserve, Alan Greenspan is

that the key to winning the RHU coins is to suggest a solution that the most people like.<sup>312</sup> It should be noted, I have written about PeopleClaim before,<sup>313</sup> and they reached out to me to help adjust the system. In fact the asterisk in the statement above was added after our discussion about issues with this “informed consensus”<sup>314</sup> system. The asterisk leads the attentive user to learn that the platform can reject payment to dishonest individuals within the system. The developers responded to other items of feedback as well, by not calling the outcome arbitration, not advertising it as a resolution that could be enforced in court, and clarifying the nature of the platform as really a communications opportunity for parties. However, against this author’s recommendation, PeopleClaim elected to continue to use a “name-and-shame” approach to claim filing.<sup>315</sup> Name-and-shame is the idea that big businesses are more attentive than ever to bad social media — a point PeopleClaim exploits. PeopleClaim posts all filed claims online, in a readily searchable database that is easily discovered by a simple Google search.<sup>316</sup> PeopleClaim has no mechanism in place to ensure the claim is filed based upon any real harm; in fact a quick search through the claim database suggests no one at PeopleClaim follows up with individuals who are merely venting, often in a very negative, unverified manner, about a company or individual.<sup>317</sup> Public name and shaming is a poor mechanism to encourage deliberative, non-adversarial communication and is a disaster without efforts to verify the accuracy of the filed information. There are websites that note issues with small business owners being

---

not a fan of the traditional cryptocurrencies. See Sujha Sundararajan, *Greenspan Likens ‘Irrational’ Bitcoin to Revolutionary War Currency*, Coin Desk, (Dec. 7, 2017) (discussing the volatility of the potential market).

<sup>312</sup> PeopleClaim website available at <https://www.peopleclaim.com/> (last visited Aug. 9, 2019).

<sup>313</sup> See Scott J. Shackelford and Anjanette H. Raymond, *Building the Virtual Courthouse: Ethical Considerations for Design, Implementation, and Regulation in the World of ODR*, WISCONSIN LAW REVIEW 615 (2014).

<sup>314</sup> PeopleClaim website available at <https://www.peopleclaim.com/> (last visited Aug. 9, 2019).

<sup>315</sup> See *id.*

<sup>316</sup> See PeopleClaim website available at <https://www.peopleclaim.com/faq.aspx?cID=1> (last visited Aug. 9, 2019).

<sup>317</sup> See PeopleClaim website available at <https://www.peopleclaim.com/trial/directory.aspx> (last visited Aug. 9, 2019).

wrongfully accused, without opportunity to defend themselves, yet the filing about them remains widely available.<sup>318</sup>

The PeopleClaim case illustrates two major dangers that the author contends should be avoided. First, we cannot allow intentional mis-information to be disseminated about the services that the platform provides. Second, public name-and-shame situations should never be allowed without some mechanism of verification.

## 2 Public ODR Platforms

Many of the benefits of private platforms can be replicated in the public online justice environment. Public ODR platforms, however, have distinct advantages in terms of enforcement of outcomes. Public-based and/or judicially-supported ODR platforms have several enforcement mechanisms—the simplest of which is to have the local courts assist in enforcement should the losing party fail to comply.<sup>319</sup> With this type of mechanism, no feedback, accreditation or account restrictions are necessary to encourage compliance. Instead, users agree to abide by particular rules and laws and allow a publicly supported ODR system to resolve issues.<sup>320</sup> Moreover, the close ties between the platform and the court system demand the ODR platform comport with local law,<sup>321</sup> which in most instances guarantees that it meets the requirements of the rule of law.<sup>322</sup> As such, public systems are assumed to be compliant with the basic expectations of any other justice environment.

One of the early advances in public ODR platform development occurred in Mexico. Called Concilianet, the platform is both hosted and supported by the government via

---

<sup>318</sup> See Site Jabber Reviews, available at <https://www.sitejabber.com/reviews/peopleclaim.com> (last visited Aug. 9, 2019).

<sup>319</sup> See e.g., the Civil Resolution Tribunal in British Columbia, Canada, (resolving via an ODR process, motor vehicle injury dispute and small claims (among others), available at <https://civilresolutionbc.ca/> (last visited Aug. 9, 2019).

<sup>320</sup> See *id.*

<sup>321</sup> See *id.*

<sup>322</sup> See *id.*

the Federal Attorney's Office of Consumer (PROFECO).<sup>323</sup> The Concilianet platform has been created to resolve disputes between registered merchants<sup>324</sup> and their customers.<sup>325</sup> The process is remarkable as it: (1) is straight forward; (2) has no direct filing costs; (3) allows online and in-person filing; (4) uses verified forms of personal identification for registration of a claim; (5) complies with local law on data retention and protection; (6) has trained assistants to assist the consumer through the process; (7) uses a secured electronic data capture and virtual courtroom throughout the process, and; (8) is able to resolve the majority of claims within a short, consumer-driven timeline.<sup>326</sup> In fact, even merchants not registered with the service can have a claim filed against them, but the filing must be done in person (as use of the online platform requires consent via registration) at a local satellite office.<sup>327</sup> It is also worthy of note that the entire system is both voluntary and produces a non-binding outcome that allows the consumer to retain his right to his day in court.<sup>328</sup>

One of the better examples of a public-supported, multi issue- human centered focused ODR platforms currently exists within the province of British Columbia, Canada.<sup>329</sup> Unlike previously described systems, the British Columbia ODR platform engages the business once the complaint is filed.<sup>330</sup> At this initial stage the parties to the dispute agree to proceed with the resolution process and also agree to have the outcome be final and binding (i.e., arbitration). Although the platform does not assist in the recovery, should the losing party fail to comply the final award can be taken to the judicial system

---

<sup>323</sup> Government of Mexico, Federal Attorney's Office of Consumer, Concilianet, website, available at <https://concilianet.profeco.gob.mx/Concilianet/> (last visited Aug. 9, 2019).

<sup>324</sup> For a full list, see Concilianet, list of participants, available at [https://concilianet.profeco.gob.mx/Concilianet/proveedores\\_que\\_concilian.jsp](https://concilianet.profeco.gob.mx/Concilianet/proveedores_que_concilian.jsp) (last visited Aug. 9, 2019).

<sup>325</sup> See *id.*

<sup>326</sup> See *id.*

<sup>327</sup> See *id.*

<sup>328</sup> See Author, Concilianet FAQ, <https://concilianet.profeco.gob.mx/Concilianet/comoconciliar.jsp> (last visited Aug. 9, 2019).

<sup>329</sup> The Civil Resolution Tribunal in British Columbia, Canada,(resolving via an ODR process, motor vehicle injury dispute and small claims (among others), information available at <https://civilresolutionbc.ca/> (last visited Aug. 9, 2019).

<sup>330</sup> See *id.*



for enforcement. In the first seven months, “the system . . . handled nearly 14,000 small-claims cases. Roughly 85 percent of the 700 cases resolved to date were settled. Only 12 went to decision at the tribunal.”<sup>331</sup> Since these initial figures the platform has grown, frankly at an exponential rate. As of July 2019, the CRT Statistics Snapshot reports Small Claims explorations 50,337, Strata Property explorations 25,302, Motor Vehicle Accidents and Injuries explorations 2,180, Societies and Cooperative Associations explorations 79.<sup>332</sup>

Interestingly, some of the modern platforms are developed with an eye toward taking the technology to the people. For example, a pilot program in San Francisco sought to reduce the number of evictions among its low-income population.<sup>333</sup> The pilot program created a platform that could be accessed online, and intentionally expanded its accessibility by making it available via iPad.<sup>334</sup> The Bar Association of San Francisco then hired law students to go into the affected communities and provide justice on sight.<sup>335</sup> The application was designed with an eye toward compliance with existing landlord–tenant laws in San Francisco, so the Bar Association could be confident that even the newest of law students would be able to competently and accurately provide this important legal service.<sup>336</sup>

Initiatives such as the one in San Francisco have provided important examples of success that may embolden more cities to use technology to improve access to justice through the ODR eco-system. The team who designed the British Columbia software mentioned above went on to design the largest public ODR system to be launched to

---

<sup>331</sup> Author, *British Columbia ODR System Handles 14,000 Cases In First 7 Months*, ABA NEWS, (Feb. 4, 2018) [https://www.americanbar.org/news/abanews/aba-news-archives/2018/02/british\\_columbiaodr.html](https://www.americanbar.org/news/abanews/aba-news-archives/2018/02/british_columbiaodr.html) (last visited Aug. 9, 2019).

<sup>332</sup> The Civil Resolution Tribunal in British Columbia, Canada, CRT Statistics Snapshot – July 2019, available at <https://civilresolutionbc.ca/crt-statistics-snapshot-july-2019/> (last visited Aug. 9, 2019).

<sup>333</sup> See Simon Boehme, *Housing Dispute? Handle It Quickly Online*, BAR ASSOCIATION OF SAN FRANCISCO, (Fall 2018) available at <http://www.sfbar.org/forms/sfam/q32018/conflict-intervention-service-2-SFAM-Q318.pdf> (last visited Aug. 5, 2019).

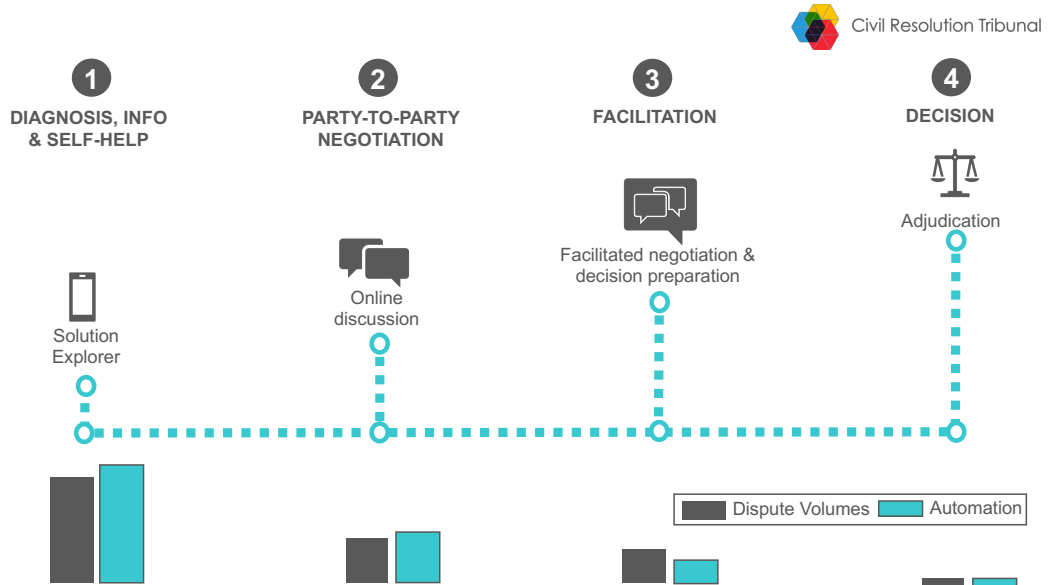
<sup>334</sup> See *id.*

<sup>335</sup> See *id.*

<sup>336</sup> See *id.*

date.<sup>337</sup> As part of Canada’s commitment to improving access to justice and embracing the use of technology, the Civil Resolution Tribunal<sup>338</sup> was started in 2016 to handle condominium disputes. It quickly expanded to handle small claims and most recently implemented a process to handle car accident and personal injury disputes.<sup>339</sup>

Shannon Salter, Civil Resolutions Tribunal Chair, describes the system:<sup>340</sup>



Civil Resolution Tribunal Canada, (Japan APEC Workshop 2018)

As can be seen from the diagram, the system involves multiple technology deployments to improve individuals’ access to the justice system, including an entire section of the platform and website that assists individuals in understanding the law and their rights within the law, and directs them toward options for resolution.<sup>341</sup> For many involved in

<sup>337</sup> The Civil Resolution Tribunal in British Columbia, Canada, (resolving via an ODR process, motor vehicle injury dispute and small claims (among others)), information available at <https://civilresolutionbc.ca/> (last visited Aug. 9, 2019).

<sup>338</sup> *See id.*

<sup>339</sup> *See id.*

<sup>340</sup> Shannon Salter, Civil Resolution Tribunal presentation, APEC ODR Conference (November 2018).

<sup>341</sup> *Id.* Called Solution Explorer, the platform is a free public legal information and tool with guided pathways , interactive question and answers, resolution or preparation for CRT process and quarterly updates on content. *See id.*

ODR design, the first stage of the ODR eco-system is dispute *avoidance*, meaning that the system should assist individuals before a problem becomes a dispute.<sup>342</sup> This is reflected in the priority given to “self-help” as the first tier of the service. In the self-help process individuals are allowed to ask questions and receive some basic legal information based on the entered description of the issue. As can be seen in the diagram, the process can then move into a party-to-party negotiation phase where the parties communicate, without intervention, within the platform. Should the parties desire, the platform can then assist in the negotiation process. The process is completed when a decision is provided to the parties. In general, many of these systems envision the use of a human as the facilitator, meaning a unique human not party to the dispute enters into the communications environment. However, as will be further explored in Chapter 4, as this technology develops it is quite possible the facilitation will occur without a human in the loop. While the CRT is one of the newer platforms, it is widely regarded as successful,<sup>343</sup> so much so that the Civil Resolution Tribunal Act was amended to give the CRT “jurisdiction over motor vehicle injury disputes \$50,000 and under, including accident benefit claims and minor injury determinations.”<sup>344</sup>

In the United States, the National Center for State Courts<sup>345</sup> notes that “ODR is utilized in only a small fraction of U.S. courts” and that “currently the U.S. lags well behind

---

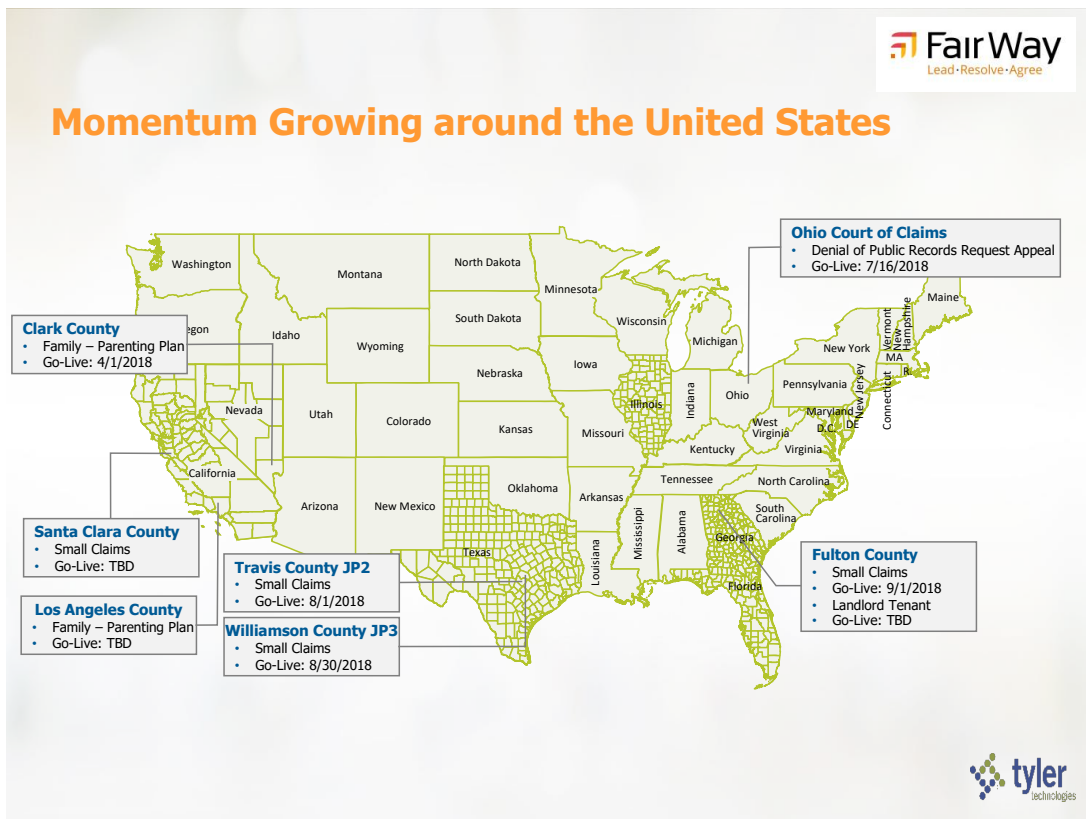
<sup>342</sup> See *id.* For further discussion, see Katsh, *DIGITAL JUSTICE*, *supra* note 251, at 17. Adherence to dispute avoidance is one reason why the ODR community rejects too narrow of definition. The author believes, as noted in Chapter 1, the better choice is to allow the ODR community to embrace an eco-system approach to regulation.

<sup>343</sup> CRT Statistics Snapshot (December 2018) available at <https://civilresolutionbc.ca/crt-statistics-snapshot-december-2018/> (last visited January 17, 2019) providing a total of 8,810 disputes with “75% are likely to recommend the CRT to others”. Participant Satisfaction Survey — April to December 2018 (Jan 2018) available at <https://civilresolutionbc.ca/participant-satisfaction-survey-april-december-2018/> (last visited Aug. 5, 2019).

<sup>344</sup> Effective January 1, 2019. See Lauryn Kerr, *The CRT’s Legislation is Changing*, (Dec. 2018) available at <https://civilresolutionbc.ca/the-crts-legislation-is-changing/> (last visited Aug. 5, 2019).

<sup>345</sup> “The NCSC is the organization courts turn to for authoritative knowledge and information, because it collaborates closely with the Conference of Chief Justices, the Conference of State Court Administrators, and other associations of judicial leaders.” NCSC Website, About US, available at <http://www.ncsc.org/About-us.aspx> (last visited Aug. 9, 2019).

Europe and the Pacific Rim in adoption of ODR.”<sup>346</sup> In fact, the report highlights: “the most ambitious and successful ODR initiatives today are going on outside the United States.”<sup>347</sup> That is not to write that ODR is non-existent in the U.S.: early experiments are appearing alongside calls for further research. In fact, several organizations have produced guidance for ODR implementation which will be drawn upon in the examination of standards later in the chapter. Rhys West of Fairway Consulting has provided a diagram of the ODR systems under development as of 2018:<sup>348</sup>



Tyler Technology Presentation (APEC Workshop, Japan 2018)

<sup>346</sup> Joint Technology Committee, Conference of State Court Administrators, the National Association for Court Management and the National Center for State Courts, *Case Studies In ODR For Courts: A View From The Front Lines*, JTC Resource Bulletin, The National Center for State Courts Publication, 1, (Nov. 2017).

<sup>347</sup> *Id.*

<sup>348</sup> Rhys West, Online Dispute Resolution, Expanding Access to Justice, APEC ODR Conference, Materials, (November 2018).

The PEW Charitable Trust has been researching and writing in the area<sup>349</sup> and has just produced a call for further research,<sup>350</sup> as well as the possible creation of a national standard for ODR in the United States.<sup>351</sup>

One of the most ambitious ODR-based systems to date was created by the European Union, when it passed a first-of-its-kind ADR Directive and corresponding ODR Regulation to facilitate the widespread use of ODR across the E.U.<sup>352</sup> The ADR Directive creates a procedure covering all contractual disputes in every market sector (e.g., travel, banking, dry cleaning), and in every Member State.<sup>353</sup> Similar to other systems,<sup>354</sup> the ADR Directive requires traders to inform consumers about the availability of ADR,<sup>355</sup> both on their websites and in their general terms and conditions,<sup>356</sup> and requires ADR entities to meet quality criteria which guarantee that they operate in an effective, fair, independent and transparent way.<sup>357</sup>

Complementing the ADR Directive, an ODR Regulation enables E.U. consumers and traders to submit disputes arising from online purchases to an online dispute resolution

---

<sup>349</sup> See Erika Rickard, *Online Dispute Resolution Offers a New Way to Access Local Courts*, Pew Charitable Trust Fact Sheet, (Jan. 4, 2019).

<sup>350</sup> See Author, *Pew Welcomes Discussion For A National ODR Standards Body*, ARTIFICIAL LAWYER, (March 6, 2019).

<sup>351</sup> The author is potentially part of the research agenda (as of May 2019).

<sup>352</sup> The ADR Directive will not cover the sectors of health and education as these areas are already heavily regulated and disputes arise are often of a more complex nature. EUROPA publication, *A Step Forward For EU Consumers: Questions & Answers On Alternative Dispute Resolution And Online Dispute Resolution*, Press Release, (Mar. 2013) [hereinafter EUROPA, Press Release]. The original website for the platform is available at [http://ec.europa.eu/consumers/redress\\_cons/adr\\_en.htm](http://ec.europa.eu/consumers/redress_cons/adr_en.htm) (last visited Aug. 5, 2019).

<sup>353</sup> See EUROPA, Press Release, *supra* note 323.

<sup>354</sup> Anjanette H. Raymond *Yeah, But Did You See the Gorilla? Creating and Protecting an 'Informed' Consumer In Cross Border Online Dispute Resolution*, 19 HARVARD NEGOTIATION LAW REVIEW 129, 129-171 (Spring 2014)..

<sup>355</sup> See Directive 2013/11/EU Of The European Parliament And Of The Council of 21 May 2013 on Alternative Dispute Resolution For Consumer Disputes, at Art 7 [hereinafter "ADR Directive"].

<sup>356</sup> See *id.* at Art 7.

<sup>357</sup> See ADR Directive, *supra* note 326 at arts. 15-16; EUROPA, Press Release, *supra* note 323.

mechanism.<sup>358</sup> The ODR Regulation attempts to create an E.U.–wide dispute resolution platform (“ODR platform”)<sup>359</sup> that will link all the national ADR entities.<sup>360</sup> This single entry point is designed to be a user-friendly and interactive website, available in all E.U. official languages and free of charge.<sup>361</sup> In practice, consumers encountering a problem with an online purchase will eventually be able to submit a complaint online in the language of their choice through the ODR platform.<sup>362</sup> The ODR platform will notify the trader that a complaint is lodged against him.<sup>363</sup> The consumer and the trader will then agree on which ADR entity to use to solve their dispute.<sup>364</sup> Importantly, as will be discussed later, it is arguable at this point as to whether the ADR submission agreement is created between the parties.<sup>365</sup> When they agree, the chosen ADR entity will receive the details of the dispute via the ODR platform.<sup>366</sup> Because the ODR platform will be connected to the *national* ADR entities,<sup>367</sup> the platform will allow national ODR advisors to provide general information to consumers on consumer rights and redress in relation to online purchases,<sup>368</sup> and assist with the submission of

---

<sup>358</sup> EUROPA, Press Release, *supra* note 323.

<sup>359</sup> Regulation (EU) No 524/2013 Of The European Parliament And Of The Council of 21 May 2013 On Online Dispute Resolution For Consumer, at Art 5. [hereinafter ‘ODR Reg.’]

<sup>360</sup> EUROPA, Press Release, *supra* note 323.

<sup>361</sup> *See* ADR Directive, *supra* note 326, at art 9; EUROPA, Press Release, *supra* note 323.

<sup>362</sup> *See* ODR Reg., *supra* note 330 at Art 7; ADR Directive, *supra* note 326, at art. 17; EUROPA, Press Release, *supra* note 323.

<sup>363</sup> EUROPA, Press Release, *supra* note 323.

<sup>364</sup> *See* ODR Reg., *supra* note 330 at Arts 5 and 8; EUROPA, Press Release, *supra* note 323. *See also*, Raymond Yeah, *But Did You See the Gorilla?*, *supra* note 325.

<sup>365</sup> Pre-dispute arbitration clauses are generally thought of as unfair contract terms when E.U. consumers and their corresponding consumer protection laws are implicated in the transaction. This is NOT the U.S. position. *See* Ronald A. Brand, *Party Autonomy and Access to Justice in the UNCITRAL Online Dispute Resolution Project* 10 LOYOLA U. CHICAGO. INT’L L. REV. 11, 11 (2012); *See also*, Raymond Yeah, *But Did You See the Gorilla?*, *supra* note 325.

<sup>366</sup> *See* ODR Reg., *supra* note 330 at Art 8; EUROPA, Press Release, *supra* note 323.

<sup>367</sup> In each Member State, a competent authority will be in charge of monitoring the functioning of ADR entities established on its territory. *See* ADR Directive, *supra* note 326, at art 27; ODR Reg., *supra* note 330 at Art 6.

<sup>368</sup> Including a high level of transparency about the process and prior outcomes, statistics, etc. *See* ADR Directive, *supra* note 326, at art 7.

complaints and facilitate communication between the parties.<sup>369</sup> Most importantly, the new rules will provide for ADR entities to settle a dispute within ninety days,<sup>370</sup> a significantly quicker process than many existing small claims courts throughout Europe.<sup>371</sup>

Some of the E.U. Member States are exploring the use of technology within the entirety of the justice system. For example, in the United Kingdom, the Ministry of Justice and HM Courts and Tribunals Service “are implementing a wide-ranging court reform and digitalisation programme across the justice system.”<sup>372</sup> This implementation, while in its early stages, is nonetheless expected to include developments in online dispute resolution.<sup>373</sup>

### 3. Supra-National/Institutional Efforts

Despite the growing success of ODR in both public and private domestic markets in many countries, the harmonized cross-border ODR market is surprisingly limited and sparse.<sup>374</sup> In fact, until recently, only a few international institutions had even attempted to craft a cross-border legal instrument.<sup>375</sup> This lag is notable given that the enormous growth and exponential potential of cross-border ecommerce has been a driver behind the development of each of the domestic initiatives so far discussed, including the

---

<sup>369</sup> See ADR Directive, *supra* note 326, at art 11; EUROPA, Press Release, *supra* note 323.

<sup>370</sup> See ADR Directive, *supra* note 326, at art 8; EUROPA, Press Release, *supra* note 323.

<sup>371</sup> For example, the Regulation on European small claims the gives the defendant thirty days to respond to the complaint. See Regulation (EC) No 861/2007 of the European Parliament and of the Council of July 11, 2007, Establishing a European Small Claims Procedure (2007).

<sup>372</sup> Robert Thomas and Dr. Joe Tomlinson, *The Digitalisation Of Tribunals: What We Know And What We Need To Know*, Public Law Project (April 2018).

<sup>373</sup> See *id.*

<sup>374</sup> U.N. Secretary-General, Rep. of Working Group III (Online Dispute Resolution), 22<sup>nd</sup> Sess., Dec. 13-17, 2010, A/CN.9/716 at 31 (2010) [hereinafter “U.N., Working Group II Report Dec 2010”]. That is not to say that several regional and domestic systems do not exist, many of which have served as an influence to the U.N. ODR text. See *id.* at paras 32-34.

<sup>375</sup> See *id.*

European Union ADR/ODR regime.<sup>376</sup> Most of these documents seek to resolve issues associated with commercial transactions,<sup>377</sup> which produce a series of predictable issues, primarily associated with consumers and consumer law within the commercial environment.<sup>378</sup> In many ways, this sole issue is a major barrier to cross-border ODR.<sup>379</sup> Taking up the challenge, several supra-national entities and international institutions have attempted to produce best practices, frameworks and guidance documents, primarily addressing conventional ODR.<sup>380</sup> For example, the Internet Corporation for Assigned Names and Numbers (ICANN) uses an online dispute process which allows individuals to file claims within the area of domain names,<sup>381</sup> while the Organization of American States created an “ODR Initiative for the Electronic Resolution of Cross-Border E-Commerce Consumer Disputes.”<sup>382</sup>

It is the United Nations Commission on International Trade Law, however, that has made the most extensive attempt to harmonize the ODR eco-system across borders. Drawing its mandate from the U.N. General Assembly’s forty-third session,<sup>383</sup> the United Nations Online Dispute Resolution (ODR) Working Group was tasked with creating, over a five-year period, a series of instruments to facilitate development of a platform offering a three-phase dispute resolution process.<sup>384</sup> The instruments originally envisioned were: (1) procedural rules; (2) accreditation standards and minimum requirements for ODR providers/platforms; (3) guidelines and minimum requirements for ODR neutrals; (4) principles for resolving ODR disputes; and (5) a

---

<sup>376</sup> See discussion, text and corresponding footnotes, *supra* notes 323-341.

<sup>377</sup> See U.N., Working Group II Report Dec 2010, *supra* note 343, at 32.

<sup>378</sup> See *id.*

<sup>379</sup> See *id.*

<sup>380</sup> See *id.*

<sup>381</sup> See ICANN, *Online Dispute Resolution Standards of Practice*, ICANN website, (2010).

<sup>382</sup> See Colin Rule, Vikki Rogers, and Louis Del Duca, *Designing a Global Consumer Online Dispute Resolution (ODR) System for Cross-Border Small Value-High Volume Claims—OAS Developments*, UNIFORM COMMERCIAL CODE LAW JOURNAL [Vol. 42 #3](2010). Interestingly, ICANN is a ‘community’ system because enforcement can be done by removing the domain name from a party who fails to comply with the outcome, while the OAS model used a traditional ODR process, culminating in arbitration. See *id.*

<sup>383</sup> See U.N., Working Group II Report Dec 2010, *supra* note 343, at para. 385.

<sup>384</sup> See *id.* at para 115.



cross-border enforcement mechanism.<sup>385</sup> The overall goal was to create a quick, simple, and inexpensive means of resolving disputes involving low-value, high-volume, cross-border, electronic commerce transactions.<sup>386</sup>

In the end, the UNCITRAL Working Group (III) did not achieve this original broad mandate; it failed to produce even one of the documents originally envisioned. It met a later, revised mandate when it produced a set of Technical Notes on Dispute Resolution.<sup>387</sup> At the February 2015 meeting, both the U.S. and the European Union suggested that work be abandoned since no consensus could be reached as to the content of the *procedural* rules.<sup>388</sup> Ultimately, it was consumer law and the newly envisioned E.U. ODR initiative that ended the conversation as the U.S. and the E.U. each recruited allies and resisted compromise.<sup>389</sup> Some commentators are quietly asking if technology will ever be able to overcome the power of local law or the significance of the state line in the psyche of legislators. The quiet discontent is partially responsible for the growth of regional attempts at public online ADR and private entrepreneurial ODR.

The Working Group refused to let their hard work go unrecorded, and as such, produced Technical Notes which provide important insights into the role of the decision maker, which we will be explored in more detail the next section. But, before proceeding, it is necessary to emphasize that cross-border issues — relating to the rule of law, access to justice, and individual protections such as consumer protection and privacy rights — will continue to be sticking points for legislative drafting.<sup>390</sup> UNCITRAL's enormous

---

<sup>385</sup> *See id.*

<sup>386</sup> *See id.*

<sup>387</sup> *See id.*

<sup>388</sup> An occurrence I sat in the room and witnessed. It has also been reported by others, for example Karim Benyekhlef and Nicolas Vermeys, *UNCITRAL Adopts Technical Notes on ODR*, SLAW (April 2016).

<sup>389</sup> It is important to note, I was part of an NGO and at the meetings and part of the group frequently consulted by the US Delegation. I am still very active in the various groups. This is my opinion and should not be imputed to others, the group, or any NGO.

<sup>390</sup> This raises the issue of to what extent the Internet should be considered a “borderless” environment. Two schools of thought play out, one depicting cyberspace as a commons largely free from governmental interference, while the other considers

scaling back of an initially lofty goal is demonstrative of the difficulty of creating a cross-border legal text. For this very reason, in this thesis I have elected to draft best practices as opposed to legal standards. The UNCITRAL Technical Notes and similar documents are a good first step toward harmonization of the ODR justice eco-system, especially when legislative processes fail.

At the Asia–Pacific Economic Cooperation (APEC) Economic Committee Workshop in Japan 2018, the ODR Working Group spent two days focused on *The Collaborative Framework for ODR in APEC Accelerating Justice, Accelerating Trade*.<sup>391</sup> The primary purpose of the Workshop was for the members and invited participants to develop several things: a collaborative framework, ODR procedural rules and standards for a pilot with ODR providers, and Regional Arbitration/Mediation Centers.<sup>392</sup>

The Collaborative Framework was specifically tasked with establishing an APEC–sponsored initiative to use ODR to help global businesses, in particular Micro, Small & Medium Enterprises (M/o MSME), resolve business-to-business (B2B) cross-border disputes, focusing on low-value disputes. ODR Providers would be expected to adhere to the Procedural Rules developed within the APEC ODR Framework, yet APEC itself would not enter the market. Instead, ODR Providers would agree to abide by the Rules but would be free to design and deploy their own platforms. Moreover, APEC was not empowered to monitor nor advise the providers; instead the Strengthening Economic and Legal Infrastructural (SELI)<sup>393</sup> Coordinating Group was tasked with maintaining a

---

cyberspace to be an extension of national territory. *See e.g.*, Dan Hunter, *Cyberspace as Place and the Tragedy of the Digital Anticommons*, 91 CALIFORNIA L. REV. 439, 519 (2003) (depicting cyberspace as a traditional commons and warning that inaction will lead to an intractable digital anti-commons); David R. Johnson & David Post, *Law and Borders—The Rise of Law in Cyberspace*, 48 STANFORD L. REV. 1367, 1367 (1996) (arguing that “[g]lobal computer-based communications cut across territorial borders, creating a new realm of human activity and undermining the feasibility—and legitimacy—of laws based on geographic boundaries”).

<sup>391</sup> Author, *APEC’s New Online Dispute Resolution Framework*, NATIONAL CENTER FOR TECHNOLOGY AND DISPUTE RESOLUTION (March 12, 2019)(discussing APEC embracing the ODR pilot designed in previous meetings).

<sup>392</sup> *See id.*

<sup>393</sup> SELI will also compile data from the various providers to measure the success of the pilot.

list of partners who are compliant with the ODR Cooperative Framework and will maintain a website so others can access this list.

The ODR Framework “creates an electronic system for MSMEs in APEC participating economies providing ODR through negotiation, mediation and arbitration of B2B claims.”<sup>394</sup> Thus, the Framework has a traditional ODR model in mind, yet supports departures from the traditional model. The Framework seeks to set out the steps needed to encourage participation and ensure legal compliance amongst APEC Member States. It is the Procedural Rules, however, which demand attention for the purposes of this thesis.

The Procedural Rules are the first clear indication that APEC may be envisioning a significant change. Article 6 proscribes: “. . . the ODR administrator shall promptly appoint a neutral in accordance with Article 9 ...”<sup>395</sup> Of course, this removes the appointment power from the parties. However, the Rules are ambiguous as to whether the appointed neutral party must necessarily be a human being.<sup>396</sup> The Rules do envision a declaration, upon receiving notice of the appointment, of the impartiality and independence of the neutral<sup>397</sup> and do allow either party to object to the appointment.<sup>398</sup> In instances of objection, the neutral is automatically disqualified.<sup>399</sup> When a party objects to multiple appointments, the party is allowed 3 automatic disqualifications, after which the ODR administrator will determine whether the neutral shall be replaced and will make such determination based on “facts or circumstances that might give rise to justifiable doubts of the neutral’s impartiality and independence.”<sup>400</sup> Although the process may sound complex, it is actually reasonably efficient and seamless, with few time delays. For example, the objecting party has two days to ask for an automatic disqualification.<sup>401</sup> Hence, an ODR platform that complies with the Procedural Rules

---

<sup>394</sup> *See id.*

<sup>395</sup> APEC, Proc. Rules, Article 9.

<sup>396</sup> *See id.*

<sup>397</sup> *See id.*

<sup>398</sup> *See id.*

<sup>399</sup> *See id.*

<sup>400</sup> APEC, Proc. Rules, Article 9; UNCITRAL Tech. Notes, para 48(b).

<sup>401</sup> *See* APEC, Proc. Rules, Article 9.

will have a process of appointment that is reasonably quick and uses technology in an advantageous manner. In addition, the standard for measuring the independence and impartiality of the neutral<sup>402</sup> is already recognized by a majority of institutions and states and hence should not cause enforceability issues. It will however, be interesting to see what happens when the system becomes fully automated, as these Rules, and others, enforce a requirement to “declare impartiality and independence.”

## B. Online Dispute Resolution: Appointment, Impartiality, Independence, and Disclosure

To date, there are very few legislative guides, best practices or other governance documents that address how to build and ensure an online decision maker’s impartiality, independence and freedom from bias. However, as the online justice environment grows, a small body of research and policy in this area is beginning to form. This section will explore first the institutional attempts to provide this guidance and then some of the attempts of ODR providers themselves to ensure an independent and impartial decision maker.

### 1. Institutional Guidance Concerning ODR Decision Maker

We have touched on the recent attempts by the UNCITRAL Working Group to delve into the appropriate mechanisms required to ensure decision maker impartiality and independence in ODR.<sup>403</sup> The Technical Notes<sup>404</sup> this group published were created by some of the most respected authorities in the field after a lengthy process of discussion, and are likely to have greater impact than their name might suggest. As such, this document will guide the development of this section.

---

<sup>402</sup> *See id.*

<sup>403</sup> *See* United Nations Commission On International Trade Law, *Technical Notes on Online Dispute Resolution of the United Nations Commission on International Trade Law*, 71/138 (Resolution adopted by the General Assembly on 13 December, 2016)(2017)[hereinafter Technical Notes].

<sup>404</sup> *See e.g.*, United Nations Commission On International Trade Law, *Report Of Working Group III (Online Dispute Resolution) On The Work Of Its Thirty-Third Session* (New York, 29 February-4 March 2016)(Section II listing participants list).

Let us first highlight two important points that provide necessary context for appropriate interpretation of these Technical Notes. First, because the Technical Notes are covering all of ODR, there was a need to consider all potential forms of technology-driven dispute resolution. While this thesis has focused on arbitration, the Working Group was examining a more expansive set of possibilities throughout the drafting process. For example, the Working Group adopted terminology, such as “neutral,” that is not specific to ADR or any other existing process, without providing an official definition. This maximizes the future adaptability of their language; any other potential term, such as “arbitrator,” would have necessarily referenced a particular species of dispute resolution. The downside of using such a non-specific term is that it may create difficulties, for example when it comes to specific protections included in one process but not another, as previously discussed. The lack of specificity means that the process addressed by the Technical Notes cannot be presumed to include rule of law protections by default, and these issues must be actively examined.

The second point of note is that the Technical Notes are aspirational, meaning that rather than producing strict proscriptions, they provide recommendations and leave the full development of the procedural rules to the platform developer or legislative body. (Chapter 5 will attempt to suggest just such a development.<sup>405</sup>)

Narrowing down the discussion, there are three areas of interest relevant to the thesis: judicial independence, appointment of the neutrals and governance. Each will be explored in turn in the following sections.

*a. Appointment, functions and independence/impartiality of the decision maker*

In a process that looks notably similar to the appointment process within many arbitration institutions, the Technical Notes proscribe:

Technical Notes Section X — Appointment, powers and functions of the neutral

...

(b) The neutral {shall} be required to declare his or her impartiality and independence and disclose at any time

---

<sup>405</sup> See *infra*, Chapter 5.

any facts or circumstances that might give rise to likely doubts as to his or her impartiality or independence;<sup>406</sup>

Note the expansive expectation of disclosure indicated by the phrasing “likely doubts.” These proscriptions are in line with the widely regarded international arbitration expectations as described in Chapter 2<sup>407</sup>; in particular, note that disclosure, in the interest of safeguarding impartiality and independence, is an ongoing obligation.<sup>408</sup> The author therefore posits that the Technical Notes are compliant with RoL expectations regarding the decision maker, as described in Chapter 1.

The Technical Notes also include a more novel expectation: that the neutral will “treat both parties equally.” This requirement has *not* been articulated in prior rules or legislation regarding commercial arbitration.<sup>409</sup> Since the Technical Notes are intended to cover a wider subject area than commercial law, it may be that the Working Group envisioned this expectation as adding to the breadth of their relevance.

The notes continue:

(c) The ODR system provides parties with a method for objecting to the appointment of a neutral;<sup>410</sup>

Presumptively, the objection would need to be based upon the previously expressed recommendation that “likely doubts” exist concerning the neutral’s impartiality and independence.<sup>411</sup> There may also be an implied right of objection when a likely doubt exists relating to treating both parties equally. But this is a mere reading and has not been put into practice and is thus an untested standard. While there is not necessarily an expectation of the ability to object to the participation of a decision maker, the inclusion is considered a further protection of the RoL expectation of the presence of an independent and impartial decision maker. Hence, the provision may be thought of as essential in the process.

---

<sup>406</sup> Technical Notes, *supra* note 374, Section X, at 48(b)

<sup>407</sup> *See supra*, Chapter 2, at Table B.

<sup>408</sup> Technical Notes, *supra* note 374, Section X, at 49(d).

<sup>409</sup> *See supra*, Chapter 2, at Table B.

<sup>410</sup> Technical Notes, *supra* note 374, Section X, at 48(c).

<sup>411</sup> *See supra*, Chapter 2, at Table B.

What entity holds the authority to make determinations relating to independence and impartiality? Like most institutional arbitration rules, the Technical Notes provide that:

(d) In the event of an objection to an appointment of a neutral, the ODR administrator {shall} be required to make a determination as to whether the neutral shall be replaced;<sup>412</sup>

This is consistent with standard practice: one of the hallmarks of arbitration is that the parties agree to make use of an authority<sup>413</sup> — an arbitration institution or ODR platform — to make the majority of procedural decisions.

This general rule is not without exception, however. Taking a somewhat different approach, the ABA Task Force focused more on the Providers as the primary source ensuring impartiality and independence, as opposed to focusing the primary disclosure requirements on the decision makers or the administrators/institutions. For example, “ODR Providers must disclose all matters that might raise a reasonable question about the impartiality of the ODR Provider or its neutral(s).”<sup>414</sup> The ABA Guidelines are also *more* specific about disclosures and conflicts of interest.

Specifically, ODR Providers should disclose the following:

A. Relationship to others concerning providing ODR services.

1. If the ODR Provider provides ODR services under a contractual relationship with other organizations, such as merchants, trade associations, etc.
2. If the ODR Provider provides any referral compensation (referral fees, rebates, commissions, etc.), and if so:
  - a) to whom it is paid, and
  - b) the amount of the compensation or the basis for calculating the amount of the compensation.<sup>415</sup>

---

<sup>412</sup> Technical Notes, *supra* note 374, Section X, at 48(d)

<sup>413</sup> *See supra*, Chapter 2, at Table C.

<sup>414</sup> American Bar Association Task Force On Ecommerce And ADR, *Recommended Best Practices For Online Dispute Resolution Service Providers*, at VI, [hereinafter ABA Task Force, ODR Best Practices] available at <https://www.americanbar.org/content/dam/aba/migrated/dispute/documents/BestPracticesFinal102802.authcheckdam.pdf> (last visited Aug. 5, 2019).

<sup>415</sup> ABA Task Force, *ODR Best Practices*, *supra* note 385, at VI.

The included list is more specific than generally included within a set of best practices designed with a brick-and-mortar arbitration environment in mind;<sup>416</sup> however, the list serves to require the *platform* to provide disclosures, a necessity in an ODR environment. A privately developed technology platform may itself be influenced in a manner that a court system is not. The disclosure of certain information is thus essential to ensure the system itself is viewed as trustworthy and so the platform can ensure RoL compliance, especially relating to the decision process.

The ABA Best Practices suggest that ODR Providers be held responsible for the disclosures of their neutrals. For example:

3. An ODR Provider shall take reasonable steps to cause its Neutrals to disclose whether the Neutral or a person closely affiliated with the Neutral (e.g., spouse, relative, or business partner, etc.) has any conflicts of interest, including but not limited to:
  - a) Any direct personal, business, professional or financial relationship with a party or its representative;
  - b) Any direct or indirect interest in the subject matter or outcome of the dispute, including contingent fee arrangements; and
  - c) Any personal knowledge that the ODR Provider or Neutral has of facts relevant to the dispute.<sup>417</sup>

Placing responsibility on the platform/provider to ensure compliance with RoL disclosure requirements is an important and innovative assurance of a trustworthy system. As briefly reviewed in Chapter 2, arbitration institutions have *not* been viewed as responsible for the actions, or inactions, of their arbitrators. Placing responsibility upon the platform ensures some level of oversight and responsibility over the decision making process. Best practices that make the platform responsible for failures to comply with RoL standards will also help to ensure that any technology deployed in the decision making processes are carefully vetted.

Neutrals themselves *are* also held to disclosure requirements under the ABA Guidelines: “Neutrals shall . . . Disclose all conflicts of interest and act with impartiality

---

<sup>416</sup> See *supra*, Chapter 2, at Table B.

<sup>417</sup> ABA Task Force, *ODR Best Practices*, *supra* note 385, at VI.



and independence . . .<sup>418</sup> However, there are no clear guidelines as to what merits disclosure. Unfortunately, this means that ABA practice follows the traditional arbitration approach to both issues.<sup>419</sup> That approach of course, should suffice to reduce or eliminate issues that may arise within the context of arbitration, award enforcement, and the New York Convention,<sup>420</sup> but it is not fully adequate to regulate the technology that may gain wide use in the near future. While the inclusion of the platform/provider as a responsible party ensuring compliance in disclosures is an important step, this thesis argues that the ABA and other bodies must go further, by insisting that providers ensure the absence of impartiality and independence (and bias)<sup>421</sup> in the technology as well.

*b. Governance*

In terms of process governance, the UNCITRAL Technical Notes do make clear that the use of ODR should not be a mechanism to subvert or otherwise ignore basic expectations already well established within the justice environment.

Technical Notes Section XII — Governance

52. It is desirable for guidelines (and/or minimum requirements) to exist in relation to the conduct of ODR platforms and administrators.

53. It is desirable that ODR proceedings be subject to the same due process standards that apply to that process in an offline context, in particular independence, neutrality and impartiality.<sup>422</sup>

However, the Technical Notes drafters were well aware of the need to strike a balance between respecting RoL and answering the need for a less onerous and costly process than those offered by existing systems.

50. While the process for appointment of a neutral for an ODR proceeding is subject to the same due process standards that apply to that process in an offline context, it may be desirable to use streamlined appointment and challenge procedures in order to address the need for ODR to provide a simple, time-, and cost-effective alternative to traditional approaches to dispute resolution.<sup>423</sup>

---

<sup>418</sup> ABA Task Force, *ODR Best Practices*, *supra* note 385, at VIII A.

<sup>419</sup> *See supra*, Chapter 2, Section B(3).

<sup>420</sup> *See supra*, Chapter 2, Section B(3).

<sup>421</sup> *See infra* Chapter 4, Section C(3) and Chapter 5, Section (A).

<sup>422</sup> Technical Notes, *supra* note 374, Section XII, at 52 and 53.

<sup>423</sup> Technical Notes, *supra* note 374, Section XI, at 50.

In other words, due process is to be respected, but not in the form of a simple copy of the existing brick-and-mortar system. Streamlining is allowable, even desirable, so long as basic due process, and by extension the rule of law, is respected. So what are the most basic features of due process? In our particular instance, the author argues that respect for the rule of law and due process demands that the platforms, institutions and individuals involved in providing ODR ensure an impartial, independent, neutral decision maker.

*c. Transparency*

To everyone from cyber security professionals to privacy experts, transparency is frequently considered an essential element in any cyber-governance.<sup>424</sup> Yet, the exact parameters of the transparency, and the specifics of what and when items should be transparent are a matter of discussion. The Technical Notes drafters were well aware of this, and the Notes elaborate on two types of transparency: disclosure of relationships (decision maker independence) and transparency in data gathering/sharing/outcomes, a somewhat different, but as we shall explore, ultimately related concern.

First, the UNCITRAL Technical Notes Section II, Principles: Transparency recommends:

It is desirable to disclose any relationship between the ODR administrator and a particular vendor, so that users of the service are informed of potential conflicts of interest.<sup>425</sup>

It is important to note that the transparency described in this section applies to the ODR *administrator* and encompasses the administrator's connections to vendors and neutrals. As such, any connection among these entities should be disclosed in order to address conflict of interest concerns. The creation of a recommendation specifically focusing on the technology, through the administrator or vendor, is an important aspect of independence that could have easily been overlooked in the emerging cyber-justice ecosystem. The assurances that a system, the entirety of the system, even those administering or otherwise creating the technology, should disclose prior relationships

---

<sup>424</sup> See e.g., Scott J. Shackelford & Steve Myers, *Block-by-Block: Leveraging the Power of Blockchain Technology to Build Trust and Promote CyberPeace*, 19 YALE J.L. & TECH, 376 (2018)(discussing monitoring occurring through transparency of the public nature of the blockchain.).

<sup>425</sup> Technical Notes, *supra* note 374, Section II Principles: Transparency, at 10

is an important one, as it demonstrates respect of the rule of law and ensures the independence of the entire process, technology included.

The second type of transparency addressed in the ABA Best Practices concerns the selection of neutrals, specifically:

How individuals are selected to become part of the panel of neutrals eligible to handle disputes.

How a particular neutral is selected to handle a particular dispute.

The process by which neutrals are required to certify that they have no conflicts of interest and have disclosed all matters that reasonably might affect impartiality with respect to a particular dispute.

Procedures for disqualification of a neutral for cause.<sup>426</sup>

And, the ABA document includes a provision not yet included in any other guidance:

Identify and/or link to the ethical rules by which the neutrals are bound.<sup>427</sup>

This provision relates to the fact that certain local rules may direct the decision maker's disclosures; transparency about which rules they are following allows parties who feel their right may have been violated to know the appropriate reporting authority. Other parts of the best practices literature address outcome transparency. For example, the Technical Notes say:

The ODR administrator may wish to publish anonymized data or statistics on outcomes in ODR processes, in order to enable parties to assess its overall record, consistent with applicable principles of confidentiality.<sup>428</sup>

All relevant information should be available on the ODR administrator's website in a user-friendly and accessible manner.<sup>429</sup>

---

<sup>426</sup> ABA Task Force, *ODR Best Practices*, *supra* note at VI B.

<sup>427</sup> ABA Task Force, *ODR Best Practices*, *supra* note 385, at VI C.

<sup>428</sup> Technical Notes, *supra* note 374, Section II Principles: Transparency, at 11. The ABA recommends a similar level of disclosure. *See* ABA Task Force, *ODR Best Practices*, *supra* note 385, at II A-C.

<sup>429</sup> Technical Notes, *supra* note 374, Section II Principles: Transparency, at 12

Those invested in the use of technology in the justice eco-system are aware of the necessity of ensuring outcome transparency, which allows for outcomes to be audited for issues such as dependency and partiality. This ensures that those viewing the process can be satisfied that the process is fundamentally fair.

The ABA has also recommended transparency in terms of outcomes

If the ODR Provider handles a substantial number of B2C disputes, regular periodic statistical reports should be published online that permit a meaningful evaluation of the proceedings and that respect the confidentiality of the participants and the individual proceedings.<sup>430</sup>

However, where the Technical Notes provide broad language to cover a range of contexts, the ABA Best Practices zoom in to a more granular level, in particular distinguishing between business-to-consumer and other types of disputes that may be resolved within ODR.

With respect to other disputes, particularly those resolved by arbitration, participants should be encouraged to allow the decisions to be published with any confidential or propriety information deleted.<sup>431</sup>

Outcome transparency will be an integral part of any successful expansion of the use of ODR in justice environments. Transparency supports compliance with fundamental principles of justice, such as adherence to the RoL, and it allows individuals to gather information and make informed decisions about their choice of providers, with the ability to assess the providers' adherence to RoL expectations. Moreover, as more portions of the system are captured behind technology-driven providers, transparency ensures outside entities can audit and monitor outcomes for adherence to the RoL as well as the absence of partiality, dependence, and bias from decision makers or the larger system itself.

## 2. ODR Providers Identified Manner of Appointment

With the technology already in implementation, the appointment process has become a matter of active concern to ODR developers. Several providers have developed and

---

<sup>430</sup> ABA Task Force, *ODR Best Practices*, *supra* note 385, at II D.

<sup>431</sup> *Id.* at II E.

deployed appointment processes that include the appointment of a human decision maker within the ODR platform. The following is a composite list of some of the appointment processes of the major institutions and providers.

*Table E: ODR Providers Selection of the Decision Maker*

<b>NAME</b>	<b>DESCRIPTION</b>	<b>ARBITRATION RULES</b>	<b>SELECTING DECISION MAKER</b>
<b>Arbitranet</b>	Brazil Online Arbitration Provider that promises a decision in 100 days. The online system chooses an arbitrator based on the indications of the parties		Any person of legal age who is in perfect mental condition and who is trustworthy can be an arbitrator. Under local arbitration law, the referee is not required to be lawyers or meet other technical requirements, but it is supposed to be a specialist in the field.
<b>Camera Arbitrale di Milano - Risolvionline.com</b>	Italian online mediation primarily for retail purposes. Consumers can file a complaint online. The vendor is invited to participate in the negotiation. An expert provides a non-binding evaluation.	RULES OF RISOLVIONLINE.COM	The Secretariat identifies the third party from a list drawn up according to criteria of competence and professionalism. The third party undertakes to comply with the Rules of RisolviOnline.com and the Rules of Conduct attached to them.
<b>FairClaims</b>	Platform assigns an experienced Arbitrator to	Ethics Standards for Neutral Arbitrators in	Arbitrators are assigned by the platform

NAME	DESCRIPTION	ARBITRATION RULES	SELECTING DECISION MAKER
	<p>render a legally binding, court-enforceable decision. The suite of tools includes a smart settlement engine (FairOffer) with mediated chat (FairChat), AI mediation (FairBot), a smart intake and assessment tool (FairTake), video arbitration and a claims dashboard.</p>	<p>Contractual Arbitration adopted by the Judicial Council of California.</p> <p>FairClaims Rules</p>	
<p><b>VirtualCourthouse</b></p>	<p>Virtual Courthouse is a web-based combination of multimedia technologies and business processes, integrated with a customer-friendly user interface.</p>	<p>VirtualCourthouse follows the pattern established by the Uniform Arbitration Act and the Federal Arbitration Act.</p>	<p>The initiating party selects 3 or more mediators/arbitrators, rank them and send the list to the opposing party along with an invitation to participate in an alternative dispute resolution process.</p>
<p><b>World Intellectual Property Organization</b></p>	<p>WIPO ECAF, which is accessible through the Center’s web site, allows for secure filing, storing and retrieval of case-related submissions in an electronic docket, by parties, neutral(s) and the Center, from anywhere in the world. WIPO ECAF also facilitates case management for</p>		<p>Parties may appoint their own arbitrators, mediators and experts, or may leave the choice to the Center. For this purpose, the Center maintains lists of mediators, arbitrators and experts, including comprehensive details of their qualifications and experience. The persons listed represent a wide variety of</p>

NAME	DESCRIPTION	ARBITRATION RULES	SELECTING DECISION MAKER
	all actors in a WIPO proceeding by providing, in addition to the online docket, a case overview, time tracking and finance information.		specializations in all fields of intellectual property and in dispute resolution. The lists, which are constantly expanding, presently comprise over 1,000 persons from 70 countries.

Author compilation of online website information (Up-to-date April 2019)

The chart describes the common process for appointing a decision maker. It is important to note several important aspects of the appointment process. First, party autonomy is the default rule and is protected in all examples provided. Chapter 2 spends considerable time on the arbitration community's commitment to party autonomy, which includes the parties' ability to appoint a decision maker.<sup>432</sup> Moreover, it is important to recall, in the context of arbitration, party autonomy is protected from contract creation because the parties select the institution and/or provider and as such are agreeing to that institution/provider's process of selecting the decision maker. Second, if the parties have not selected a decision maker, the institution/platform rules set the standard. In general, there are two options (1) institutional selection, on behalf of the parties, and (2) agreement by the parties once the dispute is initiated, by selecting from a list of approved decision makers.<sup>433</sup> Either way, two key takeaways apply: (1) the institution/platform controls the process, and (2) the institution/platform controls the individuals who are allowed to resolve disputes, through the use of qualification checks, the pool of decision makers, and the monitoring of behavior. The institution/platforms are designed to be active participants in the process, including the approval of the decision maker selection of the parties.

<sup>432</sup> See *supra* Chapter 2, Section B(2).

<sup>433</sup> See *id.*

### 3. Providers’ Rules of Independence and Impartiality

Of course, the selection process is merely the first step in the appointment of the decision maker. Next the institutions/platforms enforce whatever rules they have established to ensure the independence/impartiality of the decision maker. The following is a composite list of these rules from major providers.

*Table F: Current ODR Institutional/Platform Attempts to ensure Independence and Impartiality*

<b>ODR Institutional Rules</b>	<b>Mechanism for appointing the decision maker</b>	<b>The rules relating to the impartiality of the decision maker</b>	<b>The rules relating to the independence of the decision maker</b>
<b>U.N. Working Group</b>	A neutral is required for the dispute resolution process — the ODR administrator “promptly” appoints the neutral — the ODR administrator promptly notifies the parties of the name of the neutral and any other relevant or identifying information.	<p>The neutral is required to declare his or her impartiality and independence and disclose at any time any facts or circumstances that might give rise to likely doubts as to his or her impartiality or independence.</p> <p>The neutral is required to remain independent, impartial and treat both parties equally throughout the proceedings.</p>	Not distinguished from rules for ensuring impartiality.



ODR Institutional Rules	Mechanism for appointing the decision maker	The rules relating to the impartiality of the decision maker	The rules relating to the independence of the decision maker
<b>European Commission ODR</b>	The company will provide a list of ADR providers who can assist in resolving the complaint. (a) Initiating party selects a provider using the platform. (b) Initiating party may propose their own alternative provider if they do not wish to use any of the listed options. Both parties must agree on an ADR provider within 30 days of the initial submission to the platform, or the case is closed.	N/A	N/A
<b>Arbitration Resolution Services</b>	An arbitrator (a highly-educated and trained “private judge”) is <b>randomly selected</b> from a pool composed based on expertise and relevant knowledge.	Arbitrators undergo a <b>conflict check</b> to ensure complete neutrality. The evidence submission process meets all industry encryption standards in order to provide a completely neutral and secure arbitration.	Not distinguished from rules for ensuring impartiality.

<b>ODR Institutional Rules</b>	<b>Mechanism for appointing the decision maker</b>	<b>The rules relating to the impartiality of the decision maker</b>	<b>The rules relating to the independence of the decision maker</b>
<b>Virtual Courthouse</b>	<p>Initiate a case.            Choose your preferred Neutrals.            Invite the opposing party (as many as you want). You can also invite co-counsel to the case. You both agree to a Neutral.            We have hundreds of Neutrals with lengthy experience....</p>	<p><b>Neutral Oath</b> (that neutrals have to sign):</p> <ol style="list-style-type: none"> <li>1. Provide truthful and accurate information about my qualification, experience and credentials;</li> <li>2. Fairly, impartially and justly conduct my activities as a neutral based only upon the merits of each case;</li> <li>3. Disclose to the parties any information which may reasonably call into question my impartiality;</li> <li>4. Conduct my activities as a neutral in a prompt fashion;</li> <li>5. Comply with all applicable ethical rules, regulations and standards as from time to time adopted;</li> <li>6. Diligently apply all rules of procedure and law as agreed upon by the parties or adopted by any statute, regulation or rule of law;</li> <li>7. Keep all information obtained from the parties confidential and not disclose that information to any other individual without the express consent and authorization of all of the parties;</li> </ol>	<p>Not distinguished from rules for ensuring impartiality.</p>

<b>ODR Institutional Rules</b>	<b>Mechanism for appointing the decision maker</b>	<b>The rules relating to the impartiality of the decision maker</b>	<b>The rules relating to the independence of the decision maker</b>
<p>Internet Corporation for Assigned Names and Numbers (ICANN) ODR Standards of Practice</p>	<p>N/A</p>	<p>Information about the credentials and experience of third parties shall be made available for the parties;</p> <p>Third parties shall be independent, without any type of personal or professional relationship with any of the parties;</p> <p>The third party, whenever this is feasible, shall not be linked to successful case settlement or to a recommendation or decision in favour of any particular party. When dealing with repeat players ODR service providers shall provide with a mechanism to grant a greater level of transparency and impartiality.</p>	<p>Not distinguished from rules for ensuring impartiality.</p>

Author compilation of online website information (Up-to-date April 2019)

The chart reveals a few key takeaways for those designing policy and implementation rules for ODR platforms. First, while this author has argued that independence and impartiality are independent expectations under the RoL, in current practice the two terms have come to be measured as a single phrase and are in essence treated as the same standard. For example, return to the chart and look at the language used by Virtual

Courthouse to describe the responsibilities of the neutral: to “fairly, impartially and justly conduct my activities . . .”<sup>434</sup> and to “disclose . . . any information which may reasonably call into question my impartiality.”<sup>435</sup> It is important to not assume this standard eliminates all issues that could arise from dependence. Unfortunately, dependence itself is not explicitly addressed.

Nonetheless, there are other systemic features that can address issues with dependence and partiality. The Arbitration Resolution Services requires a conflict check (presumably an automated process, although its standards are yet to be made transparent).<sup>436</sup> It is worth noting that this system uses a random assignment, which certainly greatly lessens the chances of dependence and partiality — especially with the additional safeguard of the conflict check.<sup>437</sup> Thus, while this author would argue for dependence to be more explicitly addressed, we should not overlook design features, such as random assignment, that are also relevant.

A review of the chart reveals the ODR community’s commitment to RoL, as we can see the various providers explicitly addressing RoL issues in their process and rules, generally following well-established legal requirements from the international arbitration world. As discussed in Chapter 1, adherence to the RoL should be expected from any justice provider, including an alternative justice provider.<sup>438</sup> As described in Chapter 2, arbitration has a long history of ensuring independence and impartiality. And while the exact parameters of such protections differ slightly amongst the various institutions, the expectation of the independence and impartiality of the decision makers is unwavering.<sup>439</sup>

---

<sup>434</sup> See *supra* Table F, at Virtual Courthouse, VirtualCourthouse Rules of Arbitration and Mediation.

<sup>435</sup> See *id.*

<sup>436</sup> See *supra* Table F, at Arbitration Resolution Services, information available at <https://www.arbresolutions.com/arbitration-services-arbitration-resolution-services/> (last visited Aug. 10, 2019).

<sup>437</sup> See *supra* Table F, at Arbitration Resolution Services.

<sup>438</sup> See *supra* Chapter 1, Sections C and D. See also, *infra* Chapter 5, Section A (1).

<sup>439</sup> See *supra* Chapter 2, Section A and Table C.

### C. Smart Contracts Develop a New Standard

While the inclusion of smart contracts and blockchain driven environments may seem premature, in fact, virtual judges are in use and they do rely upon and resolve issue in a smart contract/blockchain environments. For example, it has been publicized that China is using a virtual judge- meaning a decision maker that is not human, and is doing the decision making based within a blockchain environment.<sup>440</sup> Moreover, and more directly on point, groups such as Blockchain in Transport Alliance,<sup>441</sup> which includes entities such as FedEx, Delta, Daimler, UPS, are actively considering dispute resolution in their evolving transport focused blockchain environment.<sup>442</sup>

And international institutions are joining the growing conversation and are beginning to ask the difficult questions surrounding when a human must be part of the decision process. Institutions such as the World Economic Forum, and its soon to be (recently released) report entitled: *Dispute Resolution for Blockchain-Based Transactions*<sup>443</sup> are designed recognizing the importance of dispute resolution to build a trusted mechanism. In these instances, and others, arbitration in a virtual environment is being considered, in the exact manner as described shortly and in the final stages of the thesis.

The inclusion of the smart contract environment is essential, even as the conversation in the area is in its early stages of development as these entities are the ones fully considering arbitration, online dispute resolution, dispute resolution as a mechanism of trust, and the fundamental question of when a human must be part of the decision making process.

---

<sup>440</sup> Bryan Lynn, *Robot Justice: The Rise of China's 'Internet Courts'* VOA LEARNING ENGLISH, (Dec. 11, 2019) available at <https://learningenglish.voanews.com/a/robot-justice-the-rise-of-china-s-internet-courts-/5201677.html>

<sup>441</sup> BLOCKCHAIN IN TRANSPORT ALLIANCE: DRIVING STANDARDS AND ENABLING TECHNOLOGY ADOPTION, website information available at <https://www.bitastudio/>

<sup>442</sup> *Id.*

<sup>443</sup> *See id.*

The term “smart contract” was conceived to describe agreements between two or more parties that can be automatically enforced, without an intermediary.<sup>444</sup> At its most basic, a smart contract is a *computer protocol* intended to digitally facilitate, verify, or enforce the negotiation or performance of a contract.

Smart contracts are rendered as computer programs. Users can request the execution of contracts by sending suitable transactions to the nodes of a peer-to-peer network. These nodes collectively maintain the history of all transactions in a public, append-only data structure, called blockchain. The sequence of transactions on the blockchain determines the state of each contract, and, accordingly, the assets of each user.<sup>445</sup>

Smart contracts, thus, are not merely contracts, *per se*, but computer protocols that verify a specific aspect of a party’s performance within a proscribed set of parameters.

Smart contracts, at least the theory behind the technology, have been around for well over 20 years.<sup>446</sup> Recently a resurgence occurred as distributed ledger technologies, such as Ethereum and Hyperledger, became part of a cryptocurrency boom.<sup>447</sup> There are many issues with smart contracts and distributed ledgers that lie well beyond the scope of this thesis, but for our purpose it is important to understand that even well-crafted protocols will not eliminate disputes.<sup>448</sup> Indeed smart contracts are designed with dispute resolution mechanisms in mind. For example, consider the smart contract system being built by the Accord Project.<sup>449</sup> The Accord Project is attempting to build

---

<sup>444</sup> Nick Szabo, *Formalizing And Securing Relationships On Public Networks*, First Monday 2(9)(University of Illinois at Chicago) (1997) (explaining the technology behind smart contracts).

<sup>445</sup> *See id.* at Introduction.

<sup>446</sup> *See id.*

<sup>447</sup> *See* Nicola Atzei, Massimo Bartoletti, Tiziana Cimoli, Stefano Lande, Roberto Zunino, *SoK: Unraveling Bitcoin Smart Contracts*, 7th International Conference on Principles of Security and Trust (POST), European Joint Conferences on Theory and Practice of Software (2018)(discussing smart contracts, ethereum, and consensus).

<sup>448</sup> *See e.g.*, Amy J. Schmitz and Colin Rule, *Online Dispute Resolution for Smart Contracts*, 2019 Journal of Dispute Resolution 103 (2019)(discussing possible disputes that could arise and some solutions).

<sup>449</sup> “The Accord Project’s open source technology enables the creation of legally binding agreements that can “come alive” through connections to external sources of data and software systems that enable digitized and automated contract management, performance, and related operations—the full automation of business transactions.” More information can be found at the Accord Project Blog (July 28, 2018).

a “smart legal contract”<sup>450</sup> and the Accord Project’s Dispute Resolution Working Group, of which the author is currently one of the co-chairs, is examining the dispute resolution process.<sup>451</sup> The work of the Accord Project will serve as a model for explanation and discussion.

At the current time (April 2019) the Accord ODR Working Group is beta testing based upon the following potential contract language:

Late Delivery and Penalty

In case of delayed delivery of Goods, “Betty Buyer” shall pay to “Steve Seller” a penalty amounting to 10.5% of the total value of the Goods for every 2 days of delay. The total amount of penalty shall not, however, exceed 55% of the total value of the delayed goods. If the delay is more than 15 days, the Buyer is entitled to terminate this Contract.<sup>452</sup>

For a smart contract to work within an automated environment, this contract language must be converted into protocols for the technology to execute. The above clause can be converted into protocol as follows:

Late Delivery and Penalty.

In case of delayed delivery of Goods, [{buyer}] shall pay to [{seller}] a penalty amounting to [{penaltyPercentage}]% of the total value of the Goods for every [{penaltyDuration}] of delay. The total amount of penalty shall not, however, exceed [{capPercentage}]% of the total value of the delayed goods. If the delay is more than [{maximumDelay}], the Buyer is entitled to terminate this Contract.<sup>453</sup>

In a conventional legal environment, this contract would be created by human parties. An individual would need to fill out forms, secure shipment space and begin the

---

<sup>450</sup> Accord Project, website, information page (about) available at <https://www.accordproject.org/about/> (last visited Aug. 10, 2019).

<sup>451</sup> Accord Project Dispute Resolution Working Group website, available at <https://www.accordproject.org/working-groups/dispute-resolution> (last visited Aug. 10, 2019).

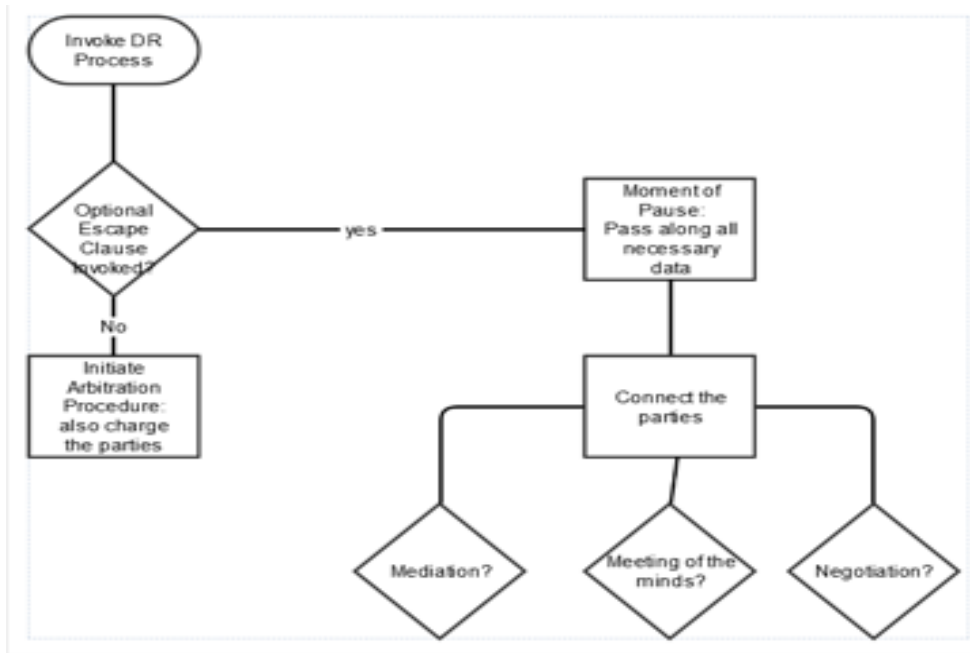
<sup>452</sup> Accord Project, Training Slides and Beta Model Explanation (2018). It should be noted, the author is one of the leaders of the Working group. More information can be found at the Accord Project website at <https://docs.accordproject.org/> (last visited Aug. 10, 2019).

<sup>453</sup> See *id.* Designed using Ergo - a domain-specific language that captures the execution logic of legal contracts. *Id.*

payment process through the banking system. Documents would need to be exchanged amongst various parties and institutions to allow shipment and payment to progress.<sup>454</sup>

In a smart contract all of this process is automated, meaning little human intervention occurs — unless or until something outside of the protocol’s scope calls for it. In our example, the protocol language both sets penalties and, in the case of late delivery, enforces that clause. The non-breaching party will not need to take any action; the system will “charge” the breaching party the penalty amount through an “automated” payment.

Some issues, however, will trigger human involvement. The following schematic shows the envisioned Accord ODR process:



Accord Project, Working Group (Mar.8, 2019)

As can be seen from the schematic, once a payment is late (late as defined in the contract) the harmed party receives a notice and then will be allowed to advance the

---

<sup>454</sup> See Ewan McKendrick, *GOODE ON COMMERCIAL LAW*, Chapter 35, 1055 (Penguin UK; UK ed. (March 1, 2017)(discussing documentary international sales).



issue and formally initiate the dispute resolution process, or to reject the process. The vast majority of developers design automations that follow this (overly simplified) outline, but their approaches to resolving disputes vary, as will be explored next.

### 1. Smart Contracts and ODR

To date, the majority of work in smart contract development has been focused on developing a real-world smart contract. A few potential market participants are seeing the value of exploring dispute resolution as a part of the smart contract process. Table G summarizes dispute resolution providers who are considering the use of ODR within the smart contract environment and who have used the traditional ODR conceptualizations of dispute resolution within their design model.

*Table G: Smart Contract Providers Considering ODR*

NAME	DESCRIPTION	KEY FEATURES
<a href="#"><u>Code Legit</u></a>	Use ready-made smart contracts from the Codelegit legal-by-design smart contracts library. Fitted with the Codelegit Arbitration Certificate, your design architecture runs on the safe side of law.	<ul style="list-style-type: none"> <li>● <a href="#"><u>Arbitration Library Datarella</u></a> “parent company” broad blockchain solution provider</li> </ul>
<a href="#"><u>Confideal</u></a>	Focused primarily on individuals and companies engaged in cryptocurrency markets. However, the platform is open to any existing business that can	<ul style="list-style-type: none"> <li>● Choose arbitrator before entering deal</li> <li>● Arbitrators are professionals from law firms, arbitration institutions and the like. Have to approved by Confideal before being registered on the platform.</li> </ul>

NAME	DESCRIPTION	KEY FEATURES
	integrate blockchain solutions into their business.	
<a href="#"><u>Cryptonomica</u></a>	Provides online dispute resolution via London-based arbitration with awards (decisions) legally enforceable under international law in almost all countries. Contracts must be signed with verified keys to be eligible for the service.	<ul style="list-style-type: none"> <li>● Rule of Arbitration developed from text of The United Nations Commission on International Trade Law (<a href="#"><u>UNCITRAL</u></a>) <a href="#"><u>Arbitration Rules</u></a></li> </ul> <p>Using verified user's public key:</p> <ol style="list-style-type: none"> <li>1. The decision maker can verify his/her signature on a document in the arbitration tribunal under our <a href="#"><u>Arbitration Rules</u></a>, and thus it is <i>legally</i> recognizable and enforceable in all counties that are parties of <a href="#"><u>The New York Convention 1958</u></a></li> <li>2. The complainant can file a case through the online portal.</li> </ol>
<a href="#"><u>Jams-EndDispute</u></a>	Endispute, JAMS Online Mediation, is a video-based mediation service.	<ul style="list-style-type: none"> <li>● Intending to launch <a href="#"><u>Smart Contracts, Blockchain, and Cryptocurrency Practice</u></a></li> <li>● Powered by CourtCall</li> </ul>
<a href="#"><u>LegalZoom</u></a>	LegalZoom <a href="#"><u>uses an API and platform developed by the startup "Clause" together with the Accord Project specification for representation of smart legal contracts to make</u></a> documents available for editing, signing, and execution.	<ul style="list-style-type: none"> <li>● Working with the startup <a href="#"><u>Clause</u></a></li> </ul>
<a href="#"><u>Sagewise</u></a>	Builds technology to efficiently	<ul style="list-style-type: none"> <li>● Start Up, new to the market</li> <li>● The designer of the eBay system — Colin Rule — is involved.</li> </ul>

NAME	DESCRIPTION	KEY FEATURES
	resolve disputes using blockchains and cryptocurrency. System includes smart contracts and a dispute vendor marketplace.	
<a href="#">SmartArb</a>	In cooperation with well-known ADR institutions like the Chartered Institute of Arbitrators, provides a roster of neutrals or the option for parties to propose their own. They are in the process of preparing procedural rules for mediation and arbitration proceedings administered by SmartArb.	

Author compilation of online website information (Up-to-date August 2019)

The chart reveals the growing universe of smart contract providers who are not only exploring the concept academically but designing real-world implementations, including a few who envision a dispute resolution process. For example, Smart Arb reports that they are preparing model clauses to be used in agreement within a smart contract environment.<sup>455</sup> Moreover, several of the smart contract developers are using existing arbitration institutions and rules. Most envision an ODR process with a human arbitrator/decision maker. For example, SageWise, which includes well-known eBay ADR creator Colin Rule among its Advisors,<sup>456</sup> envisions a full ODR process within

<sup>455</sup> See *supra* Chapter 3, Table G at Smart Arb.

<sup>456</sup> See *supra* Chapter 3, Table G at SageWise.

the smart contract/blockchain universe.<sup>457</sup> Commitments to the use of existing institutions and rules will ensure smart contract providers are able to resolve disputes within an online, automated environment will remain committed to RoL and the standards already existing within international commercial arbitration.

## 2. Smart Contract Developers Technology-Driven Solutions

Some ODR providers envision a significantly different technology and dispute resolution process, especially as it relates to the decision maker within a cryptocurrency driven system. As such, these providers merit review in their own separate section. The following chart captures some current ODR developer provisions that relate directly to the role of the decision maker.

---

<sup>457</sup> The term “smart contract” is used to “describe agreements between two or more parties, that can be automatically enforced without a trusted intermediary.” See Atzei, *SoK: Unraveling Bitcoin Smart Contracts*, *supra* note 414 at Introduction.

Two of the more common ones are built in distributed ledger technologies, led by Ethereum and Hyperledger.

In such incarnations, smart contracts are rendered as computer programs. Users can request the execution of contracts by sending suitable transactions to the nodes of a peer-to-peer network. These nodes collectively maintain the history of all transactions in a public, append-only data structure, called blockchain. . . .

Bitcoin is a decentralised infrastructure to securely transfer currency (the bitcoins, B) between users. Transfers of bitcoins are represented as transactions, and the history of all transactions is stored in a public, append-only, distributed data structure called blockchain. Each user can create an arbitrary number of pseudonyms through which sending and receiving bitcoins. The balance of a user is not explicitly stored within the blockchain, but it is determined by the amount of unspent bitcoins directed to the pseudonyms under her control, through one or more transactions. The logic used for linking inputs to outputs is specified by programmable functions, called scripts. . . Bitcoin . . . transactions are grouped into blocks, and . . . each transaction must pay a fee to the “miner” who appends it to the blockchain.

See Atzei, *SoK: Unraveling Bitcoin Smart Contracts*, *supra* note 414 at Introduction.

Table H: Smart Contract Developers Technology-Driven Attempts to Manage Independence and Impartiality of the Decision Maker

<b>BAF — Blockchain Arbitration Forum</b>
<p>BAF offers a “Smart Contract Arbitration Library” which consists of a set of coded rules (based on United Nations Commission on International Trade Law (UNCITRAL) arbitration rules) that allow counterparties to pause, resume, modify and terminate a smart contract. Additionally, the Arbitration Library connects the software with human beings acting as arbitrators. The breach of a contract provides an example of how the automated mechanism works: “one party which considers the other party to be in a breach of the legal contract pauses the execution of the Smart Contract by triggering a (function) pause and send to arbitrator in the Arbitration Library. This function automatically notifies a so-called Appointing Authority defined in the Blockchain Arbitration Rules.”</p>
<b>Brisq</b>
<p>“In order to trade, the user must select at least one arbitrator. He can only select among arbitrators with whom he shares a language. By default the checkbox for auto-selection of matching arbitrators is enabled. That way the user will always have the maximum set of arbitrators matching his language. As there is no concept of arbitrator reputation, it makes no sense to deselect arbitrators, as this will only reduce his trading possibilities. Traders will only be able to take offers of users with whom they have at least one overlapping selected arbitrator. Naturally, an arbitrator cannot be selected for his own trades.”</p>
<b>Confideal</b>

Arbitrators create their accounts on the platform and enter their personal data: languages, location, specialization, working experience, etc. An arbitrator for any particular deal is selected by both sides of the contract during its setup, before the actual signing. At the moment, it is only possible to select a single arbitrator per deal.

In order to ensure a high level of professionalism and satisfaction with arbitration services, there is a rating system for arbitrators on the Confideal platform. It allows the best arbitrators to rise to the top of the ratings of available arbitrators and, consequently, receive more requests for dispute resolution from users. The position of each arbitrator in the ratings is determined by two types of factors: objective and subjective.

Objective factors: speed of response, number of settled disputes, completeness of disclosed information about self. The weight of objective factors accounts for 70% of the arbitrator's rating.

Subjective factors: votes given to arbitrators by the CDL token owners. After the process of arbitration is over, the counterparties involved in the respective contract can evaluate the work of their arbitrator by liking or disliking. Subjective factors account for 30% of the arbitrator's rating.

### **Cryptonomica**

1. If the parties have agreed that a sole arbitrator is to be appointed, and if within 10 days after receipt by all other parties of a proposal for the appointment of a sole arbitrator the parties have not reached agreement thereon, a sole arbitrator shall, at the request of a party, be appointed by the appointing authority.

2. The appointing authority shall appoint the sole arbitrator as promptly as possible. In making the appointment, the appointing authority shall use the following list-procedure, unless the parties agree that the list-procedure should not be used or unless the appointing authority determines in its discretion that the use of the list-procedure is not appropriate for the case:

(a) The appointing authority shall communicate to each of the parties an identical list containing at least three names;

(b) Within 7 days after the receipt of this list, each party may return the list to the appointing authority after having deleted the name or names to which it objects and numbered the remaining names on the list in the order of its preference;

(c) After the expiration of the above period of time the appointing authority shall appoint the sole arbitrator from among the names approved on the lists returned to it and in accordance with the order of preference indicated by the parties;

(d) If for any reason the appointment cannot be made according to this procedure, the appointing authority may exercise its discretion in appointing the sole arbitrator.

### **Juris**

When a dispute is escalated to a SNAP (Juris Suite Billing and Accounting) the Juris protocol will facilitate a polling and discussion process among all available Jurists. They will have a limited amount of time to review the details of the case and provide their judgment on an equitable solution via anonymous vote.

If the parties are unable to resolve their conflict to the satisfaction of all, one or more of the parties may choose to escalate further to a Juris Peremptory Agreement for Neutral Expert Litigation (or PANEL) judgment. When a case is escalated, the Juris system will facilitate the selection of a panel of three High Jurists: those at the highest level of reputation, classified by domain experience and any other factors outlined initially in contract creation. This panel will be selected by U.N. mandated rules, and convene virtually through the Juris Platform.

Jurists are classified by their rank in The Juris Reputation System, and by their areas of demonstrated expertise, all of which are attached to their JurisID. The reputation system is a blockchain-based means of storing and representing the following: proof of certification, merit, competence, and engagement. This pool will initially be seeded with existing, certified arbitrators and legal professionals. But anyone is welcome to sign up to start earning (or losing) reputation by voting on cases, taking part in discussion, and helping with opinions. Classified by expertise if specified, all Jurists are notified when there is a new case. If they take part in an initial SNAP vote they will be required to take part in the final vote, and are encouraged to take part in discussion. Reputation bleeds over time, and may be docked for bad behavior, and anyone is subject to rank demotion. All Jurist signups require proof of identity, and proof of certification if applicable. There are three reputation ranks: High Jurists are pre-certified and Juris Foundation vetted. They are arbitration professionals, and those who have earned advancement and maintained standing through the system. High Jurists are eligible for the more lucrative but time-consuming PANEL judgments. Good Standing Jurists have continued to contribute fruitfully to decisions and discussions on the Juris platform. They are able to vote in SNAP judgments, and the outcome of their anonymous votes will be included in

the case files and reports to involved parties. They are able to advance to High Jurist by gaining reputation. Good Standing level reputation may be initially obtained by providing proof of a law degree or equivalent certification. Novice Jurists are new signups not able to demonstrate any prior experience. They are allowed to take part in SNAP judgments and discussions, contribute to opinions, and earn reputation to advance to Good Standing.

### **Jury Online**

Jury.Online Platform implements a smart contract protocol that includes specifications for deal execution and resolution, on the basis of the Ethereum blockchain, to connect counterparties and judges. (1) System chooses  $n$  random judges (2) Decision by absolute majority, (3) Identity of judges is unknown to the parties, but their competencies are presented. The Jury Online protocol defines the interactions between counterparties, judges and pools of judges, as well as the requirements for information encryption and side-channel communication.

### **Kleros**



Users have an economic interest in serving as jurors in Kleros: collecting the arbitration fee that every juror gets for his work. Candidates will self-select to serve as jurors using a token called pinakion (PNK)<sup>3</sup>.

The probability of being drawn as a juror for a specific dispute is proportional to the amount of tokens a juror deposits. The higher the amount of tokens deposited, the higher the probability the user will be drawn as juror. Jurors who do not deposit pinakions do not have the chance of being drawn. This prevents inactive jurors from being selected.

After candidates have self-selected specific courts and deposited their tokens, the initial selection of jurors is done randomly. Theoretically, a candidate may be drawn more than once for a specific dispute (but in practice it is unlikely). The amount of times a user is drawn for a dispute (called its weight) determines the number of votes he will get in the dispute and the amount of tokens he will win or lose during the token redistribution.

Author compilation of online website information (Up-to-date August 2019)

It is important to notice within the chart several very different approaches to resolving disputes, from traditional to more novel. For example, Brisq requires the selection of an arbitrator in the original contract; the arbitrator must be on the approved list, but beyond that there are few restrictions on the selection.<sup>458</sup> The Brisq system is similar to many or most of the systems so far reviewed in this thesis. Interestingly, Cryptonomica also uses a very traditional appointment process, in which the parties select an arbitrator.<sup>459</sup> In the absence of agreement, the institution produces a list of acceptable decision makers and the parties rank options from the provided list.<sup>460</sup> The institution then appoints based on the parties' ranking. Both of these providers are using a smart contract environment, yet have adopted decision makers and an appointment process that imitate the existing international arbitration environment discussed in Chapter 2.<sup>461</sup>

The remaining developers have developed dispute resolution systems that sit outside any particular smart contract/blockchain environment. These providers, it could be argued, are creating smart contract dispute resolution *institutions*, which function much like the International Chamber of Commerce of the LCIA discussed in Chapter 2

---

<sup>458</sup> See *supra* Chapter 3, Table H at Brisq.

<sup>459</sup> See *supra* Chapter 3, Table H at Cryptonomica.

<sup>460</sup> See *id.*

<sup>461</sup> See *supra* Chapter 2, Section A and Table B.

functions for traditional arbitration.<sup>462</sup> Setting up dispute resolution providers that are not attached to a particular smart contract/blockchain environment frees the designers to think beyond the constraints of currently available arbitration environments. These can be broken down into two main groups: (1) institutions seeking to create their own pool of arbitrators and processes, and (2) those seeking to create an environment that resembles an online jury. For example, several providers are attempting to use technology (blockchain and smart contracts) as a means to create entirely new mechanisms for appointing decision makers. In Confideal, individuals who wish to serve as arbitrators create accounts presenting their professional resume with a focus on their experience as a decision maker.<sup>463</sup> Within the system arbitrators are ranked in two ways, objective and subjective.<sup>464</sup> Objective factors allow the platform developers to rank decision makers based on traditional metrics such as time taken prior to resolution, while the subjective ranking is derived from feedback from users of the environment.<sup>465</sup> The use of both objective and subjective evaluation factors is common in reputation systems in other environments, and here it improves the reliability of the rankings and hopefully prevents gaming of the system.

In contrast, Juris uses an online polling environment in which multiple individuals participate, mimicking the function of a jury more than a traditional arbitrator.<sup>466</sup> Individuals participate in discussions and are then polled to vote on a recommended resolution.<sup>467</sup> If the parties are unable to resolve their dispute after receiving the recommendation, the process is escalated to a panel of decision makers.<sup>468</sup> Within this environment individuals are ranked, using similar measures to those within the Juris environment, and the individuals with the highest reputation rankings are allowed to decide escalated cases.<sup>469</sup>

---

<sup>462</sup> *See id.*

<sup>463</sup> *See supra* Chapter 3, Table H, at Confideal

<sup>464</sup> *See id.*

<sup>465</sup> *See id.*

<sup>466</sup> *See supra* Chapter 3, Table H, at Juris.

<sup>467</sup> *See id.*

<sup>468</sup> *See id.*

<sup>469</sup> *See id.*

Finally, Kleros provides a crowdsourced online dispute resolution environment that focuses on incentivizing individual jurors.<sup>470</sup> Individuals bid to become a juror using a platform-based coin/token, with those bidding the highest number of tokens having a higher chance of being selected.<sup>471</sup> Individual jurors are then further incentivized to resolve disputes via consensus-based incentives, meaning that individuals vote and those who vote in the same manner as the majority receive rewards, while those who vote against the majority are penalized.<sup>472</sup>

#### D. Online Dispute Resolution Summary Conclusion

While Online Dispute Resolution lacks a clear definition, some patterns can be discerned that allow this review existing online dispute environments to be broadly divided into two categories: public systems annexed to the brick-and-mortar court systems, and private systems that in general adopt arbitration as a final stage in the resolution process. While this differentiation might seem unimportant, it is in fact fundamental to the conversation, as public ODR environments can be assumed to be governed by existing laws relating to any justice environment, while private systems are only regulated in the limited circumstances that any arbitration agreement and award are regulated. As such, the newest ODR private environments may present interesting dilemmas in terms of their commitment to the RoL.

The chapter provided an overview of the better known online environments, the rules applicable to each and the institutional rules put in place to cover the independence and impartiality of the decision makers. Of course, some of the newest systems have moved beyond traditional models which position appointed human decision makers as the final authority, instead embracing the newest thinking and getting to a decision using online jury pools, incentives and the total automation of the entire contract performance and

---

<sup>470</sup> See *supra* Chapter 3, Table H, at Juris. The author has also provided consultation for the Kleros developers.

<sup>471</sup> See *supra* Chapter 3, Table H, at Juris.

<sup>472</sup> See *id.* See also Kleros website, FAQs, available at <https://medium.com/kleros/kleros-frequently-asked-questions-about-peer-to-peer-justice-5a921cb76abe> (last visited Aug. 10, 2019).

dispute process. Time will tell if these providers are embracing RoL and a commitment to ensuring an impartial and independent decision maker, be it human or not.

Because at the time of writing ODR was mimicking ADR environments, this chapter was written without demanding much understanding of technology, other than an understanding of the existing ODR providers that many are undoubtedly familiar with, such as eBay and Amazon. The newest issues that are arising in this area, by contrast, do demand that policy makers, and others, have a basic grasp of the technology and the design process that give rise to them. Chapter 4 seeks to provide just such a background, while Chapter 5 seeks to create best practices and suggests a potential model for deployment.

#### Chapter 4: The Next Step — A Non-Human Actor as an Assistant to the Decision Maker in the Online Justice System

Artificial Intelligence and the Law has existed as a research field since the early 1980s.<sup>473</sup> Programs such as IBM’s Watson serve as examples of the evolving use of AI for “question answering, information extraction, and argument mining from text. . .”<sup>474</sup> Despite incredible advances, the newest technology and algorithms still cannot *yet* perform legal reasoning.<sup>475</sup> Now is the time to get ahead of the curve and consider the design principles that should underlie the AI-based judicial decision making model of the future.

This chapter will explore the next steps in the evolution of online justice environments by considering what arbitration might look like with a *non-human actor* as an assistant or as a full contributor to the decision making process. The chapter will briefly explore relevant history of artificial intelligence and the law, and review traditional technology-driven approaches to design, by explaining software models, machine learning, various types of model risk and approaches to the reduction of model risk. It will then describe examples of emergent systems, to assist in a fuller understanding of what technology

---

<sup>473</sup> See Ashley, *AI AND LEGAL ANALYTICS*, *supra* note 135 at 3.

<sup>474</sup> *Id.*

<sup>475</sup> *See id.*

can actually accomplish, and will conclude by providing a substantive analysis of two key aspects of model design: transparency and socio-technical aspects of use. The information in this chapter forms the foundation for the development of the best practices and a potential model presented in Chapter 5.

## A. The Basics of AI and the Law

As automated data analysis replaces human supervision and intuition in decision making, there is a growing need to incorporate many of the concepts discussed in this thesis into algorithmic decision making frameworks. In particular, the concepts of dependency, partiality, bias, transparency and accountability will have to be built into new technologies. This discussion will begin by offering a basic orientation to how these concepts fit into algorithmic models.

Artificial Intelligence (AI) is generally defined as the “capacity of a computer to perform operations analogous to learning and decision making in humans.”<sup>476</sup> As such, AI is a computer-based learning and decision making operation. In 1956 at the Dartmouth Artificial Intelligence Conference, the technology was described as: “Every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.”<sup>477</sup> This sentence captures a variety of nuances that underlie the original AI “mission of endowing computers with human intelligence.”<sup>478</sup> For example, as noted data scientist Fatemeh Zahedi writes: “AI focuses on how a person logically justifies actions and decisions, and attempts to give the machine a similar logical and reasoning capability.”<sup>479</sup> Crucially, for our discussion, Zahedi explains:

---

<sup>476</sup> The definition of AI is a contentious one. See William Vorhies, *What Exactly Is Artificial Intelligence And Why Is It Driving Me Crazy*, DATA SCIENCE CENTRAL, (Jan. 23, 2018) (discussing the varied definitions).

<sup>477</sup> Martin Childs, *John McCarthy: Computer Scientist Known As The Father Of AI*, INDEPENDENT, (Nov. 2011).

<sup>478</sup> Fatemeh Zahedi, *An Introduction to Neural Networks and a Comparison with Artificial Intelligence and Expert Systems*, Interfaces, Vol. 21, No. 2, 25-38, (Mar. - Apr., 1991).

<sup>479</sup> Zahedi, *Introduction to Neural Networks*, *supra* note 445 at 34. In contrast to neural network researchers who “opted for simulating human intelligence by studying how the brain functions.” *Id.* at 34

AI has a macro scope, treating the brain as a black box. Its structure is closer to models of human decision making. In decision science and cognitive science, researchers focus on actions, underlying motives, and the formal process that links the actions to motives. The axiom of these sciences is logic, its consistency, and its completeness.<sup>480</sup>

AI can be further subdivided based on the tasks that the mechanism is deployed to resolve. In general, current uses of AI can be subdivided into two broad areas, “Artificial Narrow Intelligence (Weak AI) which is AI that specializes in one area,”<sup>481</sup> and “Artificial General Intelligence (Strong AI or Human-Level AI) [referring to] a computer that is as smart as a human across the board, a machine that can perform any intellectual task that a human being can.”<sup>482</sup> We have yet to successfully develop anything even close to true AGI, but in 2014 a survey of scientists asked when they believed we would “more likely than not to have reached AGI?”<sup>483</sup> Their estimate: 2040.<sup>484</sup>

In the meantime automation, defined in our discussion in Chapter 1 as the “technique, method, or system of operating or controlling a process by highly automatic means, as by electronic devices, reducing human intervention to a minimum” is already widely used in many decision making processes. Paul Beswick, Global Head, Digital Labs at Oliver Wyman, explains:

Many companies will look for the potential to automate their own internal processes and decisions, but you can also directly automate customer interactions, and the prize in terms of lower costs and an enhanced experience is often even greater. Rather than connecting with a broker, for example, who then shops for an insurance policy, a customer can input data and get a direct and immediate quote from a specific carrier. . . Instead of talking to a bank manager, potential borrowers can connect their business accounts to the bank, proving their financial position, and get immediate approval for loans up to, say, \$50K, with only exceptions directed to a bank officer.<sup>485</sup>

---

<sup>480</sup> Zahedi, *An Introduction to Neural Networks*, *supra* note 445 at 35

<sup>481</sup> Tim Urban, *The AI Revolution: How Far Away Are Our Robot Overlords?*, GIZMODO (Feb. 9, 2015) available at <https://gizmodo.com/the-ai-revolution-how-far-away-are-our-robot-overlords-1684199433> (last visited Aug. 4, 2019).

<sup>482</sup> *Id.*

<sup>483</sup> *Id.*

<sup>484</sup> *Id.*

<sup>485</sup> Oliver Wyman Digital Labs, *Automated Decision Making: Manage By Exception And Reserve Brainpower For The Tough Call*, Oliver Wyman Blog, (Aug, 2016).

Individuals will likely already have experienced basic automated decision making: when shopping for an insurance policy one can now enter data and get an immediate quote, in contrast to the prior system that required a phone call with an insurance broker; a bank can automate loan accreditation by using the automated system to verify income and financial position, the primary factors in determining the loan amount.

Until recently, it was accurate to describe the “processing method in AI [as] inherently serial,”<sup>486</sup> often best understood as “knowledge coded in in the form of IF–THEN rules.”<sup>487</sup> This requires the knowledge to be “stored in an explicit manner in the form of production rules, frames, objects, or semantic nets.”<sup>488</sup> All knowledge creation occurs outside the AI system,<sup>489</sup> while the system deploys a deductive reasoning protocol to the knowledge it is fed to arrive at a conclusion.<sup>490</sup> It is only recently that “modern” AI has really been able to move beyond this kind of logic-based deductive reasoning.<sup>491</sup> The more sophisticated AI of today has the capacity known as “machine learning.” The core principle of machine learning is that machines take data inputs and develop adapted protocols, *i.e.* “learn,” for themselves. This, for example, “makes it possible to ask human language questions of broad data sets like MRI images and have [the AI] determine if that is cancer or not cancer.”<sup>492</sup>

There are several different ways to “teach” AI, which we will examine in order to better understand its capacities and limitations. According to data scientist William Vorhies machine learning “is using a variety of supervised and unsupervised techniques to find, exploit, and optimize actions based on patterns we are able to find in the data.”<sup>493</sup>

According to Greg Corrado from Google:

Machine learning *systems* are made up of three major parts: (1) Model: the system that makes predictions or identifications; (2) Parameters: the signals or factors used by the model to form its decisions; and (3)

---

<sup>486</sup> Zahedi, *An Introduction to Neural Networks*, *supra* note 445 at 35.

<sup>487</sup> *Id.*

<sup>488</sup> *Id.*

<sup>489</sup> *See id.*

<sup>490</sup> *See id.*

<sup>491</sup> *See* Zahedi, *An Introduction to Neural Networks*, *supra* note 445 at 36.

<sup>492</sup> Vorhies, *What Exactly is Artificial Intelligence*, *supra* note 443.

<sup>493</sup> *Id.*

Learner: the system that adjusts the parameters — and in turn the model — by looking at differences in predictions versus actual outcome.<sup>494</sup>

Let us briefly delve into each of these three aspects of machine learning systems.

## 1. Models

Computer models seek to break down a complex set of human tasks into discrete steps.

According to noted authority Kevin Ashley:

The models specify how a problem is input and the type of [legal] result to output. In between, the *model builders* have constructed a computational mechanism to apply domain knowledge to perform the steps and transform the inputs to outputs.<sup>495</sup>

It is important to note the italicized wording: “model builders,” that is the humans who develop the model from steps to interpretations to parameters.

In general, a model builder does not just copy a schematic, but rather is guided by *rules of action*.<sup>496</sup> They begin from a set of principles, asking questions and developing requirements specific to the process at hand.<sup>497</sup> The various issues explored in this thesis constitute principles that would necessarily generate rules of action for model design: the use of a neutral decision maker, the use of legal formalism or legal realism, the importance placed on independence and impartiality.<sup>498</sup> Moreover, a model builder would need to determine the parameters that define clear cases of each, as well as the nature of outliers.<sup>499</sup>

From these guiding principles the model builder derives parameters, assumptions, constraints and boundaries as an initial set of design considerations.<sup>500</sup> Parameters are

---

<sup>494</sup> Danny Sullivan, *How Machine Learning Works, As Explained By Google*, *supra* note 1.

<sup>495</sup> See Ashley, *AI AND LEGAL ANALYTICS*, *supra* note 135 at 4 (emphasis added).

<sup>496</sup> See *id.*

<sup>497</sup> See *id.*

<sup>498</sup> See Chapter 1, *supra* at sections

<sup>499</sup> See *id.* Each of these discussions will be used in Chapter 5 to create a set of principles that will guide the model builders when constructing the decision maker in online justice environment. See *infra* Chapter 5, Section A.

<sup>500</sup> See Ashley, *supra* note 135 at 4-8.



an attempt to break down the principles into a systematic set of rules.<sup>501</sup> Often, to create these rules, the model builder must make some assumptions. An assumption, to a model builder, is basically a truth, that was taken as the truth — for the purposes of this model, absent any further evidence.<sup>502</sup> For example, Jon Kleinberg, during his keynote at the ACM Conference on Fairness, Accountability, and Transparency entitled *Fairness, Rankings, and Behavioral Bias*, raised the question of the evaluation of the Rooney Rule.<sup>503</sup> The Rooney Rule is a well-known National Football League policy that “requires league teams to interview ethnic-minority candidates for head coaching and senior football operation jobs.” When discussing the real world application of the rule, Kleinberg noted that in many instances of examining the impact on affirmative action in the NFL, the mathematical model builder “assumed” a single group of applicants.<sup>504</sup> The assumption, while not really universally true (any single group could easily be divided into sub-groups), was taken as the truth for the purposes of the evaluation.<sup>505</sup> Of course, if evidence arises that alternatives to the assumption not only exist but are likely impactful, the model is adjusted. Moreover, some assumptions are industry-specific, based on collective understandings,<sup>506</sup> these types of assumptions may become endemic to all models in a given sector.<sup>507</sup>

In explaining the basics of model building, we can look to existing models for development and lessons. For example, the finance industry has used a model-based approach to institutional risk, and their models have already been subjected to regulation.<sup>508</sup> For regulatory purposes, a model in the financial sector is defined as “a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into

---

<sup>501</sup> See Ashley, *AI AND LEGAL ANALYTICS*, *supra* note 135 at 4.

<sup>502</sup> *See id.*

<sup>503</sup> See Jon Kleinberg, *Fairness, Rankings, and Behavioral Bias*, Key Note Speaker, ACM FAT\*2019, available at [https://fatconference.org/2019/livestream\\_capitol.html](https://fatconference.org/2019/livestream_capitol.html) (last visited Aug. 5, 2019).

<sup>504</sup> *See id.*

<sup>505</sup> *See id.*

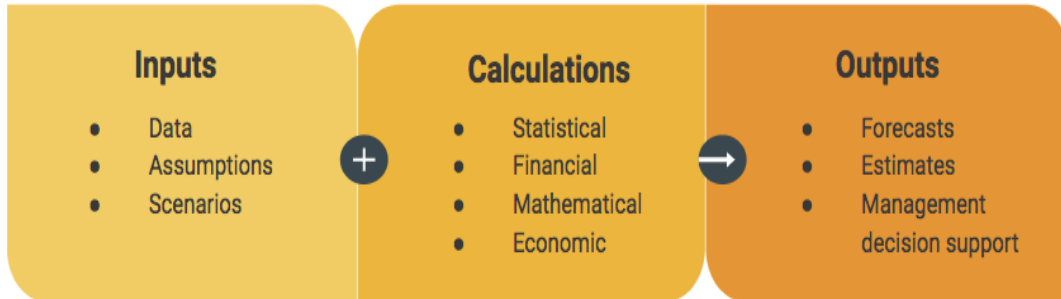
<sup>506</sup> *See id.*

<sup>507</sup> *See id.*

<sup>508</sup> See Seph Mard, *What’s Model Risk and Why Does It Matter*, DataRobot (Jan. 2018)(explaining model risk).

quantitative estimates.”<sup>509</sup> *Data Robot* author Seph Mard uses the following schematic to represent this model building:

*Schematic 1: Financial Institution Model*



Taken from Mard, *What’s Model Risk*<sup>510</sup>

The model builder has used a set of principles (not captured in the schematic) to make decisions about the appropriate model to be deployed.<sup>511</sup> To accomplish this particular task — that is to measure financial institutional stability — the model builder will use a set of financial assumptions and pre-determined financial scenarios.<sup>512</sup>

The schematic progresses to show the “transformative” processes that occur — in this financial model that is the application of a set of statistical, financial, mathematical and/or economic formulas.<sup>513</sup> These parameters (in this case, calculations) “transform” the inputs represented in schematic box one into the outputs represented in schematic box three.<sup>514</sup> And of course, as a final note — the outcomes are then monitored so that adjustments can be made to the model — this “learning” process will be discussed in detail below.

In terms of *legal decision making models*, the field is so new that the community of developers have yet to arrive at standardized assumptions. In April of 2018 the

---

<sup>509</sup> *Id.*

<sup>510</sup> See Mard, *What’s Model Risk*, *supra* note 468.

<sup>511</sup> See *id.*

<sup>512</sup> See *id.*

<sup>513</sup> See *id.*

<sup>514</sup> See *id.*

European Research Council awarded Vrije Universiteit Brussel Research-Professor Mireille Hildebrandt a foundational research grant to come up with a new legal hermeneutics to create “a framework for lawyers to approach computational law architectures intelligently; to understand limitations and implications, and be able to ask the right questions to assess technologies that are increasingly being put to work assessing us.”<sup>515</sup> Although the area is not necessarily new, detailed research in this area is just beginning — and more research, commentary and development will undoubtedly be coming in the near future. However, it is too early to truly anticipate any advanced legal decision making models.

## 2. Learning

Ronald van Loon of Data Science Central describes three types of learning: supervised, unsupervised and reinforcement.<sup>516</sup> In terms of *supervised learning*, which might be more familiar to a wider audience as the Bruner, Goodnow, & Austin phrase *concept learning*, which is “the search for and listing of attributes that can be used to distinguish exemplars from non-exemplars of various categories.”<sup>517</sup> In many ways, supervised machine learning seeks to do the exact same thing, with a computer process. In supervised learning, the desired output is already known.<sup>518</sup> As such, “all that a system needs to do is to work out the steps or process needed to reach from the input to the output.”<sup>519</sup> In this type of learning, “All of the input, the output, the algorithm, and the scenario are being provided by humans.”<sup>520</sup>

---

<sup>515</sup> Natasha Lomas, *What Do AI And Blockchain Mean For The Rule Of Law?*, TECHCRUNCH (April 2018). In fact, we are still in the very early stages, thus it is difficult to “address such questions as how to represent what a legal rule means. . . how to distinguish ‘hard’ from ‘easy’ legal issues, and the roles that cases and values play in interpreting legal rules.” Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 4.

<sup>516</sup> Ronald van Loon, *Machine Learning Explained: Understanding Supervised, Unsupervised, and Reinforcement Learning*, DATA SCIENCE CENTRAL (Jan. 16, 2018).

<sup>517</sup> J.S. Bruner, J.J. Goodnow, & G.A. Austin, A STUDY OF THINKING, Science Editions (1967).

<sup>518</sup> *See id.*

<sup>519</sup> Ronald van Loon, *Machine Learning Explained*, *supra* note 476.

<sup>520</sup> *Id.*

As I described in my paper, *Building a Better HAL 9000: Algorithms, the Market, and the Need to Prevent the Engraining of Bias*:<sup>521</sup> the SAS Institute notes supervised learning:

Supervised learning algorithms are trained using labeled examples, such as an input where the desired output is known. For example, a piece of equipment could have data points labeled either “F” (failed) or “R” (runs). The learning algorithm receives a set of inputs along with the corresponding correct outputs, and the algorithm learns by comparing its actual output with correct outputs to find errors. It then modifies the model accordingly. Through methods like classification, regression, prediction and gradient boosting, supervised learning uses patterns to predict the values of the label on additional unlabeled data.<sup>522</sup>

As summarized by Michele Sebag, supervised learning occurs by “training a classifier from examples labelled by the expert in order to automatically label further examples.”<sup>523</sup> Supervised learning accounts for approximately seventy percent of machine learning applications.<sup>524</sup> It is commonly used in applications where historical data predicts likely future events.<sup>525</sup> For example, it can anticipate when credit card transactions are likely to be fraudulent or which insurance customer is likely to file a claim.<sup>526</sup>

In contrast:

Unsupervised learning is used against data that has no historical labels. The system is not told the "right answer." The algorithm must figure out what is being shown. The goal is to explore the data and find some structure within.<sup>527</sup>

---

<sup>521</sup> See Anjanette H. Raymond, Emma Arrington Stone Young, and Scott J. Shackelford, *Building a Better HAL 9000: Algorithms, the Market, and the Need to Prevent the Engraining of Bias*, 15 NORTHWESTERN. J. TECH. & INTELL. PROP. 215, 221-2, (2018).

<sup>522</sup> *Machine Learning, What it is and Why it Matters*, SAS INSTITUTE, available at [http://www.sas.com/en\\_us/insights/analytics/machine-learning.html](http://www.sas.com/en_us/insights/analytics/machine-learning.html) (last visited Aug. 10, 2019).

<sup>523</sup> Michele Sebag, *A Tour Of Machine Learning: An AI Perspective*, AI Communications 27, 11–23 (2014).

<sup>524</sup> See *id.*

<sup>525</sup> See *id.*

<sup>526</sup> See *id.*

<sup>527</sup> *Machine Learning, What it is and Why it Matters*, *supra* note 481.

During the process of unsupervised learning, the system does not have concrete data sets; in fact, the outcomes to most of the problems are largely unknown. As science fiction as it may sound:

unsupervised learning has the ability to interpret and find solutions to a limitless amount of data, through the input data and the binary logic mechanism present in all computer systems. The system has no reference data at all.<sup>528</sup>

Unsupervised learning, (exploratory data analysis), is less advanced as a field than supervised learning.<sup>529</sup>

Following a very different path, reinforcement learning is built upon the concept that the optimal behavior or action is reinforced by a positive reward.<sup>530</sup>

Depending on the complexity of the problem, reinforcement learning algorithms can keep adapting to the environment over time if necessary in order to maximize the reward in the long-term.<sup>531</sup>

In reinforced learning training models, rewards and consequences are established first, and as the algorithm attempts to accomplish a prescribed task, it executes different paths. The algorithm learns when it receives a reward or punishment.<sup>532</sup> The system then tries again to achieve the outcome and again learns from success or failure.<sup>533</sup> Unfortunately, reinforced learning is time consuming and requires a lot of data, hence, it is most often used in areas such as gameplay and robotics.<sup>534</sup>

The type of learning that will be deployed is a key consideration when evaluating potential technologies, as this is one of the mechanisms that will drive system evolution and thus introduce new issues that may not have been foreseen in the original design.<sup>535</sup> For example, supervised learning indicates that the designers have an intended outcome in mind at the time of creation; the data set is deployed so as to build the process

---

<sup>528</sup> Ronald van Loon, *Machine Learning Explained*, *supra* note 476.

<sup>529</sup> See Sebag, *A Tour Of Machine Learning*, *supra* note 482 at 17.

<sup>530</sup> Bernard Marr, *Artificial Intelligence: What Is Reinforcement Learning - A Simple Explanation & Practical Examples*, FORBES, (Sept, 2018).

<sup>531</sup> *Id.*

<sup>532</sup> *See id.*

<sup>533</sup> *See id.*

<sup>534</sup> *See id.*

<sup>535</sup> *See id.*

necessary to produce the desired outcome. Both the outcome and the data set can be infused with unintended and implied bias, by designers or others. And, while some bias can be reduced (as we will see below), other forms of bias are difficult to design around. This is especially true of unintentional bias.

Unsupervised learning, where the outcome is unknown at the time of design, offers incredibly powerful potential, but comes with its own issues. In these cases the designers have a data set. They deploy a process upon the data set to discover patterns and clusters. These discoveries are then used to create rules as the system learns. While this may seem an ideal way to arrive at unbiased conclusions, this type of deployment is incredibly dependent on “good” data. Unsurprisingly, data from the justice system contains all of the biases, discrimination and elisions that have historically characterized law and justice.<sup>536</sup> Keep in mind, there is nothing necessarily wrong with unsupervised learning as such. In fact, often the environment can highlight and make concrete the very clusters and patterns of bias that social scientists have long known to exist in the justice system, but that practitioners rarely actively identify. However, it is imperative that designers be aware of the potential for unsupervised learning to amplify and re-embed forms of bias that most would agree we should be seeking to reduce or eliminate. Understanding the type and nature of machine learning at play is fundamental to any discussion of technology in a justice environment.

### 3. Model Risk

Also important for consideration is an emerging area of model development: model risk. Model risk is “the potential for adverse consequences from decisions based on incorrect or misused model outputs and reports.”<sup>537</sup> Model risk is a robust area of consideration for model developers, especially in the areas of bias and discrimination, which in the model development lexicon are often thought of in terms of a measure of

---

<sup>536</sup> See e.g., Samuel R. Sommers, Satia A. Marotta, *Racial Disparities in Legal Outcomes: On Policing, Charging Decisions, and Criminal Trial Proceedings*, POLICY INSIGHTS FROM THE BEHAVIORAL AND BRAIN SCIENCES, Vol. 1(1), 103-111 (2014)(including an in-depth review of the historical aspects and the corresponding research).

<sup>537</sup> See Mard, *What’s Model Risk*, *supra* note 468.

fairness. Hence, it is important to understand some basics in relation to model risk as it arises in areas of key consideration for model selection and monitoring. The next section will turn to the concepts of input risk, model risk, outcomes and complexity. The section following will further break down and address some specific issues related to model abstraction: the Framing Trap, the Portability Trap, the Formalism Trap, the Ripple Effect Trap and the Solutionism Trap. It will conclude with suggestions for reducing some types of model risk.

*a. Model Risk: Biased Inputs*

First, model risk can arise from both incorrect (or incomplete) inputs *and* outputs.<sup>538</sup> It should not be surprising that models built upon data that is incomplete, incorrect, or has been organized without attention to appropriate, well-defined standards, can (and likely will) produce outcomes that are less than accurate.<sup>539</sup> Beyond the narrow question of accuracy in the initial data set lie issues of omission, false and misleading representation, and implicit bias.<sup>540</sup> These are issues that arise not from model selection nor the model builder's own bias, but instead from the use of historical data as the training data.<sup>541</sup> It is also important to understand that, due to the nature of machine learning, outputs become inputs, and so errors and biases in the initial training data do not necessarily remain limited and easily isolated.<sup>542</sup> If the data used for training is biased by the discrimination already taking place in the world it represents, this bias gets reproduced in the outputs and the evolving protocols of the algorithm.<sup>543</sup>

Policing technology deployments can be used as a first point of consideration as the early deployments of technology have revealed troubling trends in technology design when the algorithm relies upon data gathered and used within a prediction model. In the United States, police departments have historically focused disproportionately on

---

<sup>538</sup> See Mard, *What's Model Risk*, *supra* note 468.

<sup>539</sup> See *infra* Chapter 4, Section and Chapter 5, Section C(3). See also Carolyn Healey, *Garbage In, Garbage Out – How Bad Data Hurts Your Business*, Service Objects Website, (Oct. 9, 2018).

<sup>540</sup> See Mard, *What's Model Risk*, *supra* note 468.

<sup>541</sup> See *id.*

<sup>542</sup> See *id.*

<sup>543</sup> See *id.*

communities of color, their data over-represents crime by minorities and under-represents crime by whites. Whether designers intend it or are simply unaware of implicit bias in their data set, the result is the same: *The decision making algorithm inherits and perpetuates the inequalities of the society in which it was created.*<sup>544</sup> This relates back to the nature of supervised machine learning, which:

works by recognizing patterns of relations in big data sets. But to do this, the algorithm has to be trained by feeding it massive amounts of *historical* training data. Using this data, the algorithm constructs a model of the world that is then used to infer similar patterns in new sets of data.<sup>545</sup>

Because algorithms are always trained on historical data, it is virtually impossible to find a “clean” dataset on which an algorithm can be trained to be “bias-free.”<sup>546</sup> The historical biases in the justice system have been studied extensively, and found to be systemic, pervasive and long-lasting.

Because these algorithms are in many ways, the first attempts to create prediction models in justice environments, many scholars look to these as key places to begin discussion. Many argue the absence of bias free prediction models in environments where these models have been used for some time, should give everyone pause of wider deployments in the justice environment. These are frequently cited as individual's argue against the use of technology and prediction as a decision maker.<sup>547</sup>

Some commentators argue that measures may be taken to reduce the issue of data bias:

---

<sup>544</sup> See Henrik Chulu, *Let Us End Algorithmic Discrimination*, MEDIUM, Tech Festival, (Aug. 3, 2018)(discussing the issues and potential approaches to reducing the discriminatory aspects of algorithms).

<sup>545</sup> See *id.*

<sup>546</sup> Lokke Moerel, *Algorithms Can Reduce Discrimination, But Only With Proper Data, Privacy Perspectives*, International Association of Privacy Professionals (2019).

<sup>547</sup> See, e.g., Neil Sahota, Will A.I. Put Lawyers Out Of Business? *Forbes* (Feb. 9, 2019) (reviewing areas of appropriate technology and AI uses while rejecting the use of AI as a decision maker); John Morison, and Adam Harkens, Re-Engineering Justice? Robot Judges, Computerised Courts and (Semi) Automated Legal Decision-Making, *Legal Studies* Volume 39, Issue 4, pp. 618-635 (Dec. 2019) (discussing the role of data, processing and building automated decision making)



To solve this, group indicators such as race, gender, and religion are often removed from the training data. The idea is that if the algorithm cannot “see” these elements, the outcome will not be discriminatory.<sup>548</sup>

This argument however, is based on a misunderstanding of how algorithms work. “The algorithm will soon find derived indicators — proxies — to explain this bias.”<sup>549</sup> For example, eliminating race as a measured attribute fails to be a solution, as the system can combine several other measured data, such as location, community demographics, and participation in community groups as proxies for race.<sup>550</sup> Thus, simple elimination of attributes is not recommended as the complexity of the world in which we live is reflected in a multitude of attributes. Instead, the only solution is to first make biases transparent in the training data.<sup>551</sup> Creating such a robust set of training data requires a lot of work and preplanning. Group indicators have to be collected, usually at the initial time of data collection, so the data can be assessed for biases, and the “algorithm must be trained against selecting these factors, by means of ‘adversarial training.’”<sup>552</sup> The process is difficult to perfect.

Fortunately, designers are working on reducing biases in data sets and learning environments; in fact, it is a hot topic in engineering and system design. In this author’s opinion, policy makers can and must insist that design explicitly commit to recognizing and reducing the impacts of biased data sets in machine learning environments. The best solutions will likely require policy maker involvement, but the efforts of engineers and systems designers will be absolutely fundamental.

---

<sup>548</sup> *Moerel, supra note 497.*

<sup>549</sup> *Id.*

<sup>550</sup> *See id.*

<sup>551</sup> *See id.*

<sup>552</sup> *Id.* In general, adversarial training is a technique used in model building to train systems to identify those that do not belong. When model builder train a system, they feed ‘good’ data into the machine so it learns. But, what happens when ‘bad’ data is feed in, the machine can just accept anything its fed. So, it has to learn to exclude things, which is adversarial training. Its easiest to understand, (in my opinion) in the training done with photos and how the machine identifies people, (or cats, or dogs) and labels them correctly. *See* Adar Kahiri, *Adversarial Training: Creating Real Pictures of Fake People With Machine Learning*, MEDIUM, Towards Data Science, (Dec. 29, 2018).

*b. Model Risk: The Challenge of Social Complexity*

Even setting aside concerns about input data, model risk originates in the existing complexities of society, which can be extremely tricky to represent in the terms required for an algorithm to function. Benthall and Haynes explain:

there is a substantive difference between systems that result in controversial or unfair outcomes due to the racial bias of their designers and those that do so because they are reflecting a society that is organized by racial categories.<sup>553</sup>

This of course, raises the question: can any model ever properly resolve the complexities of society? Benthall and Haynes sharply outline the dilemma:

Machines that attempt to correct unfairness through explicit use of racial classification do so at the risk of reifying racial categories that are inherently unfair. Machine learning systems that allocate resources in ways that are blind to race will reproduce racial inequality in society.<sup>554</sup>

Race represented as a single factor, whether embedded within a model or eliminated as an acceptable data category, will always fail to reflect the complexity of race within society. Hence, having an understanding of social complexity is an essential element of creating valid models.

Modeling social complexity is an under-researched area in computational engineering, but one that is now gaining prominence. Selbst and others identify five errors that arise from the failure “to consider how social context is interlaced with technology in different forms,”<sup>555</sup> and thus the remedies also require a deeper understanding of “the social to resolve problems.”<sup>556</sup> The five errors the community identifies are: the Framing Trap, Portability Trap, Formalism Trap, Ripple Effect Trap, and Solutionism Trap. Each will now be examined in turn.

---

<sup>553</sup> Sebastian Benthall and Bruce D. Haynes, *Racial Categories In Machine Learning*, published in Conference on Fairness, Accountability, and Transparency (FAT\* '19)(2019).

<sup>554</sup> *Id.* at sec. 4.

<sup>555</sup> See Andrew Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, Janet A. Vertesi, *Fairness and Abstraction in Sociotechnical Systems*, 59, published in Conference on Fairness, Accountability, and Transparency (FAT\* '19), (2019).

<sup>556</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at 59.

i. The Framing Trap

The Framing Trap, that is the “failure to model the entire system over which a social criterion, such as fairness, will be enforced”<sup>557</sup> is a common area of focus for designers.

The frame (or view) asks the designer to:

Recognize, explicitly, that a machine learning model is part of a sociotechnical system, and that the other components of the system need to be modeled. By moving decisions made by humans and human institutions within the abstraction boundary, fairness of the system can again be analyzed as an end-to-end property of the sociotechnical frame.<sup>558</sup>

In other words, framing the designer’s responsibility too narrowly in terms of computer code fails to capture the full function and impact of their work. Using a sociotechnical frame allows designers to examine not only the expected deliverable and other standards of technology development, but also all the factors that come into play in deployment in a real-world setting. For example, in our instance, the design of a risk assessment protocol for the justice system must account for how judges *use and respond* to risk assessment scores. Researchers within the field sometimes call this part of the design “human–computer interaction.” Human–computer interaction researchers examine the design and use of computer technology, focusing on the interfaces between people (users) and computers.

Recent research also allows us to take a close look at the manner in which prediction tools are currently being used in the decision making process of a court, in a very narrow category of risk assessments, which are commonly used in within the judicial decision process in the United States. As Green and Chen note: “risk assessments do not make definitive decisions — *they inform judges*, who are the final arbiters.”<sup>559</sup> Green and Chen’s findings ought to lead to adjustments within the model, when humans are asked to apply technology produced predictions, especially when those predictions are intended to influence decision making. For example, their study found several

---

<sup>557</sup> *Id.* at Section 2.1.

<sup>558</sup> *Id.* at 60.

<sup>559</sup> Ben Green and Yiling Chen, *Disparate Interactions: An Algorithm-in- the-Loop Analysis of Fairness in Risk Assessments*, 90, in submission Conference on Fairness, Accountability, and Transparency (FAT\* ’19)(2019).

behaviors that called into question the fairness of the risk assessment.<sup>560</sup> In fact, the participants were unable to effectively evaluate the accuracy of the predictions.<sup>561</sup> And, more importantly, the “use of risk assessments led to higher risk predictions about black defendants and lower risk predictions about white defendants.”<sup>562</sup>

While these findings pertain to a narrow set of circumstances, that being the use of risk predictions, the use of the predictions was in an area of the justice environment where the decisions had significant impact upon individuals. Green and Chen’s work on real-world deployment leads to conclusions will be incredibly important within Chapter 5 as the author builds a neutral decision maker model within a justice environment.

It is also worth noting another field research study, which suggests that while some people may be reluctant to use AI systems that are presented as providing recommendations, when the information is instead presented as additional data to be used within an existing decision making process, people more willingly engage with the use of the AI system.<sup>563</sup> Other studies have shown that judges do not consistently take the recommendations into account,<sup>564</sup> automation bias is sometimes present,<sup>565</sup> and judges deviate from recommendations in biased ways.<sup>566</sup>

---

<sup>560</sup> *Id.* at 98.

<sup>561</sup> *Id.*

<sup>562</sup> *Id.*

<sup>563</sup> See Cliff Kuang, *An Ingenious Approach To Designing AI That Doctors Trust*, Co.Design, (Jan. 17, 2018)(explaining the importance of considering presentation as an influencing choice).

<sup>564</sup> Angèle Christin, *Algorithms In Practice: Comparing Web Journalism And Criminal Justice*, BIG DATA & SOCIETY 4, 2 (2017).

<sup>565</sup> Danielle Keats Citron, *Technological due process*, WASHINGTON UNIVERSITY LAW REVIEW, 85, 1249-1313 (2008).

<sup>566</sup> *Recognizing and Eliminating Bias from Court Operations*, SUPREME COURT STANDING COMMITTEE ON FAIRNESS AND DIVERSITY (Florida) (2008) (“The report indicates that the overall perception of those with long-standing experience in the Florida court system is that, although there have been significant improvements in reducing discrimination, disparate treatment continues to occur.”) *Id.* at 1.

This body of research suggests that algorithm-in-the-loop aids must be considered and evaluated within the context of real world impacts and considered specifically within a sociotechnical context.<sup>567</sup>

ii. The Portability Trap

The Portability Trap occurs when designers fail to “understand how repurposing algorithmic solutions designed for one social context may be misleading, inaccurate, or otherwise do harm when applied to a different context.”<sup>568</sup> The Portability Trap is one of the easier risks to mitigate against, yet it is often an overlooked aspect of risk within the technology community. It is often assumed that a well-designed model can be repurposed or simply undergo minor adjustments, and be used in an adjacent, yet distinct, environment or context.

That is not to write that no algorithm can be appropriately re-purposed. It is instead to highlight the importance of asking questions about the similarities between contexts and applying a full sociotechnical framing treatment to a new deployment of an existing algorithm in order to determine its appropriateness. Of course, the issue becomes how model designers are to know the full context of a model that will be deployed in more than one context.

In response to model risk arising from trained machine learning models being used in instances for which they are not well suited, the Mitchell, Wu, Zaldivar, Barnes, Vasserman, Hutchinson, Spitzer, Raji and Gebru report recommends that released models be accompanied by documentation detailing their performance characteristics.<sup>569</sup> The authors make the following suggestions for relevant information to be captured in the model card: (1) model details, such as basic information about the

---

<sup>567</sup> Green, *Disparate Interaction*, *supra* note 507, at 90.

<sup>568</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at Section 2.2.

<sup>569</sup> Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru, *Model Cards for Model Reporting*, published in Conference on Fairness, Accountability, and Transparency, (2019).

model, (person or organization developing model, model date and model type);<sup>570</sup> (2) information about training algorithms, parameters, fairness constraints or other applied approaches, and features;<sup>571</sup> (3) the intended use;<sup>572</sup> (4) the evaluation data, such as, details on the training dataset used and the dataset(s) used to produce any other quantitative analysis reported on the card;<sup>573</sup> and finally, (5) any ethical considerations or recommendations relevant to deployment.<sup>574</sup>

The transparency of the information and the wide availability of such information could, in theory, improve the use of models and encourage those using models to consider intended use. In addition, in the United Kingdom, the *Data Ethics Framework* has captured a similar idea in a set of principles, guidance and a workbook to guide the design of appropriate data use in the public sector.<sup>575</sup>

More importantly, the information being widely available can allow additional parties to critically evaluate the extended use of the model and to question inappropriate uses. The transparency allows those using the model to be aware of underlying assumptions that formed the model and to determine the acceptable tolerance for assumptions and parameters in the new use.

### iii. The Formalism Trap

---

<sup>570</sup> See *id.* at 222.

<sup>571</sup> See *id.*

<sup>572</sup> See *id.*

<sup>573</sup> See *id.*

<sup>574</sup> See *id.*

<sup>575</sup> United Kingdom, Department for Digital, Culture, Media & Sport, *Government Efficiency, Transparency And Accountability* (June 2018) available at <https://www.gov.uk/government/publications/data-ethics-framework> (last viewed Aug. 10, 2019) Built upon the U.K. Cabinet Office, *Data Science Ethical Framework*, (withdrawn) (2016) available at <https://www.gov.uk/government/publications/data-science-ethical-framework> (last viewed Aug. 10, 2019).

The Formalism Trap occurs when designers fail to “account for the full meaning of social concepts such as fairness, which can be procedural, contextual, and contestable, and cannot be resolved through mathematical formalisms.”<sup>576</sup>

As we have seen, some attributes of the real world are unavailable for use within models, for a variety of reasons, giving rise to work-arounds or proxies.

Generally, any model that imputes the missing protected attribute value based on other, observed variables is known as a proxy model, and such a model that is based on predicting conditional class membership probabilities is known as a probabilistic proxy model.<sup>577</sup>

Unfortunately, some phenomena appear deceptively easy to integrate into a model design this way, masking the challenge of fully appreciating the actual complexity of the concept. Ideally, a proxy should never be used in a manner that oversimplifies and misrepresents complex social conditions. Indeed, it is plausible that some complex social conditions should never be represented by a single attribute or proxy.

However, poorly designed regulation can effectively demand the creation of proxy models, and when designers fail to fully understand the nature of the regulatory request, this can be incredibly problematic. Measuring outcomes of the “decision making system with respect to a protected class, such as gender or race, is challenging when class membership labels are unavailable,”<sup>578</sup> for example. Under U.S. Federal law, lenders must demonstrate that credit decisions comply with fair lending laws.<sup>579</sup> However, gathering information on membership in those protected classes runs afoul of other U.S. law. For instance, “credit card and auto loan companies must demonstrate that the way they extend credit is not racially discriminatory, yet are not allowed to ask

---

<sup>576</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at Section 2.3.

<sup>577</sup> Jiahao Chen, Nathan Kallus, Xiaojie Mao, Geoffry Svacha, and Madeleine Udell. *Fairness Under Unawareness: Assessing Disparity When Protected Class Is Unobserved*, published in Conference on Fairness, Accountability, and Transparency (FAT\* '19), (2019).

<sup>578</sup> *Id.*

<sup>579</sup> Two different federal laws deal with discrimination in lending: the Fair Housing Act (FHAct) and the Equal Credit Opportunity Act (ECOA). These fair lending laws prohibit lenders from discriminating in credit transactions on the basis of race, color, national origin, religion, sex, and other specified.

applicants what race they are when they apply for credit.”<sup>580</sup> This dilemma leaves the creation of a proxy for race as one of the companies’ only viable solutions. Yet race is a prime example of a complex concept that is harmful when over-simplified.

#### iv. The Ripple Effect Trap

The Ripple Effect Trap occurs when designers fail to “understand how the insertion of technology into an existing social system changes the behaviors and embedded values of the pre-existing system.”<sup>581</sup> The ripple effect may be one of the biggest concerns in model design within the emerging technology-driven justice environment. As was surveyed in Chapter 3, very little technology has been deployed in the justice environment, and what has been so far has usually been deployed with an eye toward improving access to justice — thus the individual users are the focus of evaluation. More work needs to be done concerning how other human actors, such as decision makers, interact with the technology in the justice environment.

There has been some limited research in this area: for example, in response to Kentucky mandating that risk assessments be used to inform all pretrial release decisions, Stevenson and a group of researchers set out to measure the impact of the introduction of technology-driven risk assessments on pretrial release.<sup>582</sup> Stevenson found that while an initial impact may have occurred, shortly thereafter, the release rate returned to original numbers.<sup>583</sup> In 2019, Green and Yen performed a study designed to review how risk assessments influence human decisions within criminal justice adjudication.<sup>584</sup> Since risk-based assessments and predictions have been used within the context of the U.S. criminal justice environment for some time now, researchers are beginning to ask how the use of these instruments is working within decision making process. For example, using an Amazon Mechanical Turk–based experiment, Green and Yen

---

<sup>580</sup> Chen, *Fairness Under Unawareness*, *supra* 525 at Section 1.1.

<sup>581</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at Section 2.4.

<sup>582</sup> See Megan Stevenson, *Assessing Risk Assessment in Action*, MINNESOTA LAW REVIEW 103 (forthcoming).

<sup>583</sup> See *id.*

<sup>584</sup> See Green, *Disparate Interaction*, *supra* note 507, at 90.



measured how laypersons make predictions about risk, both with and without the aid of a computer-generated risk assessment.<sup>585</sup> Using over 550 participants, they found that people underperform compared to the risk assessment tool, even when provided with its advice;<sup>586</sup> are unable to evaluate the performance of themselves or of the risk assessment tool;<sup>587</sup> and that participant interactions with the risk assessment itself introduced new forms of bias into decision making.<sup>588</sup> These early studies suggest more research needs to be done in this area to fully appreciate the impact of technology upon a decision maker.

There is a growing research interest in how humans use technology-driven predictions, which can be broken down into two main areas: information filters and the way the prediction is incorporated into the decision making process.

First, information filters have been the topic of decades of research across many disciplines and cultures.<sup>589</sup> In general, the research suggests that information presumed to help people make fairer decisions can in fact fail, as individuals tend to filter information through their existing biases.<sup>590</sup> For example, policies preventing employers from asking job applicants whether they have a criminal record (“ban the box”) have actually increased racial discrimination because employers rely more on pre-existing stereotypes<sup>591</sup> and overestimate how many black applicants have criminal

---

<sup>585</sup> See Green, *Disparate Interaction*, *supra* note 507, at 90.

<sup>586</sup> See *id.*

<sup>587</sup> See *id.*

<sup>588</sup> See *id.*

<sup>589</sup> See e.g., Uri Hanani, Bracha Shapira, Peretz Shoval, *Information Filtering: Overview Of Issues, Research And Systems*, USER MODELING AND USER-ADAPTED INTERACTION, 11, 203–259 (2001)(creating a framework for considering issues within the growing area).

<sup>590</sup> See e.g., Giovanni Luca Ciampaglia, Filippo Menczer, *Biases Make People Vulnerable to Misinformation Spread by Social Media*, THE CONVERSATION, U.S. (June 21, 2018)(describing ground breaking research conducted at Indiana University to help fight impact of bias and silos).

<sup>591</sup> See Amanda Agan and Sonja Starr, *Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment*, THE QUARTERLY JOURNAL OF ECONOMICS 133, 1 191-235 (2017).

records.<sup>592</sup> In another troubling instance, interpretations of police-worn body camera footage seem to be directly related to the interpreting individual's prior attitudes about police.<sup>593</sup> Specific to our interest is research conducted concerning decision makers, and again there is a growing interest in the area. It is well documented that judges harbor implicit biases. For example, in 1997 Trope and Thompson developed a study that explored the degree to which preexisting stereotypes determined the willingness of a judging subject to seek individuating information before making a decision.<sup>594</sup> They found that

when negatively stereotyped people are asked questions, they are asked fewer questions, and the questions are asked in a way that tends to elicit confirmation of the stereotype rather than information that would individuate the subject or challenge negative stereotyping. In sharp contrast, questions to positively stereotyped people are more symmetric and therefore likely to elicit responses that would either confirm or disconfirm a particular stereotype.<sup>595</sup>

Trope and Thompson highlight: "In essence, our participants gave stereotyped targets relatively few opportunities to express their personal views on the issues at hand."<sup>596</sup>

As Professor Mills suggests:

Trope and Thompson's study suggests that a judges' formula for approaching . . . cases involving people who are easily stereotyped, such as African-American women . . . may be influenced by their unconscious tendency to search only for a confirmation of the stereotype rather than for a more complete picture. . . <sup>597</sup>

---

<sup>592</sup> See Jennifer L. Doleac and Benjamin Hansen, *The Unintended Consequences of 'Ban the Box': Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden*, (2018) available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2812811](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2812811) (last visited Aug. 5, 2019).

<sup>593</sup> See Roseanna Sommers, *Will Putting Cameras on Police Reduce Polarization?*, *YALE LAW JOURNAL* 125, 1304, 1309 (2016).

<sup>594</sup> See Linda G. Mills, *A PENCHANT FOR PREJUDICE: UNRAVELING BIAS IN JUDICIAL DECISION-MAKING*, 14 University of Michigan Press, (1999).

<sup>595</sup> *Id.* at 14, relying upon Yaacov Trope and Eric Thompson, *Looking For The Truth In All The Wrong Places? Asymmetric Search Of Individuating Information About Stereotyped Group Members*, *JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY*, 73, 229–241, 240 (1997).

<sup>596</sup> *Id.*

<sup>597</sup> Mills, *Penant for Prejudice*, *supra* note 542 at 14.

The impact of this kind of bias is severe: at least one study has found that white judges give harsher sentences to black defendants than white ones, even when they have committed the same crime and received the same score from the formula used to set criminal punishments.<sup>598</sup> Moreover, various research projects have found that judges often make initial intuitive judgments which they might, or might not, override with deliberation.<sup>599</sup> Judges are influenced by irrelevant numerical anchors,<sup>600</sup> the way outcomes are framed,<sup>601</sup> and even irrelevant emotional cues.<sup>602</sup>

Such instances of bias are not limited to courtrooms. In fact, at least one study has found somewhat similar patterns in the decision making of arbitrators.<sup>603</sup> The authors of the study comment:

In this first-ever psychological experiment involving international arbitration, we found that arbitrators often made intuitive and impressionistic decisions rather than the fully rational and deliberative decisions that might be normatively desirable. This finding, though

---

<sup>598</sup> See Josh Salman, Emily Le Coz, and Elizabeth Johnson, *Florida's Broken Sentencing System*, SARASOTA HERALD-TRIBUNE (2016).

<sup>599</sup> See generally, Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L.REV.1 (2007); Chris Guthrie, Jeffrey J. Rachlinski & Andrew J. Wistrich, *The "Hidden Judiciary": An Empirical Examination of Executive Branch Justice*, 58 DUKE L.J. 1477 (2009).

<sup>600</sup> Jeffrey J. Rachlinski, Andrew J. Wistrich, & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 INDIANA L.J. 695 (2015).

<sup>601</sup> See e.g., Birte Englich, Thomas Mussweiler and Fritz Strack, *Playing Dice With Criminal Sentences: The Influence of Irrelevant Anchors on Experts' Judicial Decision Making*, PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN, 32:2, 188-200, (2016).

<sup>602</sup> See Jeffrey J. Rachlinski, Andrew J. Wistrich & Chris Guthrie, *Altering Attention in Adjudication*, 60 UCLA L. REV.1586 (2013) (identifying how directing judicial attention shapes outcomes); Jeffrey J. Rachlinski, Chris Guthrie & Andrew J. Wistrich, *Contrition in the Courtroom: Do Apologies Affect Adjudication?*, 98 CORNELL L.REV.1189 (2013) (finding apologies encourage judges to be more lenient); Andrew J. Wistrich, Jeffrey J. Rachlinski & Chris Guthrie, *Heart Versus Head: Do Judges Follow the Law or Follow Their Feelings?*, 93 TEXAS L. REV.855, 862(2015) (finding that "judges' feelings about litigants influence their judgments").

<sup>603</sup> See Susan D. Franck and Anne van Aaken, and James Freda, and Chris Guthrie, Chris and Jeffrey Rachlinski, *Inside the Arbitrator's Mind*, EMORY LAW JOURNAL, Vol. 66, (2017)(attempting to frame the logic behind opinion driven decision making).

perhaps disappointing, is unsurprising. Arbitrators are people, and they make judgments and decisions the way other people do.<sup>604</sup>

The biases and external, non-case-specific influences occur, regardless of if the decision maker is a human or a model build upon existing data sets.

While the aforementioned examples suggest that humans often favor their own biased intuition over more objective factors, a second feature of the Ripple Effect Trap arises from granting too *much* credence to computational outputs. Humans often suffer from what has come to be known automation bias. Automation bias arises most commonly in two ways, omission errors, and commission errors. Omission errors occur when “people do not recognize when automated systems err.”<sup>605</sup> Commission errors occur when “people follow automated systems without considering contradictory information.”<sup>606</sup> Numerous studies have demonstrated the presence of automation bias in multiple environments. For example, automated systems can alter a person’s relationship to a task by creating what is commonly called a “moral buffer” distancing the individual from the impact of their decisions.<sup>607</sup> Cummins posits that automated decision support tools can cause individuals to believe the automation is “in charge,” which in turn causes them to abandon their sense of responsibility.<sup>608</sup> Moreover, people prefer to rely upon other peoples’ judgement, even in the face of accurate algorithms.<sup>609</sup>

---

<sup>604</sup> Franck, *Inside the Arbitrator's Mind*, *supra* note 551 at 1169.

<sup>605</sup> See Kathleen L. Mosier, Melisa Dunbar, Lori McDonnell, Linda J. Skitka, Mark Burdick, and Bonnie Rosenblatt, *Automation Bias and Errors: Are Teams Better than Individuals?*, in Proceedings of the Human Factors and Ergonomics Society Annual Meeting 42, 3, 201-205 (1998).

<sup>606</sup> *Id.* at 201.

<sup>607</sup> *Id.*

<sup>608</sup> See Mary L. Cummings, *Automation and Accountability in Decision Support System*, THE JOURNAL OF TECHNOLOGY STUDIES, Vol 32(1) 23-30, 25 ( 2006).

<sup>609</sup> See Joa Sang Lim and Marcus O’Connor, *Judgmental Adjustment of Initial Forecasts: Its Effectiveness and Biases*, JOURNAL OF BEHAVIORAL DECISION MAKING 8, 3, 149-168, 149, (1995).

In fact, in general individuals cannot consistently distinguish between reliable and unreliable predictions<sup>610</sup> and many deviate incorrectly from algorithmic forecasts.<sup>611</sup>

Clearly, existing research is available to guide designers when considering the manner in which a decision maker will be impacted by deployment of technology, especially algorithmic prediction technology. Existing research has already lead to questions about the use of technology, specifically risk predictors, in the decision making portion of the justice environment. More research is needed to fully assess potential negative impacts in justice environments. Thus, designers must consider the user of the technology, the manner of use, the design and the method of delivery of the prediction or recommendation.

#### v. The Solutionism Trap

The Solutionism Trap occurs when designers fail to “recognize the possibility that the best solution to a problem may not involve technology.”<sup>612</sup> The Solutionism Trap is one that few are ready to consider in the entrepreneurial technology-driven world in which we current live. It is however, time to give real thought to the ultimate usefulness of technology that is built upon historically biased data, and/or upon a poorly designed model that seeks to capture complex societal problems in an oversimplified set of attributes, or with several other significant problems that have been reviewed in this chapter. While the author is not yet ready to insist that algorithms should not be used in the justice environment at all, the crucial question is: if it is to be used, how can the system as a whole be designed to mitigate risk and avert negative impact on the system’s users and those whose fate may determine by its outputs? A neutral decision maker can in fact use technology as an aid to make decisions, if the system recognizes and accounts for the concerns elaborated above. However, a watchful eye must be kept

---

<sup>610</sup> See Paul Goodwin and Robert Fildes, *Judgmental Forecasts of Time Series Affected by Special Events: Does Providing a Statistical Forecast Improve Accuracy?* JOURNAL OF BEHAVIORAL DECISION MAKING Vol. 12(1), 37-53, 37 (1999).

<sup>611</sup> See Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey, *Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them*, MANAGEMENT SCIENCE, Vol. 64,(3) (2016).

<sup>612</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at Section 2.5.

on the deployment and, ultimately, model designers and justice environment monitors must be ready to jettison the use of technology if it is found that it cannot be deployed without negative impact.

Ultimately, adherence to the rule of law clearly demands consideration being given to the use of technology within the decision maker portion of the technology-driven justice environment. As we update the justice system, rather than aiming to replace human decision making with technology, we should aim to use technology in order to reduce burden, guide, assist, and educate human decision makers.

## B. Technology Based Decision Makers

In the recent boom in technology-based decision making,<sup>613</sup> three distinct approaches to the use of technology have emerged: (1) expert systems, (2) argument retrieval, and (3) cognitive computing.<sup>614</sup> Each will be discussed in turn.

### 1. Expert Systems

---

<sup>613</sup> This should be distinguished from some of the more well established, lower order processes that have been occurring for some time now. As Tim Pullan, CEO and Founder of legal AI and risk analysis company, ThoughtRiver notes:

AI-driven clause search in M&A and lease review is now an established process for many law firms. It may soon become a de facto minimum professional service requirement to deploy AI in certain types of work. But in many cases this technology deployment is really just a way of lawyers producing the same thing for less. . . new technology . . . has the capability to help lawyers produce more insightful analysis of a type never before served up as advice in the legal industry.

Author, *ThoughtRiver: Legal AI + The Big Changes Ahead For 2018*, ARTIFICIAL LAWYER, (Jan. 18, 2018).

<sup>614</sup> See Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 8 and 11.

Expert systems deal with a narrow area of law and are designed with enough nuance and heuristic<sup>615</sup> knowledge to ask the user pertinent questions about the problem and then customize an appropriate answer.

As noted by Davis in 1984:

To build expert systems is to attempt to capture rare or important expertise and embody it in computer programs. It is done by talking to the people who have that expertise. In one sense building expert systems is a form of intellectual cloning. Expert system builders, the knowledge engineers, find out from experts what they know and how they use their knowledge to solve problems. Once this debriefing is done, the expert system builders *incorporate the knowledge and expertise in computer programs*, making the knowledge and expertise easily replicated, readily distributed, and essentially immortal.<sup>616</sup>

Expert systems have been part of the growing AI movement for many years now — and in many industries. In the legal sector, the growth occurred primarily in the 1980s,<sup>617</sup> at which time it was rejected by many noted scholars.<sup>618</sup> However, recent advances in several areas of technology<sup>619</sup> have allowed somewhat of a resurgence of legal expert systems.<sup>620</sup>

---

<sup>615</sup> Heuristics, in this instance, is thought of as ‘rules of thumb,’ frequently used by skilled practitioners use to in applying specific rules to the facts at hand. *See* Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 4.

<sup>616</sup> Randall Davis, *Amplifying Expertise With Expert Systems*, p. 18, in Winston P.H., & Prendergast K.A., (eds), THE AI BUSINESS: COMMERCIAL USES OF ARTIFICIAL INTELLIGENCE, MIT Press, Cambridge, Mass (1984).

<sup>617</sup> *See* Philip Leith, *The Rise And Fall Of The Legal Expert System*, EUROPEAN JOURNAL OF LAW AND TECHNOLOGY, Vol 1, Issue 1, 2010.

<sup>618</sup> *See e.g. id.*

<sup>619</sup> For example, text analytics. *See* Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 4-6.

<sup>620</sup> The firms announcing the growing use of AI is difficult to keep up with in terms of numbers. For example, in just a short period in 2017 the firms Chapman Tripp, New Zealand’s largest full-service commercial law firm. *See* Author, *Top Kiwi Firm, Chapman Tripp, Picks Luminance’s Legal AI Tech*, ARTIFICIAL LAWYER, (Nov. 20, 2017); Addleshaw Goddard. *See* Author, *Addleshaw Goddard Innovation Head: Baking In Legal AI*, ARTIFICIAL LAWYER, (Nov. 20, 2017); U.S. law firm, Latham & Watkins, *see* Author, *Latham & Watkins Picks Kira Systems For Legal AI Review Work*, ARTIFICIAL LAWYER, (Nov. 28, 2017); Slaughter and May. *See* Author, *Legal AI Co. Luminance Gets \$10m Funding, Brings in Ex-MI5 Boss*, ARTIFICIAL LAWYER, (Nov. 29, 2017); White Shoe, Davis Polk & Wardwell, Latham & Watkins. *See* Author, *The Legal AI ‘Barbarians’ Have Already Taken the Gates*, ARTIFICIAL

Expert systems can perform case analysis and prediction by using “natural language processing (NLP) to filter key data from litigation documents and then leverage this information into actionable and searchable insights for users.”<sup>621</sup> For example, in November of 2017 Berkeley Bridge announced a “downloadable platform that allows lawyers to build their own rules-based, expert systems with minimal need for outside input.”<sup>622</sup> Several firms have already begun trials with this system.<sup>623</sup> Global law firm Linklaters was an integral part of the creation of the newest player in the market — Eigen Technologies.<sup>624</sup> Eigen describes itself as a technology company on the forefront of AI-based legal review services,<sup>625</sup> but this phrasing does not do justice to the full scope of their activities. Eigen is a technology suite that, they assert, can offer more than mere compliance-based document review, “[allowing] a company, law firm, or bank to gain visibility of its data core, or contract stack, to read its own DNA. . .” Dr. Lewis Z. Liu, CEO and Co-Founder, notes:

(Dr. Liu) wants Eigen to appeal to the C-suite of execs in financial service businesses and their top lawyers, who want to know the answers to key questions, such as: “What is the risk exposure of all these OTC derivatives? What are we really legally liable for according to this vast mass of futures contracts?”<sup>626</sup>

Assuming the hopes of Dr. Liu are realized, calling Eigen an expert system in the old sense may not fully capture its power. While we may be years away from an expert

---

LAWYER (2017); Clyde & Co. *See* Author, *Global Firm Clyde & Co Forms Legal Data Lab with UCL*, ARTIFICIAL LAWYER, (Dec. 12, 2017); TLT. *See* Author, *UK Law Firm, TLT, Takes Stake in U.S. Legal AI Co. LegalSifter*, ARTIFICIAL LAWYER, (Dec. 13, 2017); Holland & Knight. *See* Author, *Legal AI Co. Luminance Bags Holland & Knight in U.S. Firm Win*, ARTIFICIAL LAWYER, (Dec. 18, 2017).

<sup>621</sup> Author, *California Legal AI Co. Gavelytics Aims to Be Case Prediction Local Hero*, ARTIFICIAL LAWYER, (Nov. 17, 2017).

<sup>622</sup> Author, *Berkeley Bridge Offers DIY Expert System Platform to Lawyers*, ARTIFICIAL LAWYER (Nov. 1, 2017).

<sup>623</sup> For example, U.S. law firm Cadwalader Wickersham & Taft and Dutch law firm, Stibbe. *See id.*

<sup>624</sup> Author, *Legal AI's Dark Horse, Eigen Technologies, Comes Into The Light*, ARTIFICIAL LAWYER, (Nov. 3, 2017).

<sup>625</sup> *See id.*

<sup>626</sup> *See id.*



system that is truly able to “solve” the user’s legal problems, these new products on the market are much more complex than the text-based decision trees of the 1980s.<sup>627</sup>

## 2. Computational Models of Legal Reasoning

Where expert systems generally seek to solve complex problems in a particular domain, the newest technology seeks to extract semantic information from legal texts and use it to help humans solve legal problems.<sup>628</sup> For example, Casetext, a litigation analysis platform, now includes a legal research tool that uses natural language processing.<sup>629</sup> Casetext has launched a feature called “Black Letter Law” which enables users to “quickly identify axiomatic statements of law.”<sup>630</sup>

In addition, Casetext has launched a feature called “Holdings” which is asserted to have captured the “largest searchable collection of concise case summaries ever assembled.”<sup>631</sup> Interestingly, to create the Holdings feature, Casetext used a process they have called “judicial language processing”<sup>632</sup> which is designed to exploit “patterns within the U.S. case law corpus to excerpt summaries directly from judicial opinions.”<sup>633</sup> While case search capabilities have existed for some time,<sup>634</sup> AI-driven case law systems are relatively new. Interestingly, the Casetext team claims that both law firms and judges have expressed interest in their product.<sup>635</sup> Assuming this is true, AI case systems are already influencing judicial decision making.

---

<sup>627</sup> Eigen Technologies is not alone in using such technology. For example, Mexican legal tech firm, Laboralisto, is building a case management platform for employment disputes that will eventually provide case prediction capabilities. *See* Author, *Mexico’s Laboralisto Paves Way To Labour Law Predictive Capability*, ARTIFICIAL LAWYER, (Nov. 13, 2017) While legal AI case prediction start-up, Gavelytics, has created a predictive platforms that focuses solely on and California. Author, *California Legal AI Co. Gavelytics Aims to Be Case Prediction Local Hero*, ARTIFICIAL LAWYER, (Nov. 17, 2017).

<sup>628</sup> *See* Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 11.

<sup>629</sup> *See* Author, *Artificial Lawyer Interview: Jake Heller, CEO, Casetext*, ARTIFICIAL LAWYER, (Jan. 17, 2018).

<sup>630</sup> *Id.*

<sup>631</sup> *Id.*

<sup>632</sup> *See id.*

<sup>633</sup> *Id.*

<sup>634</sup> Consider Westlaw and LexisNexis to name but two.

<sup>635</sup> *See* Author, *Interview: Jake Heller*, *supra* note 577.

### 3. Analytics, Feedback and Rankings

One of the newer deployments of AI in the justice environment involves creating constructive feedback loops. We can see an example of this in the realm of Online Dispute Resolution specifically. The process works as follows: ODR platforms gather data as they are used, including data points on interfaces, operations and users. This data can now be channeled into feedback to create a system that better reflects the needs of the population that the system serves. For example, the British Columbia ODR system, called the Civil Resolution Tribunal, produces monthly data analytics and has a feedback system integrated into the platform.<sup>636</sup>

Feedback loops can also be important to arbitration decision makers as well as the institutions that provide their appointments. For example, outcomes could be monitored and could be used to provide specific, granular data to the decision maker, pointing out particular times when attention is lacking, or when decisions tend to be more harsh. These are biases that are well researched in the existing brick-and-mortar justice system, as described previously. In these instances feedback loops could help individual decision makers become more aware of patterns in their own decision making and the events that may be impacting their decision making process.

One of the newest systems, the “Arbitrator Intelligence Questionnaire” is a “feedback questionnaire designed to facilitate systematic collection of information about arbitrator case management and decision making.”<sup>637</sup> AIQ is designed to cover a wider range of decision makers than the CRT system, specifically covering arbitrators in the international commercial arena. Since the international commercial arbitration environment is based on party selection of arbitrators, systems such as this can be

---

<sup>636</sup> The Civil Resolution Tribunal uses various mechanisms to release information, their monthly Participant Satisfaction Survey as well as their CRT Statistics Snapshot, can both be found on their website. *See* Civil Resolution Tribunal, Blog, CRT Website, available at <https://civilresolutionbc.ca/blog/> (last visited Aug. 12, 2019).

<sup>637</sup> *See* Arbitrator Intelligence, AIQ webpage, information available at <https://www.arbitratorintelligence.org/about-us> (last visited Aug. 12, 2019).

immensely helpful in creating a transparent decision making environment, even when the process itself is shielded in confidentiality.

The AIQ describes its goal as follows:

AI's core mission is to promote fairness, transparency, accountability and diversity in arbitrator appointments. By making systematically gathered data about arbitrator decision making more equally available, the AIQ will help level the playing field in arbitrator selection and make the entire process more transparent. Meanwhile, more information about arbitrators will provide market-based accountability for arbitrators and allow newer and more diverse arbitrators a meaningful opportunity to build their reputations based on actual performance.<sup>638</sup>

While feedback loops that solicit user comment are often viewed as positive, there is research in other areas in which this type of feedback has been found to be less than helpful or accurate. In fact, several online systems that rely upon feedback and transparent comments from users have been manipulated in ways that undermined the effectiveness of the feedback loop. For example, the online marketplace Amazon has been consistently criticized for the prevalence of people gaming its review system, despite the efforts undertaken by Amazon to curtail the problem.<sup>639</sup> And, of course, the existence of gender bias in feedback-based reviews, such as job performance, is well researched as a major detriment to women.<sup>640</sup> Other biases in performance-based reviews are also well researched.<sup>641</sup> One can surely imagine these same biases occurring in justice-based performance reviews as well. Efforts ought to be undertaken to combat the replication of such identified biases in new systems.

### C. Rule of Law Considerations in Light of Ubiquitous Technology

---

<sup>638</sup> See *id.*

<sup>639</sup> See Emma Woollacott, *Amazon's Fake Review Problem Is Now Worse Than Ever, Study Suggests*, FORBES, (Sept. 9, 2017)(discussing the findings of recent surveys revealing widespread misleading, untrue and downright fake reviews.)

<sup>640</sup> See Paola Cecchi-Dimeglio, *How Gender Bias Corrupts Performance Reviews, and What to Do About It*, HARVARD BUSINESS REVIEW, (April 12, 2017).

<sup>641</sup> See Steffen Maier, *4 Unconscious Biases That Distort Performance Reviews*, ENTREPRENEUR, (Sept. 22, 2016)(discussing performance reviews style, outline, and language that distorts review outcomes).

As artificial intelligence augments or even replaces human decision making, legislators and policymakers must account for the risks that AI poses to human rights principles and the Rule of Law (RoL). As examined in Chapter 1, the Rule of Law principles demand the law to be clear, publicly visible or known, and applicable to all people, including lawmakers themselves.<sup>642</sup> The section that follows returns to the RoL principles established within Chapter 1 in an effort to extend consideration into the deployment of technology and AI in the justice eco-system, especially as it relates to the decision maker. The section begins by exploring due process within the technology-infused justice eco-system. It then moves to consider the role of society “in the loop”; issues arising when code replaces the common law of precedent; the inclusion of powerful quasi-governmental actors without accountability within the justice eco-system; and finally a review of RoL in terms of decision making processes that include a human in the loop.

### 1. Rule of Law: Due Process and Technology

The introduction of technology into a justice eco-system, especially AI deployed as part of the decision making process, will have consequences for the justice system being viewed as compliant with the RoL which may extend to issues of due process and basic human rights. For example, the Wisconsin Supreme Court, in *Wisconsin v. Eric L. Loomis*,<sup>643</sup> ruled that a judge’s use of closed-source recidivism assessment software in sentencing does not necessarily breach the constitutional right to due process,<sup>644</sup> so long as the judge does not rely on the score exclusively and receives written warnings about the software and the scoring.<sup>645</sup> The Wisconsin Supreme Court did set out limits on the use of proprietary scores in future sentencing decisions.<sup>646</sup> First, risk scores “may not be considered as the determinative factor in deciding whether the offender can be

---

<sup>642</sup> See Chapter 1(B).

<sup>643</sup> *Wisconsin v. Eric L. Loomis*, 881 N.W.2d 749 (Wis. 2016). The U.S. Supreme Court denied judicial review of the decision in June 2017.

<sup>644</sup> Loomis’s attorneys asserted it was it was unfair to rely on a score whose accuracy cannot be assessed, interrogated, and challenged and as such, their clients due process rights had been violated. See *id.*

<sup>645</sup> See *Loomis*, 881 N.W.2d at 749.

<sup>646</sup> See *id.*

supervised safely and effectively in the community.”<sup>647</sup> Second, to avoid the denial of due process, the following warning must be given to sentencing judges:

(1) the proprietary nature of COMPAS has been invoked to prevent disclosure of information relating to how factors are weighed or how risk scores are to be determined;<sup>648</sup> (2) risk assessment compares defendants to a national sample, but no cross-validation study for a Wisconsin population has yet been completed; (3) some studies of COMPAS risk assessment scores have raised questions about whether they disproportionately classify minority offenders as having a higher risk of recidivism; and (4) risk assessment tools must be constantly monitored and re-normed for accuracy due to changing populations and subpopulations.<sup>649</sup>

The Court argued that the warning would “enable courts to better assess the accuracy of the assessment and the appropriate weight to be given to the risk score.”<sup>650</sup> While these limitations are noteworthy, they likely fail to fully appreciate the impact of automation bias in technology deployments. As explained above, “automation bias,” that is the “use of automation as a heuristic replacement for vigilant information seeking and processing,”<sup>651</sup> is a well-researched concern in the socio-technological ethics. In situations such as described above, there is little doubt that automation bias effectively “turns a computerized suggestion into a final, authoritative decision.”<sup>652</sup> As such, system design must identify automation bias as a risk and consider the mitigation appropriate under the circumstances. Because this is a justice eco-system, mitigation must be appropriate in light of the significance of the impact of automation bias on the individuals within the system. Mere notice is never enough.

The model proposed in this thesis will directly take on this issue and attempt to mitigate potential negative impact by calling for a human in the decision making loop, monitoring of the user, monitoring of the outcomes, training for the user,<sup>653</sup> and

---

<sup>647</sup> *Loomis*, 881 N.W.2d 749 (Wis. 2016).

<sup>648</sup> *Id.*

<sup>649</sup> *Id.*

<sup>650</sup> *Id.*

<sup>651</sup> Danielle Keats Citron, *Technological Due Process*, U OF MARYLAND LEGAL STUDIES RESEARCH PAPER No. 2007-26; WASHINGTON UNIVERSITY LAW REVIEW, Vol. 85, pp. 1249-1313, (2007).

<sup>652</sup> *Id.*

<sup>653</sup> Studies have shown that individuals who receive such training are more likely to scrutinize an automated system's suggestions. *See id.*

auditing for patterns that indicate a failure to use the technology correctly,<sup>654</sup> such as attention lapses and unreasonably quick decision making.

## 2. Rule of Law: Consensus and Society-in-the-Loop

As highlighted by the Law Society of England and Wales in its report entitled: *The Use Of Algorithms In The Justice System In England And Wales* (as summarized by the Law Gazette) “We need to build a *consensus rooted in the rule of law*, which preserves human rights and equality, to deliver a trusted and reliable justice system now and for the future.”<sup>655</sup> Consensus rooted in the rule of law demands examination of several aspects of technology as a decision maker in a justice-based eco-system.

First, one must imagine current technology design as primarily designed *without* consensus. In fact most technology, especially AI, is designed without consultation of interested parties, or local communities and constituents. The Rule of Law demands attention to the requirements of constitutional democracy<sup>656</sup> which is “democracy of, by, and for the people, such that all citizens, *rather than favored individuals or groups*, have the right to politically participate.”<sup>657</sup> One can argue that at the current time, designers and technology industry financiers are a narrow group wielding broad and powerful influence over technology design and deployment decisions. Moreover, those same entities are able to hide behind intellectual property rights to obscure important portions of the design. Thus, although machine learning systems adapt and hopefully improve over time, the original influences are difficult to determine and their effects may be longer lasting than society anticipates.

---

<sup>654</sup> See Danielle Keats Citron, *Big Data Should Be Regulated by ‘Technological Due Process’*, NEW YORK TIMES (July 2016).

<sup>655</sup> Michael Cross, *Justice Algorithms ‘Need Urgent Oversight’*, LAW SOCIETY GAZETTE (June 4, 2019) (speaking about The Law Society, *Algorithm Use In The Criminal Justice System*, June (2019)).

<sup>656</sup> When decision systems are introduced into public contexts such as criminal justice, it is important they are subject to the scrutiny expected in a democratic society. See Law Society, *Algorithm Use In The Criminal Justice System*, (June 2019) at Section 5.4.3.2.

<sup>657</sup> *Id.*

Similarly, to date there have been very few consensus-building events around the deployment of algorithmic systems. There is little to no consultation of the panoply of stakeholders impacted or interested in the deployments. Fortunately, some academics and others are beginning to focus attention on the need to be more inclusive in designing and deploying new technology. For example, MIT Media Lab author Rahwan has put forward the idea of “society-in-the-loop artificial intelligence,” which embeds the judgment of society as a whole in the development and use of AI systems,<sup>658</sup> and the Law Society, Technology and the Law Policy Commission makes similar calls for action.<sup>659</sup>

As described in Chapter 3, ODR can be designed by a multitude of entities and can be deployed in either a public or private provider environment. Public deployments of ODR exist within the existing justice eco-system and as such are required to be compliant with the same standards as brick-and-mortar courts. In contrast, private providers are outside the traditional brick-and-mortar system and are governed by contract and consent to arbitration, bound by fewer standards and protections. Society-in-the-Loop is a concept that becomes perplexing to apply to the ODR context the more that design and deployment is privately sourced. Citizens have rights and the ability to participate in the election process, which gives them some voice in deployments managed by publicly elected officials, whereas private providers are not required to be concerned with citizens as such, and are expected to focus their consideration on more narrow client groups. In these situations it is important that government undertake the role of protecting citizens in the face of consent-driven contractual agreements, holding those that design and deploy, even from the private sector, to a standard that includes society in the loop.

### 3. Rule of Law: Code Replaces the Common Law of Precedent

---

<sup>658</sup> See Iyad Rahwan, *Society-in-the-Loop Programming the Algorithmic Social Contract*, MIT Media Law (Aug. 12, 2016).

<sup>659</sup> See Law Society, *supra* note 603 at page 23 citing Veale M, Van Kleek M, and Binns R, ‘Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making’ in Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI 2018 , 7-8, (ACM Press 2018).

Within the Law Society Report authors noted a particularly relevant observation made by Barrister Jacob Turner:

It is of course true that morals shift and ethics change but laws can be updated to fit new morals. When it became socially acceptable in wider with regards to ethics for AI. With regards to the technical implementation of these, this is why we need to set-up the governance structures, including and involving experts who are able to track the changing laws and the changing rules and apply them with regards to the technology as it stands, at any given point, but this is an ongoing, dynamic process.<sup>660</sup>

As we have discussed machine learning is dynamic, but its initial algorithms, especially those designed with prediction as the design focus, are built upon past data. Prediction algorithms:

run the risk of stagnation, holding the evolution of justice anchored in the past rather than free to evolve. In computing, this is known as “concept drift”, and is a challenge to understand and cope with in high-stakes environments.<sup>661</sup>

Concept drift is a risk that needs to be assessed and addressed in the mitigation plan. It may be the case, however, that concept drift presents yet another reason to be skeptical of the deployment of algorithms in the justice eco-system. Certainly more research must be done in this area.

#### 4. Rule of Law: Powerful Quasi-Governmental Actors without Accountability

Following on from the need to include society in the loop, a commitment to democracy and citizen engagement demands a mechanism of accountability. On July 2, 2019 the Technology and the Law Policy Commission found “a lack of explicit standards, best practice, and openness or transparency about the use of algorithmic systems in criminal justice across England and Wales.”<sup>662</sup> While these might seem to be considerations that lie outside the scope of AI standards, they are important as these systems are being used by governmental and quasi-governmental authorities, and accountability is a cornerstone of any justice eco-system participant. The absence of standards is a global issue, as will be discussed more fully in Chapter 5.

---

<sup>660</sup> Law Society, *supra* note 603 at Section 5.4.4.2 page 23.

<sup>661</sup> *Id.*

<sup>662</sup> Law Society, *supra* note 603 at Summary page 4.



Fortunately, a few entities are taking on the issue and some consensus on basic necessary standards is beginning to emerge. For example, the National Institute of Standards and Technology has published a report stating that “stakeholders in the development of this plan expressed broad agreement that societal and ethical considerations must factor into AI standards, . . .”<sup>663</sup> While the report notes that it is not “clear how that should be done”<sup>664</sup> nor is there yet “sufficient scientific and technical basis to develop those standards,”<sup>665</sup> it does note two areas of consensus amongst those stakeholders surveyed:

- (1) The degree to which ethical considerations might be incorporated into standards should be tied tightly to the *degree of risk to humans*, and
- (2) Privacy considerations should be included in any standards governing the collection, processing, sharing, storage, and disposal of personal information.<sup>666</sup>

The proposal that risk to humans be used as the measure of requisite ethical scrutiny, of course, calls for both risk and scrutiny to be further defined. As the NIST Call for Public Comment proceeds, its commitment to consideration of societal and ethical standards as part of technological design governance is a major step in the right direction. More needs to be done: unfortunately the report fails in sketching these broad commitments to even hint at considerations of responsibility, an issue the author hopes to see further clarified.

Responsibility and liability should be established for those that *build and deploy* technology as justice eco-system providers. As prescribed by The Law Society report “duties and statutory oversight and scrutiny bodies . . . would provide key safeguards to the integrity of criminal justice in the digital age.”<sup>667</sup> Merely allowing entities to seek recourse for failure to provide a deliverable under the build contract terms is not enough. Instead, individuals and groups must be given a right of recourse when they

---

<sup>663</sup> National Institute of Standards and Technology (NIST), *U.S. Leadership In AI: A Plan For Federal Engagement In Developing Technical Standards And Related Tools*, DRAFT FOR PUBLIC COMMENT, line 265 (July 2, 2019).

<sup>664</sup> *Id.* at line 266.

<sup>665</sup> *Id.* at line 267.

<sup>666</sup> *Id.* at line 268-71.

<sup>667</sup> Law Society, *supra* note 603 at Summary page 5.

have been injured or negatively impacted by the deployment of AI. The recourse cannot be based on the narrow traditional demonstrations of an economic or discriminatory harm, because AI systems are typically opaque and often fail to demonstrate recognizable impact upon individuals and groups until long after the deployment occurs. Thus, legal recourse must be available based on a new and expanded understanding of harm. Both accountability and clear chains of responsibility demand transparency in AI design, deployment and monitoring, including clear explanations for how AI systems may impact *all* stakeholders.

The Canadian government in February of 2019 passed a federal directive entitled *Directive on Automated Decision-Making*<sup>668</sup> which puts in place a tool that can “accurately measure the impact an automated process may have on the citizens it was created to serve.”<sup>669</sup> The Algorithmic Impact Assessment requires companies to complete an online survey about the platform they have developed, answering questions such as “Does the system enable override of human decisions?” and “Is there a process in place to document how data quality issues were resolved during the design process?” The company then receives an impact level rating. The impact level of the AI platform is ranked one through four, with higher ranking demanding a high level of mitigation. For example,

If an automated decision process receives an impact assessment of level four, it will require two independent peer reviews, a public plain language notice, a human intervention failsafe, and re-occurring training courses for the system.<sup>670</sup>

It is believed this is the first of its kind<sup>671</sup> tool which provides a high level of accountability for those who *deploy* AI systems. Although limited to those entities that are contracted by/for the government, it is a first step toward preventing entities who

---

<sup>668</sup> See Government of Canada, *Directive on Automated Decision-Making*, (effective April 1, 2019).

<sup>669</sup> Max Greenwood, *Canada’s New Federal Directive Makes Ethical AI a National Issue*, TECH VIBES (March 9, 2019).

<sup>670</sup> *Id.*

<sup>671</sup> Before its full adoption, it was in use by Mexico and Portugal and the U.K. may seek to as implement the tool. See Greenwood, *Canada’s New Federal Directive*, *supra* note 616.

create technology with deep impacts to escape accountability behind a veil of intellectual property rights and design and dash creations.

The model that in Chapter 5 takes a very similar approach to accountability, asking those that design and those that deploy technology to be held responsible for their creations within the justice eco-system. Moreover, for any deployment occurring in the justice eco-system, the author assumes significant potential impact (a four in the parlance of the Algorithmic Impact Assessment) and as such, the author has incorporated the implementations recommended, including training, meaningful human intervention and transparency. Peer review has not been recommended, *per se*, but auditing throughout the process has been recommended in order to tackle the same issue.

It should be noted, the author does not agree with the government taking on the role of impact assessment designer. This is primarily because tools such as this may allow entities to respond to the questions, receive an impact score with the corresponding mitigations required, and then shield itself from negative impacts that are yet to be captured in the scoring. Gaps in scoring adjustments seem inevitable. Moreover, the government as the impact assessment designer will likely move scoring and mitigations into a political and not a constituent based conversation. Additionally, one can imagine entities designing platforms based on achieving a favorable score, without real focus on the intent of the impact assessment. Thus, the author takes the position that the development of industry driven best practices, and robust auditing by both government and industry, is the best design practice with ultimate accountability residing with those that *design and deploy* the technology.

##### 5. Rule of Law: Human-in-the-Decision-Making-Loop

The author reviewed and struggled mightily efforts to embrace a “robot judge,”<sup>672</sup> that is, a purely AI-based decision maker, and ultimately rejected this possibility in favor of a hard human-in-the-loop requirement. As Steven’s articulates:

---

<sup>672</sup> The process of consideration has been long. See Raymond, *Technology, Ethics And Access To Justice*, *supra* note 58.

Legislators and policymakers should carefully but swiftly define “artificially intelligent” technology in order to account for any human rights abuses facilitated by such technologies. This is because AI systems can challenge human agency either by standing in for us or nudging us, as well as ultimately disrupt the rule of law by forming a part of a given regulatory framework—but with such opacity that the logic of the engineer or machine cannot be challenged.<sup>673</sup>

The model proposed by this thesis recognizes these limitations, seeks to create a framework and model that is built to identify, reduce, and mitigate against negative impacts upon humans. Ultimately the current limitations of AI research and technology and the lack of confidence in those who seek to build and deploy the technology serve as a significant barrier to deploying technology without a human in the loop. The author hopes that, one day, those who design and research AI can demonstrate that this lack of trust is misplaced. For today, however, the potential negative consequences are too great to embrace a justice environment without human decision making.

This does not need to be considered a defeat or cause to reject technology within the justice eco-system altogether; instead it should be viewed as an opportunity to use technology to improve the justice environment. As was explored in Chapter 3, ODR is evolving — some of the best in the field are taking justice directly into underserved communities, are improving access to justice, and are beginning to conduct research to further advance the field. None of these systems embrace the absence of a human, but these ODR systems do deploy technology in ever expanding ways. And, prediction of justice outcomes is a growing area of research. For example, in 2016 researchers at University College London and the universities of Sheffield and Pennsylvania were able to design AI which correctly predicted verdicts to an accuracy of 79%, from hundreds of cases heard at the European Court of Human Rights.<sup>674</sup> Of course, as described earlier in this chapter, prediction systems replacing human decision makers is likely a ways off, but technology that notices patterns and assists in screening and decision making is growing in use across the global justice eco-system.

---

<sup>673</sup> Yuan Stevens, *The Promises and Perils of Artificial Intelligence: Why Human Rights and the Rule of Law Matter*, MEDIUM (Sept. 5, 2017).

<sup>674</sup> See Jane Wakefield, *AI Predicts Outcome Of Human Rights Cases*, BBC, (Oct. 23, 2016).

## D. Regulations of Artificial Intelligence

Although regulation of artificial intelligence is a very new endeavor, the United Kingdom,<sup>675</sup> France,<sup>676</sup> Australia,<sup>677</sup> and the United States<sup>678</sup> all have recently drafted or passed legislation to hold tech companies accountable for their algorithms. It should be noted that in many instances the regulation is narrowly crafted mainly to make tech companies responsible for monitoring content for violent or anti-Semitic speech. However, the United States has taken an arguably more broad regulatory approach in the new bill, called the Algorithmic Accountability Act,<sup>679</sup> which would require tech companies to audit not only content but their machine-learning systems for bias and discrimination and to take corrective action in a timely manner if such issues were identified.<sup>680</sup> The Bill is in direct response to several high-profile news stories from 2018 and '19 that have revealed the far-reaching damage algorithmic bias can do in various contexts.<sup>681</sup> This Bill and several more that are likely to be introduced in the coming months in the U.S. are intentionally expansive in order to encompass different AI products and data processes across a variety of domains and to account for the migratory nature of technologies that are often used for drastically different purposes across multiple industries. As such, the full impact of these new regulations will be difficult to gauge in the short term.<sup>682</sup> The U.S. Congress is trying to rework the traditional regulatory framework to accommodate this new reality<sup>683</sup> and this will undoubtedly lead to various AI-based systems coming under new regulation. The

---

<sup>675</sup> See James Ball, *The UK's Online Laws Could Be The Future Of The Internet—And That's Got People Worried*, MIT TECHNOLOGY REVIEW, (April 9, 2019).

<sup>676</sup> See, Lester Feder and Zorro Maplestone, *France Has Recruited Facebook To Help Solve Its Anti-Semitism Problem*, BUZZFEED NEWS, (March 11, 2019).

<sup>677</sup> See Author, *Australia's New Law Threatens Jail If Social-Media Firms Don't Remove Violent Content*, The Download, MIT TECHNOLOGY REVIEW, (April 4, 2019).

<sup>678</sup> See U.S. Senate, *Algorithmic Accountability Act*, 116 Congress, Session 1, (2019).

<sup>679</sup> See *id.*

<sup>680</sup> See *id.* The Act would also require those companies to audit all processes beyond machine learning involving sensitive data for privacy and security risks. Interestingly, the bill would place regulatory power in the hands of the U.S. Federal Trade Commission which is the agency in charge of consumer protections and antitrust regulation in the United States. See *id.*

<sup>681</sup> Staff Author, *Algorithmic Accountability Act, The Algorithm*, MIT TECH REVIEW, (April 12, 2019) .

<sup>682</sup> See *id.*

<sup>683</sup> See *id.*

newest regulation across the globe is sending a clear message to tech designers: you will no longer be allowed to unleash technology-driven systems on the world without monitoring, adjusting and using impact assessments as part of your design.

In addition to regulation, many organizations have launched initiatives to establish ethical principles focusing on socially beneficial AI. The proliferation of these efforts is so great that research is now being undertaken to consider if they can be synthesized into a unified framework. In July of 2019, Floridi and Cowls argued such a unified framework has in fact emerged.<sup>684</sup> In their article *A Unified Framework of Five Principles for AI in Society*,<sup>685</sup> they identify these principles as: Beneficence: promoting well-being, preserving dignity, and sustaining the planet;<sup>686</sup> Non-maleficence: privacy, security and “capability caution”;<sup>687</sup> Autonomy: the power to decide;<sup>688</sup> Justice: promoting prosperity, preserving solidarity, avoiding unfairness;<sup>689</sup> Explicability: enabling the other principles through intelligibility and accountability.<sup>690</sup>

Of interest in the context of the thesis are several important takeaways from the framework. First, the Floridi and Cowls framework introduces a “decide-to-delegate” model, recommending that:

humans should retain the power to decide which decisions to take: exercising the freedom to choose where necessary, and ceding it in cases

---

<sup>684</sup> They developed the framework through an examination of well respected, broad stakeholder engagement, efforts crafted within the last three years. They are: Future of Life Institute, *Asilomar AI Principles*, (2017); The Montreal Declaration for Responsible AI; *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, IEEE’, IEEE, p6 (2017); European Commission’s European Group on Ethics in Science and New Technologies, *Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems*, pp. 16-20, (March 2018); UK House of Lords Artificial Intelligence Committee’s report, *AI in the UK: Ready, Willing And Able?*, §417; *The Tenets of the Partnership on AI*, (2018).

<sup>685</sup> Luciano Floridi and Josh Cowls, *A Unified Framework of Five Principles for AI in Society*, HARVARD DATA SCIENCE REVIEW, Vol. 1, (2019).

<sup>686</sup> *Id.* at 3.1.

<sup>687</sup> *Id.* at 3.2.

<sup>688</sup> *Id.* at 3.3.

<sup>689</sup> *Id.* at 3.4.

<sup>690</sup> Floridi, *A Unified Framework*, *supra* note 632, at 3.5.

where overriding reasons, such as efficacy, may outweigh the loss of control over decision-making.<sup>691</sup>

This is also in line with the “society in the loop” principle where society should be an integral part of the decision of when to use AI and what norms and standards should be captured within the deployment.

Second, as explained in the Montreal Declaration, “the development of AI should promote justice and seek to eliminate all types of discrimination.”<sup>692</sup> This is both a noble goal and a necessary minimum standard. A guiding principle of this thesis is that technology deployments should be done in a manner that improves society. In this author’s opinion too much AI and technology seeks to merely replicate prior discriminatory practices, cloaking them behind a veil of “accurate prediction.” We must demand better of technology, especially AI.

Finally, as Floridi and Cowls note when commenting on the principle of explicability:

“explicability,” incorporating both the epistemological sense of “intelligibility” (as an answer to the question “how does it work?”) and in the ethical sense of “accountability” (as an answer to the question “who is responsible for the way it works?”), is the crucial missing piece of the AI ethics jigsaw.

Incorporating both intelligibility and accountability is essential in the creation of an alternative justice environment that will rely heavily upon technology, especially AI. As will be discussed in Chapter 5, it is an essential consideration within the model as the trustworthiness of an alternative to justice must be essential and it is an expectation under the RoL.

## E. Law and AI Chapter Conclusions

As we conclude this chapter, having reviewed their many possible fallacies, it is incredibly important to note that predictive AI systems can be remarkably accurate compared to human predictors. For example, in October of 2017 U.K.-based legal tech

---

<sup>691</sup> *Id.* at 3.3.

<sup>692</sup> Floridi, *A Unified Framework*, *supra* note 632, at 3.4 *citing* The Montreal Declaration for Responsible AI; *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*, IEEE, p.6, (2017).

start-up CaseCrunch challenged lawyers to see who could predict with greater accuracy the outcome of a number of financial product claims.<sup>693</sup> The outcome:

CaseCrunch's predictive algorithms and modelling of legal issues came out on top, scoring almost 87% accuracy in terms of predicting the success or failure of a claim. The English lawyers who were beaten got overall an accuracy level of around 62%.<sup>694</sup>

That is not to write that all predictive AI systems are more accurate than their human counterparts. As we briefly explored above, several of the more commonly used recidivism prediction tools are actually poor predictors of recidivism.<sup>695</sup> It is much more accurate to say that in pursuing the potential value of prediction, we must take note both of the potential of AI and of the cautionary tales that provide a realistic picture of the capabilities, limitations and sociotechnical context of technology.

The author's key argument is that adequately complex models based on appropriate framing, and ongoing rigorous evaluations of impact must be made essential aspects of all design. Without extreme care, the risks of algorithmic discrimination are high, and the potential consequences are grave. Regardless of where bias, dependency or partiality arise, their impact will become more and more insidious in a technology-driven environment, and could lead to further loss of trust in our judiciary. As Chulu asserts:

Living in a democratic society, we expect public institutions to be fair, transparent, and accountable. We expect decisions that affect our lives to be bound by the rule of law, making it possible to demand explanation, and appeal and redress of those that are based on faulty premises.<sup>696</sup>

Rule of law and trusted judicial agency are essential to any "alternative" to the justice system. There simply can be no legitimate argument against such a commitment. The outstanding question is what regulation can best ensure a commitment to rule of law in these alternative systems. The next chapter seeks to flesh out those aspects of the

---

<sup>693</sup> See Author, *AI Beats Human Lawyers in CaseCrunch Prediction Showdown + Data Updates*, ARTIFICIAL LAWYER, (Oct. 28, 2017).

<sup>694</sup> *Id.*

<sup>695</sup> See Ed Yong, *A Popular Algorithm Is No Better at Predicting Crimes Than Random People*, THE ATLANTIC, (Jan. 17, 2018).

<sup>696</sup> Chulu, *Let Us End Algorithmic Discrimination*, *supra* note 495.



system, in our specific area of focus: partiality, dependence and bias in the decision maker.

## Chapter 5: Design Guidance for the Reduction of Partiality, Dependence, and Bias in an Online Justice Decision Maker

This chapter will combine all of the essential topics so far explored into recommendations for a complex model of decision making within an online justice environment. As the thesis is primarily concerned with the decision maker, the focus has been on independence, impartiality and absence of bias within the decision making process, and specifically in terms of the decision maker. This chapter brings those discrete topics together to propose a single, forward-looking model for decision making in ODR.

The use of a model, any model, can be fraught with risk, even if done correctly. And in the case of justice environments, the risks are significant: individuals and/or business may lose large sums of money, personal liberty, or suffer significant long-term negative consequences; furthermore, a secondary effect of underappreciated risk is that it can lead to society to lose trust in the justice system. Thus, managing model risk will form the backbone of this chapter's analysis and recommendations.

The chapter will begin by identifying, explaining and supporting eight main *rules of action* that must guide all of the design choices that go into the complex decision making model. The chapter will then take key terminology, such as *decision maker* and *automation*, which have been used in context throughout the thesis, and provide them with specific operative definitions in the context of the model proposed. Having established both the basic principles and technical components that will constitute it, the thesis will then present a unified model of online justice that focuses on providing assurance of compliance with the Rule of Law. Finally, the chapter concludes with some additional items of design guidance aimed at addressing concerns that may arise relating to aspects of the model other than our primary focus, the decision making process.

**\*Modeling for this Project: Limitations and Necessary Adjustments**

As discussed in Chapter 4, modeling – that is the process of breaking down a complex task into discrete steps has limitations. This is especially true when one of the complex tasks is decision making. In crafting the model, below, the model builder is guided by a set of specific rules of action. Yet, in creating these rules of action, especially in a setting such as this, the model building is significantly limited in terms of perspective, experience and personal bias. For example, it is important to recall, as described in Chapter 4 social complexity is a major issue in model building and this is especially true in justice environment. The model below, is limited due to the nature of the thesis, and therefore likely suffers at the hands the fairness trap.<sup>697</sup> In this situation, the model is focused on a specific narrow aspect of a much larger system. And while this may, in general, be appropriate in some settings, in a complex justice environment it is a potential issue. To address this issue the model designer has attempted to be clear that this model does not intend to capture the social complexity of the entire justice environment. Nor, does the model attempt to capture all decision making. Instead, the model builds upon – and specifically uses- existing rules, parameters, and bounds to replicate existing systems. The choice to replicate, with adjustments and additions, of an existing system, is however, also a bias as it presupposes the existing systems are worthy of replication.<sup>698</sup>

However, the model also embraces many of the aspects of social complexity as model limitations that must be considered and reflected in the design. For example, the solutionism trap, that is where technology is assumed to be the best solution,<sup>699</sup> is specifically rejected as the model uses technology in well- defined areas and rejects the use of technology when the technology fails to full capture social and legal complexity. In fact, the model inserts a human when complex decision making is needed to resolve a dispute. As explored in Chapter 4, current algorithm driven systems, even expert

---

<sup>697</sup> See Chapter 4, *supra* at A(3)(b)(i).

<sup>698</sup> Note, this is not the portability trap, where an algorithm is repurposed.

<sup>699</sup> See Chapter 4, *supra* at A(3)(b)(v).

systems,<sup>700</sup> have yet to fully capture the complexity of complex, nuanced, legal prediction.<sup>701</sup> Currently, there is no outcome predictive system that can model legal reasoning in complex legal decision making that is widely released and verified as accurate.<sup>702</sup> Moreover, the data these systems are built upon is often less than robust, complete and/or bias free.<sup>703</sup> As such, the model sets out specific Rules of Action that set limits and provide guidance to the model design. For example, Rule of Action 6, proscribes a human in the decision-making loop.<sup>704</sup> While, the definitions create limits, such as the definition of a dispute (the limits to when a human will be sued) and Chapter 5, section 3(b) adjusts the decision making model to account for the risk that arises for bringing a human into the decision making loop.<sup>705</sup>

Finally, the Ripple Effect, that is the failure to understand how the insertion of technology into a system will change the system<sup>706</sup> has been accommodated in the model, however, as an early model it is clear more research will need to be done to fully identify how the deployment of technology, in the way proscribed in Chapter 5 will impact the complex justice environment. Within the current model, the human decision maker insertion and the impact of the technology will hopefully be discovered through robust auditing and by the reduction of conflict in the appointment process. First, the system, of both appointments, human interaction with the system, and outcomes are all part of monitoring and auditing.<sup>707</sup> The model is designed to discover issues and to mitigate against these issues through the use of an institution of authority.<sup>708</sup> The authority has the ability to take corrective action and to provide feedback, both to design

---

<sup>700</sup> See Chapter 4, *supra* at B(1-3).

<sup>701</sup> See Chapter 4, *supra* at A(3)(b)(v).

<sup>702</sup> See Chapter 4, *supra* at B(2).

<sup>703</sup> See Chapter 4, *supra* at A(3).

<sup>704</sup> See Chapter 5, *infra* at A(6).

<sup>705</sup> See Chapter 5, *infra* at C(3)(b)(i).

<sup>706</sup> See Chapter 4, *supra* at A(3)(b)(iv).

<sup>707</sup> See Chapter 5, *infra* at C(3)(1)(3).

<sup>708</sup> See Chapter 5, *infra* at C(3)(b)(1).

and to the human in the loop.<sup>709</sup> In addition, the impact of the technology deployment in the entirety of the system, including the impact upon the users of the system and those who receive outcomes, is also discovered and monitored, through the use of ongoing impact assessment<sup>710</sup> and user risk arising from the human interacting with technology itself.<sup>711</sup> Impact assessment are specifically designed tools that are often used in technology environments to note and adjust systems to mitigate negative impacts,<sup>712</sup> such as unintended consequences that might arise from unintended disclosures or similar events.

Undoubtedly, there is one key point of potential risk that arises from model design such as this, the risk that system is embedded into a system and the model fantastically fails. Unfortunately, other than recognizing such risk exists, and insisting that robust research, testing, and monitoring of the system occurs, this risk is a reality of technology design. The entire system is designed with best practices in mind, monitoring,<sup>713</sup> auditing,<sup>714</sup> and outcome driven surveys.<sup>715</sup> But systems fail and one must concede the only way to prevent widescale negative impacts is to catch those negative impacts as early and as quickly as possible through a well-designed process of monitoring and mitigating to reduce negative impacts.<sup>716</sup>

## A. Rules of Action

Rules of Action are the rules that govern a design undertaking. Rules are taken as true, until proven otherwise.<sup>717</sup> This section will set out the eight main rules of action the

---

<sup>709</sup> *See id.*

<sup>710</sup> *See* Chapter 5, *infra* at C(4).

<sup>711</sup> *See* Chapter 5, *infra* at C(3)(iii).

<sup>712</sup> *See id.*

<sup>713</sup> *See* Chapter 5, *infra* at A(5).

<sup>714</sup> *See id.*

<sup>715</sup> *See id.*

<sup>716</sup> *See* Chapter 5, *infra* at B(6).

<sup>717</sup> *See supra* Chapter 4, Section A(1).

author proposes ought to govern the design of a decision making model in a justice environment.

Both these rules and the resulting model are designed with the following goals in mind. (1) Design Objective: To produce a machine learning-based model that assists individuals in making decisions relating to the resolution of identified legal conflicts. (2) Business Use: To design a model that can be deployed in both public and private alternative justice environments as an *assistant* to the decision maker. (3) Thesis Focus: To ensure a model in which decision makers are both impartial and independent, without exception, even if adherence to this rule of law principle conflicts with the efficiencies of the overall ODR system. (4) Risk Management: to mitigate risks associated with potential dependence, partiality and bias. Based upon these objectives, the following sub sections set out the rules of action, both expanding upon topics and resolving concerns raised in prior chapters.

1. The Rule of Law (RoL) MUST be protected within the online justice environment.

As previously discussed in Chapter 1, the RoL embodies several protections that are considered essential within any justice environment. The absence of RoL principles implies, *prima facie*, negative impacts upon the people inside the justice environment. Moreover, when the RoL is absent, the trust of the users of the justice environment will erode.<sup>718</sup> Thus, ANY justice environment that fails to fulfill the expectation of adherence to the RoL is problematic to the institution of justice as a whole.

In this specific case, it is important to recall several aspects of the thesis. First, this ODR system being designed is envisioning a private dispute resolution mechanism, with arbitration as a final stage of the process. The inclusion of such this specific type of ODR system sets the stage for several key aspects to be carried through the rules and the model. With private arbitration as the key component, voluntary, consent of the parties, and choice between providers and institutions is key to considerations of authority, design, and monitoring. As described in Chapter 2, arbitration and arbitration

---

<sup>718</sup> See *supra* Chapter 1, Section A, C, and D.

institutions are considered, by the global legal community as appropriate dispute resolution providers and as appropriate authorities of such services. The providers, under the authority existing within the arbitration community and supported by the global legal community as part of the very fabric of ODR. It is these institutions tasked with maintaining and protecting the trust placed in them.

## 2. An Independent And Impartial Decision Maker Is An Essential RoL Principle, Such That No Departure From The Principle Should Be Allowed

As demonstrated in Chapter 1, the RoL is a highly regarded, widely recognized set of principles which embody the expectations of what a justice system should entail. The absence or failure of a single principle may be, and often is, sufficient to raise doubts about the entire system's adherence to the RoL.<sup>719</sup> One of the essential attributes of the RoL is that judicial decisions are made within an independent judiciary by impartial and independent decision makers.<sup>720</sup> The absence of this attribute constitutes a systemic failure of RoL. And, while there are various important concerns in the world of justice that demand to be balanced against each other, these fundamental principles are so significant that there can be no balancing of interests. It is not possible to balance a quicker or less costly process against an independent and impartial tribunal; the absence of an impartial tribunal not only negatively impacts but categorically undermines the entire endeavor. An impartial and independent decision maker is a non-negotiable aspect of any justice environment, even a private online one.

## 3. It Is The Method Of Ensuring The Independence And Impartiality Of The Decision Maker That Is Open For Greater Debate

The exact method of successfully providing the various attributes of the RoL is, however, a matter of legitimate debate. For example, while an independent judiciary and an independent and impartial decision maker are essential attributes, the features that make up the standard measure for independence and impartiality are less than universal. Prior chapters, such as Chapters 2 and 3, have attempted to explore and find

---

<sup>719</sup> See *supra* Chapter 1, Section B.

<sup>720</sup> See *supra* Chapter 1, Section C, and D.

commonality amongst the various systems.<sup>721</sup> And while some commonality does exist, grey areas remain. Thus, the model must work to ensure the mitigation of risk associated with those gray areas.

Returning to Chapter 3's presentation of data on actual challenges made to the impartiality and independence of arbitrators,<sup>722</sup> the following summary graph is revealing:

<b>Reason for Challenge</b>	<b>Count</b>
Ties to Legal Practice	17
Prior Arbitration Work	5
Other Professional Ties	9
Social Ties	6
Multiple Arbitrators	5

With the exception of multiple arbitrators, the significant majority of these cases turned on one party believing the decision maker's impartiality to be compromised by the presence of a prior or current relationship. As such, current cases and existing arbitration rules demonstrate that parties expect the justice environment to eliminate dependent relationships. These cases *suggest* that partiality is presumed if dependence exists.

However, in Chapter 1, the author explained and supported the assertion that independence and impartiality are deserving of separate guarantee.<sup>723</sup> As such, a system ought to go beyond current expectations and be designed to ensure *both* the independence and impartiality of the decision maker, as both are required to be compliant with the RoL. Moreover, as arbitral institutions are an essential part of the ODR design envisioned later in the Chapter, it is these institutions that are tasked with ensuring compliance with these expectations. And as arbitration institutions, their outcomes are part of the check and balance system of limited review of outcomes

---

<sup>721</sup> See *supra* Chapter 2, Section A, and B; Chapter 3, Section B.

<sup>722</sup> See *supra* Chapter 2, Section D.

<sup>723</sup> See *supra* Chapter 1, Section A.

afforded all awards at the time of enforcement.<sup>724</sup> As such, local courts serve to ensure arbitration institutions meet the appropriate standard.<sup>725</sup>

Later sections will begin to frame this issue in practical terms. In general, it is this author's contention that AI-driven online systems will and should use data analytics to monitor and improve all aspects of the system, including the behavior of the human decision maker. As such, technology can and should be used to reduce partiality in a number of ways.

Conflicting out decision makers based upon prior relationships could be a reasonable proxy for reduction of dependency, and is a task that can be greatly abetted by use of AI.<sup>726</sup> Moreover, auditing of the parties and the decision makers use of the system, including written communications, and auditing of outcomes are operations that technology can perform with great positive impact. If, for example, an algorithm can identify patterns that were otherwise indiscernible to the decision maker or their governing institution, this information from the algorithm presents the decision maker with an opportunity to become aware of potential issues. At this point they may ask themselves whether those patterns legitimately reflect the circumstances of the cases they are assigned, or if they reveal potential partiality issues. This would create a constructive feedback loop between algorithmic and human intelligence.

#### 4. It is the Model Design that Mitigates Dependence and Bias

As explained in Chapter 4, bias can emerge within algorithms and their outputs as well as within humans. Algorithmic bias originates, (1) from biases pre-existing in the training data,<sup>727</sup> (2) from limitations built into the processes and terms of the algorithm itself,<sup>728</sup> and (3) as emergent biases that result from human interactions with the

---

<sup>724</sup> See *supra* Chapter 2 (A).

<sup>725</sup> See *supra* Chapter 2 (A).

<sup>726</sup> A task that also should fall to the responsibility of the appointing institution as it is the arbitration institution that registers individuals to become part of the roster of arbitrators and it is the institution that creates conflict rules, enforces conflict rules and ultimately determines if a conflict exists.

<sup>727</sup> See *supra* Chapter 4, Section A(3).

<sup>728</sup> See *supra* Chapter 4, Section A(3).



technology.<sup>729</sup> Model design must consider all three of these areas to reduce bias in the entirety of the system. In our specific instance, that includes the reduction of bias in the entire system as well as reduction of dependence and partiality in the human decision maker. The following sections will specifically speak to bias in the non-human aspects of the decision making process, while also seeking to reduce dependence and partiality in the human decision maker. For example, dependence will be reduced through requiring human decision makers to comply with a carefully crafted arbitration institution conflict declaration protocol and through a random selection process. Partiality will be reduced by arbitration institutions programmatically monitoring the outcomes and deploying machine learning to discover patterns. In addition, the author seeks to address bias in the process of decision making by embracing a strong preference for human-centered design, which demands the system itself be monitored for emergent bias by the appointing institution. For example, the arbitration institution technology will be able to present multiple models of prediction and suggestions that allow the human decision maker to select from various potential outcomes. The pattern of final selections will also be monitored by the arbitration institution technology, so the system can react to emergent bias within the human portion of the model. Ensuring these steps are enshrined within the model from the beginning allows auditors to discover emergent issues and ensures the biases are addressed from inception, thereby ensuring a commitment to the reduction of partiality and dependence, an essential commitment to the RoL.

##### 5. The Reduction of Negatively Impactful Bias is the Current “Best Practice” Standard

While dependence can likely be greatly reduced through a well-designed nomination, selection and conflict process, bias and impartiality present greater difficulty. As such, special focus must be given to impartiality and bias beyond the measures listed in Rule of Action 4.

---

<sup>729</sup> See *supra* Chapter 4, Section A(3).

Impartiality, as defined in Chapter 1, is a “state of mind or attitude of the tribunal in relation to the issues and the parties in a particular case.”<sup>730</sup> As Mill explains, it is an obligation of justice and an essential concept with the RoL to ensure that the decision maker is “exclusively influenced by the considerations which it is supposed ought to influence the particular case at hand.”<sup>731</sup> One can thus argue that the state of mind of the decision maker is an important consideration. Yet, the state of mind is obviously a very difficult attribute to measure; in fact it is likely impossible to measure, as implicit bias and hidden agendas are rarely admitted. Transparency and technology may provide a way to reveal some of the decision maker’s state of mind through monitoring behavior, as described above.<sup>732</sup>

The question quickly becomes, however, not merely how to measure but exactly what to allow and what to prohibit. For example, it is easy to argue that outcomes of all decisions must be transparent and to achieve this objective by publishing outcomes. It is however, more difficult to determine what will be considered bias or partiality. The author contends that the standard must be no “negatively impactful bias.” Negatively impactful bias/partiality occurs when considerations that should not influence the decision become involved in the decision making process in some way. For example, gender may be an important consideration as part of the decision process when discussing disparity in pay amongst genders; yet it has nothing to do with decisions relating to a car accident. So, if this particular attribute has appeared in the decision making process on a car case and has potentially influenced the outcome, it should be considered to be impactful. And, the system should work to eliminate the consideration and the impact.

This may provide us with our first solid example of beneficial technological auditing, which could be done within arbitration institution technology. Consider an auditing algorithm that discovers patterns in outcomes in situations where men are the victim of

---

<sup>730</sup> See *supra* Chapter 1, Section D.

<sup>731</sup> Mill, UTILITARIANISM, *supra* note 97.

<sup>732</sup> See *supra* Chapter 4, Section C(3).

violence by a female. Based on this discovery, human decision makers<sup>733</sup> could choose to adjust the system parameters so that certain types of information (gender) are not provided to the decision maker in certain cases (car accidents). Alternatively, the algorithm might discover patterns in the decisions made by a particular individual. In this instance the system may be designed to provide additional information to the individual to assist them in recognizing and overcoming the pattern.

While these examples may seem far-fetched, in fact systems already do behave in this way. As my co-authors and I explained in *The Angel on Your Shoulder: Prompting Employees to Do the Right Thing Through the Use of Wearables*<sup>734</sup> employers already widely use devices and the monitoring of behavior to measure employment targets and employee behaviors. In fact, many have created nudges to increase employee compliance with directives about everything from attentiveness to healthy behavior.<sup>735</sup> Transparency and auditing can and should be important aspects of ensuring that negatively impactful bias and partiality are reduced.

#### 6. There Must Be A Human Within The Final Step of the Technology-Driven Decision Making Model

As was previously explored in Chapter 1 and Chapter 4, current deployments of AI and machine learning cannot fully capture legal reasoning and/or discretion in decision making.<sup>736</sup> It is therefore the author's contention that as of the writing of this thesis, there must be a human in the decision making process to act as the final decision maker. This section will break down the arguments leading to this determination, explore its consequences for the overall model, and finally will explore possible solutions for the necessity of using human judgment.

---

<sup>733</sup> As designers of the institution and within the expectations auditing of the technology to improve the system.

<sup>734</sup> See Timothy Fort, Anjanette Raymond, and Scott Shackelford, *The Angel on Your Shoulder: Prompting Employees to Do the Right Thing Through the Use of Wearables*, NORTHWESTERN JOURNAL OF TECHNOLOGY AND INTELLECTUAL PROPERTY, 14(2) (2016).

<sup>735</sup> See *id.*

<sup>736</sup> See *supra* Chapter 1, Section 4; Chapter 4, Section B.

*a. Rejecting the Application of an Individual Jurisprudential School*

As should be recalled from Chapter 1, legal formalists consider law to be “rationally” determinate, *i.e.*, the class of legitimate legal reasons available for a judge to offer in support of his decision justifies one and only one outcome, either in all cases or in some significant range of cases; and (2) adjudication is thus “autonomous” from other kinds of reasoning, that is, the judge can reach the required decision without recourse to non-legal normative considerations of morality or political philosophy.<sup>737</sup> As such, the use of a formalist school of jurisprudence as the foundation of model design will likely lead model developers, in conjunction with legal experts, to distill the law down to a concrete set of rules, sub-rules and exceptions, with outcomes measured against a metric designed to lead to a consistent outcomes. Discretion, in its most basic sense, would be rejected; instead objective, measurable, consistent outcomes would be the hallmark of a successful justice technology.

Should ODR design elect to use a legal formalism approach to design, the technology-based decision makers will be programmed within a rule-based structure. The rules, drawn from a variety of published legal codes, will remain static. The system will be tested to ensure that the outcome remains consistent — and consistency across similar situations will be paramount to success. Like the solution to a simple math problem, the outcome will be fully determined once all the relevant rules have been engaged.

In contrast, legal realists would expect the system model to embrace “an unsentimental and honest account of what judges really do.”<sup>738</sup> In this type of system, designers may elect to work backwards — examining existing cases to discover patterns and correlations that can be used to create rules that reflect the existing judicial system. These discovered rules would likely be layered on top of the law in a way that would help mimic the decision making process of human justices.

Of course, discussion of the law and rules, when considered within jurisprudence, is often much too simplistic, primarily because of the highly variable and intricately

---

<sup>737</sup> See *supra* Chapter 1, Section E(1).

<sup>738</sup> See *supra* Chapter 1, Section E(2).

nuanced micro-judgments that occur throughout the decision making process. Consider the case of *Winfield v. City of New York*,<sup>739</sup> in which United States Magistrate Judge Katharine H. Parker was faced with a dispute concerning e-discovery, mainly a challenge to the predictive coding driving the search engine.<sup>740</sup> The heart of the dispute focused on the active machine learning process that required an algorithm to identify “relevant” documents.<sup>741</sup> Active machine learning is built on input from subject matter experts (SME), that is, persons who understand the area of interest being explored. If the SME misdefines an essential term, in this case relevance, then so will the AI. The argument advanced by the plaintiffs was that improperly coded documents in the training dataset resulted in a multitude of miscoded documents in the algorithm’s output.<sup>742</sup> The result, it was argued, was that many documents had been erroneously identified as irrelevant<sup>743</sup> and thus were not produced.<sup>744</sup>

It is important to note that in the *Winfield* case the judge had issued several rulings relating to relevance<sup>745</sup> — and yet, the rule used to drive e-discovery proved problematic. At this stage in the discussion, the specifics of the *Winfield* ruling are of little concern; what is important to note is that this is a debate about the threshold at which human discretion becomes necessary. Certain essential and widely used terms within the justice system, such as relevance, are incredibly difficult to define, and in practice the nuances are often decided within the parameters of a specific case, relying on human interpretation and judicial discretion.

In addition, when designing technology to predict outcomes or assist humans in basic analysis, it is important to note that issues can arise within the context of “interpretation” of statutory language. Despite the existence of what appears on the face to be very specific language, interpretation is an essential consideration when analyzing legal text. The nuisance of interpretation leads to problems within the technology-

---

<sup>739</sup> 2017 U.S. Dist. LEXIS 194413 (SDNY, Nov. 27, 2017).

<sup>740</sup> See *Winfield*, U.S. Dist., at para. \*1.

<sup>741</sup> See *id.* at para. \*4.

<sup>742</sup> See *id.* at para. \*2.

<sup>743</sup> See *id.*

<sup>744</sup> See *id.*

<sup>745</sup> See *Winfield*, U.S. Dist., at para. \*2.

driven system. Applying specific language presents difficulties throughout the justice system, this can be demonstrated by the U.S. Supreme Court case of *Digital Realty Trust v Paul Somers*.<sup>746</sup> In the case, the justices ruled unanimously that employees are not protected from retaliation if they “blow the whistle” on alleged corporate misdeeds without going to the Securities and Exchange Commission.<sup>747</sup> The ruling was based upon the *plain language* of the Dodd-Frank law, which limited those eligible for whistle-blower protections to “any individual who provides . . . information relating to a violation of the securities laws to the Commission . . .”<sup>748</sup> Despite the plain language, and the reading falling clearly within Dodd-Frank’s purpose and design<sup>749</sup> two lower federal courts and the SEC themselves had misapplied the statute.<sup>750</sup> As can be inferred from this specific case, even highly specific, plain language statutes are often subject to widely different and incredibly variant interpretations. Introducing technology into the decision making process does not eliminate the issue, it merely hides the interpretation difficulties behind a veil of code.

Technology in the justice system must bring about a convergence of hard and fast rules with **appropriately bounded discretion**, which can alternatively be defined as **non-negatively impactful bias**. As U.S. Supreme Court Justice Sonia Sotomayor describes:

But there are situations in which some experiences are important in the process of judging because the law asks us to use those experiences.<sup>751</sup>

---

<sup>746</sup> *Digital Realty Trust, Inc. v. Somers*, 583 U.S. \_\_\_\_ (2018).

<sup>747</sup> See *Digital Realty Trust*, 583 U.S. at 2.

<sup>748</sup> See *Digital Realty Trust*, 583 U.S. at 2 relying upon 15 U. S. C. §78u– 6(a)(6).

<sup>749</sup> The core objective of Dodd-Frank’s whistleblower program is to aid the Commission’s enforcement efforts by “motiv[at]ing people who know of securities law violations to tell the SEC.” S. Rep. No. 111–176, p. 38 (emphasis added) *cited in Digital Realty Trust, Inc. v. Somers*, 83 U. S. (2018).

<sup>750</sup> See Richard Wolf, *Supreme Court Deals Blow To Wall Street Whistle-Blowers With Unanimous Ruling*, USA TODAY (Feb. 21, 2018).

<sup>751</sup> Sotomayor, *Confirmation Hearings*, at 71, 120 (statement of J. Sonia Sotomayor) contained in WASHINGTON POST TRANSCRIPT, *Sen. Jeff Sessions Holds a Hearing on the Nomination of Judge Sonia Sotomayor to Be an Associate Justice of the U.S. Supreme Court*, CQ TRANSCRIPTIONS, (Tuesday, July 14, 2009) available at <http://www.washingtonpost.com/wp-dyn/content/article/2009/07/14/AR2009071401155.html?noredirect=on> (last visited Aug. 5, 2019).

Thus, we as a society, and specifically the designers of models for technologically-enhanced justice, *must* be able to distinguish between concepts that are eligible to be represented as rules and areas where discretion is essential. We must build both of these aspects into the technology-assisted justice system. In making this argument, the author supports a hybrid jurisprudence, which is certainly far from an uncontested ideal. Such models have come under fire for years, as highlighted by jurist John Selden in a well-trodden notation:

'Tis all one as if they should make the standard for the measure we call a foot, a Chancellor's foot; what an uncertain measure would this be? One Chancellor has a long foot, another a short foot, a third an indifferent foot: 'tis the same thing in a Chancellor's conscience.<sup>752</sup>

This criticism, however, fails to distinguish subjectivity from discretion. And this distinction is in fact precisely where properly deployed technology might be able to help. Technology allows for the creation of specific rules that bound the permissible level and particular application of discretion, limiting the potential for random, unpredictable, or footwear-based judicial vagaries. For example, the system could be designed to ask for input from a human, yet to limit the input to a specific list of choices, or could ask the human to record an explanation of the choice made. In this way, discretion could be permissible, in a specific set of circumstances, yet would be bounded by the choices presented. As discussed in Chapter 1, Lord Bingham strongly advocated for bounded discretion,<sup>753</sup> and if done correctly the system could ensure this essential RoL principle is part of the design process.<sup>754</sup>

*b. Rejection of the “Robot” Decision Maker*

In 2001, a speech given by then Supreme Court nominee Sonia Sotomayor at the University of California, Berkeley drew sharp focus upon a longstanding belief that impartiality demands a neutral and detached decision maker and, more particularly, one

---

<sup>752</sup> John Selden, Table Talk, quoted in M. B. Evans and R. I. Jack (eds), *Sources of English Legal and Constitutional History*, BUTTERWORTHS, Sydney, 223-224, (1984).

<sup>753</sup> See Bingham, *supra* note 56. See discussion Chapter 1, Section A.

<sup>754</sup> One can argue, this particular aspect of model design will be left to the model designer, i.e. the person tasked with creation by the arbitration institution. It is an open question, however, how this would be- or if it should be- an expectation for the enforcement of the outcome of the process.

who is in no way guided by emotion.<sup>755</sup> Nominee Sotomayor states: “I would hope that a wise Latina woman with the richness of her experiences would more often than not reach a better conclusion than a white male who has not lived that life.”<sup>756</sup> This talking point raised its head during the confirmation hearings in which Nominee Sotomayor was asked by Sen. Jeff Sessions (R-AL) about whether judges should allow their “prejudices” to “impact their decision making.”<sup>757</sup>

Sotomayor addressed this controversy:

Never their prejudices. I was talking about the very important goal of the justice system is to ensure that the personal biases and prejudices of a judge do not influence the outcome of a case. What I was talking about was the obligation of judges to examine what they’re feeling as they’re adjudicating a case and to ensure that it’s not influencing the outcome. Life experiences have to influence you. We’re not robots to listen to evidence and not have feelings. We have to recognize those feelings and put them aside. ... But there are situations in which some experiences are important in the process of judging because the law asks us to use those experiences.<sup>758</sup>

She went on to add:

I think the system is strengthened when judges don’t presume they’re impartial, but when judges test themselves to identify when their emotions are driving a result, or their experiences are driving a result, and the law is not.<sup>759</sup>

Nominee Sotomayor’s statement flew in the face of the classical liberal view that depicts judges as simply dispassionate dispensers of objective rules.<sup>760</sup> The ideal of the emotionless, “dispassionate” judge has a very long pedigree. More than three centuries ago, Thomas Hobbes wrote in *Leviathan* that the ideal judge is “divested of all fear, anger, hatred, love and compassion.”<sup>761</sup> As previously discussed, within this classical view, the judge disinterestedly “declared” a law that existed external to and

---

<sup>755</sup> See Sotomayor, *Confirmation Hearings*, *supra* note 672 at 71.

<sup>756</sup> *Id.*

<sup>757</sup> See *id.*

<sup>758</sup> *Id.*

<sup>759</sup> *Id.*

<sup>760</sup> See *supra* Chapter 1, Section D, and Section E(1).

<sup>761</sup> Thomas Hobbes, *LEVIATHAN* 203 (A.R. Waller ed., 1904) (1651).



independently of him, for law was a science of rules, and judgment was the deductive enterprise of subsuming particular facts under general norms.<sup>762</sup>

However, Justice Sotomayor was not the first to reject the classical view of law as rules rather than of men. Supreme Court Justice Robert Jackson once wrote, “dispassionate judges” are “mythical beings,” like “Santa Claus or Uncle Sam or Easter bunnies.”<sup>763</sup> In 1921 Justice Benjamin Cardozo recognized that “[d]eep below consciousness are other forces, the likes and the dislikes, the predilections and the prejudices, the complex of instincts and emotions and habits and convictions, which make the man, whether he be litigant or judge.”<sup>764</sup> Justice William Brennan also rejected the classical view and instead acknowledged the important role that qualities other than reason must play in the judicial process. “In ignoring these qualities, the judiciary has deprived itself of the nourishment essential to a healthy and vital rationality.”<sup>765</sup> In 2006, then nominee and now Supreme Court Justice Samuel Alito offered a specific example when he said “when I get a case about discrimination, I have to think about people in my own family who suffered discrimination because of their ethnic background or because of religion or because of gender.”<sup>766</sup>

The technology assisted decision making process here proposed supports and enhances the judicial self-reflection championed by Justice Sotomayor and others. It will be designed to monitor patterns and provide judges with a reflection that encourages them to be mindful of the borders between experience, passion and impartiality. It will provide bounds within which judicial discretion is permissible, and indeed, this author contends, necessary.

---

<sup>762</sup> See *supra* Chapter 1, Section D, and Section E(1).

<sup>763</sup> *United States v. Ballard*, 322 U.S. 78, 93-94 (1944) (Jackson, J., dissenting).

<sup>764</sup> Benjamin N. Cardozo, *THE NATURE OF THE JUDICIAL PROCESS* 167-68 (1921)

<sup>765</sup> William J. Brennan, Jr., *Reason, Passion and the “Progress of the Law*, 10 *CARDOZO L. REV.* 3, 9 (1988) (arguing that by ignoring “the range of emotional and intuitive responses to a given set of facts or arguments . . . the judiciary had deprived itself of the nourishment essential to a healthy and vital rationality.”). *Id.*

<sup>766</sup> Emily Bazelon, *Mysterious Justice*, *NEW YORK TIMES MAGAZINE*, referencing his testimony in his confirmation hearings, (March 18, 2011)(commenting on his lone dissent in the free-speech suit Albert Snyder brought against the Westboro Baptist Church for picketing the funeral of his son, a soldier who was killed in Iraq.) *Id.*

*c. The Emotionally Intelligent Decision Maker*

Having determined that a purely technological, rule-focused neutral decision maker is ultimately not capable of performing the full range of judicial duties, we now turn to clarifying what the components that lie outside of the technological sphere comprise. This section will begin to frame a theory of the “emotionally intelligent” decision maker.

Professor Maroney provides a launching point for the argument:

Within law, we have inherited some hefty cultural baggage, weighted down with the belief that a good judge is emotionless. This article has unpacked that baggage and suggested that it is that belief, not emotion, that should be put aside. We need a new ideal: that of the emotionally intelligent judge. The emotionally intelligent judge is self-aware and is able to think coherently about her emotions and to be in control of their expression. She is willing to seek the opinions and support of others and approaches the emotional challenges of the job with openness and flexibility.<sup>767</sup>

It is time we reject the historic concepts of a detached decision maker, an unobtainable ideal that only serves to cloak the impact of bias, prejudice and partiality in judicial decision making, converting them into supposed objectivity and so embedding bias into the system. Instead, we need decision makers well trained to be self-aware of their own potential leanings. It is the ideal of “self-awareness” that we must ensure is embedded into the decision makers within the online justice system.

With the use of technology it may be possible to go a step further — requiring the arbitrator to be aware of their own thought processes, but also aiding them in that responsibility with a system that can signal the potential introduction of bias.

It bears revisiting an important utterance from U.S. Supreme Court Justice Sonia Sotomayor who argues:

I think the system is strengthened when judges don't presume they're impartial, but when judges test themselves to identify when their

---

<sup>767</sup> Terry A. Maroney, *The Emotionally Intelligent Judge: A New (and Realistic) Ideal*, 49 COURT REVIEW 100 (2013); VANDERBILT PUBLIC LAW RESEARCH PAPER No. 15-33 (2013).

emotions are driving a result, or their experiences are driving a result, and the law is not.<sup>768</sup>

Ultimately, there are times when a human must be able to intervene in an AI decision making process to exercise discretion in areas where concepts are too nuanced for the rigid rules of the system to fully capture, and there are times when the human should be guided by feedback from the AI decision making program to recognize partiality and steer back toward the clear bright lines of the rules. There are of course, immediate concerns about the introduction of a new source of potential bias when a human is introduced into the decision making process. As such, at all times the AI decision making process and the intervening human decision maker should be monitored for the introduction of bias. Thus, the intervening human decision maker should also be monitored by those deploying the assistive technology to ensure that the AI outputs are used as guidance and not a shortcut to arriving at conclusions/decisions.

The current version of the guidelines regarding algorithmic decision making and the General Data Protection Regulation (GDPR) accords with our view that humans are in fact required to intervene in AI decision making.<sup>769</sup> However, it goes much too far in requiring that a human reviewing an algorithmic decision consider “*all* the available input and output data” (emphasis added).<sup>770</sup> This renders a major benefit of the technology, its processing power, moot, and is likely no less unrealistic than the expectation of pure disinterest. Instead, higher-level auditing can control for bias in the data, the process, the outcome and the impact of the decisions. As argued by the Center for Data Innovation:

They (humans) can look at the algorithm’s aggregate behavior to check for evidence of bias: for example, by checking whether particular outcomes correlate to any protected characteristics (such as ethnicity, religion, or sexuality), and if so, by finding the data points related to those characteristics that lead to those outcomes, deciding whether they are fair, and if not, adjusting the algorithm accordingly. If necessary, a human auditor can

---

<sup>768</sup> See Sotomayor, *Confirmation Hearings*, *supra* note 672 at 71.

<sup>769</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, *On The Protection Of Natural Persons With Regard To The Processing Of Personal Data And On The Free Movement Of Such Data*, at Art. 22, (repealing Directive 95/46/EC (General Data Protection Regulation))(2016).

<sup>770</sup> *Id.*

also consider the most important or sensitive data involved in an individual case, which will vary depending on the nature of the decision—but aggregate auditing will likely remain a far more effective way of identifying bias.<sup>771</sup>

Auditing will identify problematic algorithmic behavior and humans will determine how to respond. Auditing of outcomes, especially those identified as highly impactful, for negative impact is especially necessary. Thus, regulation specific to the identification of algorithmic problems and the human response to the problems is a necessary step in the governance process.

#### 7. Accountability is Essential as ODR is part of the Justice Eco-System

This thesis has presented accountability as a guiding aspect of the creation of an ODR environment that seeks to be compliant with RoL principles.<sup>772</sup> This is a concern in several key areas: mainly the accountability of those that deploy the technology and accountability of those that use technology as a tool to assist in the decision making process. Both aspects of accountability are essential to the overall ODR eco-system.

Accountability as it relates to the design and deployment of technology is a topic of growing interest in the AI community. Although few regulations currently place such a burden on designers or those that deploy technology, it is hopefully the case that more legal accountability will occur in the near future.

Despite the lack of regulation, there is an emerging set of best practices, reports and guidelines on the use of AI. For example, the Partnership on AI released a *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System* in which the partnership examines the scope of the data-driven problem in the use of prediction-based AI in the U.S. justice environment and makes specific recommendations.<sup>773</sup> The

---

<sup>771</sup> Center For Data Innovation, *Comments To The Article 29 Working Party(WP29) On Its Guidelines Regarding Algorithmic Decision-Making And The General Data Protection Regulation (GDPR)*, at 3, (Nov. 2017).

<sup>772</sup> See *supra* Chapter 5, Section A(7).

<sup>773</sup> Partnership on AI, *Report on Algorithmic Risk Assessment Tools in the U.S. Criminal Justice System*, Partnership on AI Publication (2019).

Berkman Klein Center has released a series of topical papers focusing on the *Ethics and Governance of AI*<sup>774</sup> and has quickly emerged as a leader when it comes to considerations of the impact, governance, ethics and accountability of these technologies, terminology that will become increasingly familiar as the capture moves ahead.

Government has taken notice as well, for example, the European Commission recently released a *Draft Ethics Guidelines For Trustworthy AI*<sup>775</sup> which focuses on two components:

(1) it should respect fundamental rights, applicable regulation and core principles and values, ensuring an “ethical purpose” and (2) it should be technically robust and reliable since, even with good intentions, a lack of technological mastery can cause unintentional harm.<sup>776</sup>

And, the UK Information Commissioner's Office has created the *AI Auditing Framework*<sup>777</sup> and the most recently the Law Society of England and Wales undertook the task of exploring and reporting upon *Algorithms in the Criminal Justice System*.

Within the report they note:

When decision systems are introduced into public contexts such as criminal justice, it is important they are subject to the scrutiny expected in a democratic society. Algorithmic systems have been criticized on this front, as when developed in secretive circumstances or outsourced to private entities, they can be construed as rule making not subject to appropriate procedural safeguards or societal oversight.<sup>778</sup>

The report clearly supports several aspects of this thesis’ assertions. Although it is limited to criminal justice, the same concerns are directly applicable to the role of the

---

<sup>774</sup> The Berkman Klein Center for Internet & Society at Harvard University, *Ethics and Governance of AI*, (2017) available at <https://cyber.harvard.edu/topics/ethics-and-governance-ai> (last viewed Aug. 13, 2019).

<sup>775</sup> European Commission’s High-Level Expert Group on Artificial Intelligence (AI HLEG), *Draft Ethics Guidelines For Trustworthy AI*, E.C. Publication, (2018).

<sup>776</sup> *Id.* at i.

<sup>777</sup> U.K. Information Commissioners Office, *AI Auditing Framework*, ICO publication, (2018)( A formal consultation will be published towards the end of 2019.).

<sup>778</sup> The Law Society of England and Wales, *Algorithms in the Criminal Justice System*, *supra* note 602, at Section 5.4.3.2 at page 22.

decision maker within any justice environment. The report goes on to highlight a finding consistent with the author's assertion:

Few provisions currently support civil society organisations or forms of collective oversight of algorithmic systems directly, *leaving a significant accountability gap in need of remedy.*(emphasis added)<sup>779</sup>

As noted, despite the growing body of guidance, policy makers have yet to enact an enforcement mechanism intended to ensure a specific level of accountability. Consequences for failing to live up to the expectations of society must be part of any ODR model that seeks to stand within the justice eco-system. As such, those who deploy technology must ensure that the designers are committed to following the best practices that are already established in the AI governance area.<sup>780</sup> For example, dependence, partiality and bias in models must be measured and mitigated and the outcomes should be presented in a manner that is easily understandable, does not conflate multiple outcomes, and includes confidence estimates. These best practices are embedded throughout the rest of the chapter.

And, those who actually deploy the technology must take responsibility for the post-deployment evaluation, monitoring and auditing of the technology.<sup>781</sup> For example, transparency of designs, architectures and training data must be open to all (e.g., via model cards), including to internal and external auditing, research, review and criticism.

In addition, the individual human decision maker must be accountable for his interactions with the technology and be receptive to the feedback mechanisms built into the system.<sup>782</sup> While systems can (and should) be designed with an eye toward assisting the human user, that should not alleviate the accountability of the end user, in this case the decision maker. The human decision maker should be expected to attend trainings

---

<sup>779</sup> *Id.*

<sup>780</sup> In this instance, the arbitration institutions.

<sup>781</sup> In this instance, arbitration institutions are the primary focus of consideration as they build and deploy current technology, at least to date.

<sup>782</sup> Feedback should be part of the design and as such, is the responsibility of the entity deploying the system, the arbitration institution. However, the individual arbitrator should be held to a standard as well. One could imagine, in the future, specific rules of Digital Arbitrator Ethics designed by and enforced through the appointment process of arbitration institutions.

on the nature and limitations of the particular technology. And, auditing- done by the arbitration institution- should not merely include auditing of the technology based process and outcomes, but should also include monitoring and auditing of the human within the system. Monitoring and auditing of behavior and outcomes can provide an opportunity for learning and correction via feedback to the decision maker. The system can provide feedback when it notices a pattern, describe the pattern, and in some circumstances demand the human correct behavior or provide an explanation for rejecting the feedback. The feedback can escalate into a consequence-driven system should patterns not be appropriately explained by the decision maker.<sup>783</sup> As will be described later, examples of this type of monitoring can include flagging time spent on documents or other warning signs that the individual is not appropriately engaging with the presented material.

The brick-and-mortar system would likely benefit from such monitoring. For example, consider the recent U.S. Supreme Court case of *Flowers v. Mississippi*<sup>784</sup> in which a black man was repeatedly denied his right to a “criminal trial free of racial discrimination in the jury selection process.”<sup>785</sup> Mr. Flowers had been tried six times, each time with the same district attorney using his preemptory challenges to remove potential black jurors. In fact, the district attorney denied participation of 41 of the 42 black candidates.<sup>786</sup> There is widespread criticism surrounding this discriminatory behavior and commentators are dismayed that this case needed to reach the U.S. Supreme Court. Inserting technology into the process could mean that such patterns of clear discrimination are addressed earlier.

It is important to keep in mind, however, in the case of a brick-and-mortar justice environment, checks and balances are already in place, including the right of appeal. In an arbitration system, by contrast, appeals are not permitted<sup>787</sup> and the outcome is final

---

<sup>783</sup> Enforcement, via the institution, is essential as the entity approving appointment of the individual arbitrator, the institution has both authority and the ability to consequent (and create incentives) for individual to be compliant with expectations.

<sup>784</sup> 588 U. S. \_\_\_\_ (2019) No. 17–9572. (Decided June 21, 2019).

<sup>785</sup> *Id.* at Justice Kavanaugh, for the majority.

<sup>786</sup> *Id.* at Justice Thomas, dissent.

<sup>787</sup> *See supra* Chapter 2, Introduction.

and binding.<sup>788</sup> The commitment to a trustworthy, RoL-compliant alternative justice environment demands transparency, monitoring and auditing of outcomes.<sup>789</sup> The best use of technology would be to allow patterns to be discovered and revealed to an individual. In these instances, when discriminatory patterns persist, we must seek to use technology to begin to create a more just justice environment.

Finally, there is one check and balance that does apply within the ODR eco-system specifically: the balance between the institution that produces the outcome (the arbitration institution) and the institutions that enforce the outcomes, namely the justice system itself. As explained in Chapter 2, this thesis envisions arbitration as the final step in the ODR process, meaning the RoL is protected via the courts as the authority enforcing the outcome.<sup>790</sup> As such, policymakers must ensure that the technology, the human decision maker and the ODR eco-system as a whole is meeting best practices standards. They accomplish this through appropriate regulation created within the court system in the enforcing jurisdiction. An AI based system should be held to a standard that meets best practices and complies with RoL, including public policy. Policy makers can ensure this occurs by creating regulation for the enforcement of AI-based arbitration awards

#### 8. Best Practices Demand a Wider Conversation about ALL of these Issues

A survey conducted by Ari Kaplan Advisors has found that U.S. lawyers have significant knowledge gaps around legal tech and tend to see the subject through the prism of security and the need to understand e-discovery, rather than focus on more strategic tech issues.<sup>791</sup>

In a similar vein, technical literacy is lacking in many areas. For example, Lipton finds most troubling, that the rise of social media has allowed self-proclaimed “AI influencers” who do nothing more than paraphrase Elon Musk on their Medium blogs

---

<sup>788</sup> *See id.*

<sup>789</sup> *See supra* Chapter 4, Section C.

<sup>790</sup> *See supra* Chapter 2, Section Introduction.

<sup>791</sup> Author, *U.S. Survey Finds Big Legal Tech Knowledge Gap Among Lawyers*, ARTIFICIAL LAWYER, (July 20, 2018).



to cash in on this hype with low-quality, TED-style puff pieces.<sup>792</sup> “Making real progress in AI requires a public discourse that is sober and informed.”<sup>793</sup> “. . . Right now, the discourse is so completely unhinged it’s impossible to tell what’s important and what’s not.”<sup>794</sup> “There are policymakers earnestly having meetings to discuss the rights of robots when they should be talking about discrimination in algorithmic decision making. But this issue is terrestrial and sober, so not many people take an interest.”<sup>795</sup> The time is now to begin wider conversations about algorithmic decision making, especially in areas that have strong impact upon us as a society. The absence of these conversations will lead to RoL issues within the ODR community and the legal system as a whole.

## B. Defining the Key Terminology

Model design depends on well-defined terminology that is specific to the model and the area of deployment. Thus, in order to specify a model, we must narrowly define several essential terms for the sections that follow.

For example, as was examined and explained in Chapters 2 and 3:

[I]t is a fundamental and universally accepted arbitration principle of international arbitration that arbitrators have to be impartial and independent of the parties and must remain so during the proceedings.<sup>796</sup>

While *impartial* and *independent* might be considered universally understood terms, for current purposes they must be considered specifically within an online environment, online, partially automated, platform-based environment involved in the nomination, selection, and confirmation of human decision makers. Existing criteria and boundaries must be reconsidered. The following sections will create the independence and impartiality parameters appropriate for the online environment.

---

<sup>792</sup> Oscar Schwartz, *The Discourse Is Unhinged!': How The Media Gets AI Alarmingly Wrong*, THE GUARDIAN, (July 25, 2018).

<sup>793</sup> *Id.*

<sup>794</sup> *Id.*

<sup>795</sup> *Id.*

<sup>796</sup> Lew, COMPARATIVE INTERNATIONAL COMMERCIAL ARBITRATION, *supra* note 140 at 256.

## 1. Decision Maker

The use of the term “decision maker” up to this point has been intentionally unspecified as it was necessary to allow for the exploration of key issues without ascribing any additional details to the conversation. For example, in the context of arbitration and the judicial system, the terms arbitrator and judge are presumptively envisioning a human actor. However, as we develop definitions and rules for the online environment it is now time to clarify the definition of the term decision maker as a system component that includes more than a human actor.

In the context of arbitration, of course, there are instances when the decision maker is in fact a human actor and the presence of a human within the decision making loop is an important aspect of the online justice decision making model. It is also important, however, to understand that decisions can be made by the technology itself. Ethan Katsh has long called the technology the “Fourth Party” in the resolution process. As defined by Professor Katsh “the metaphor of ‘the fourth party’ emerged out of a recognition that software could play a role that might replace a mediator in simple cases and, more commonly, would assist and collaborate with the neutral third party. . .”<sup>797</sup>

In the context of this thesis, the term *decision maker* has three possible meanings: (a) a human decision maker who uses technology to gather information, assist in party communication and perform similar activities; (b) a human decision maker substantively aided by technology, such as when the technology makes some early categorizations, classifications, suggestions and recommendations, in some cases even nudging parties during the negotiation stage, but the human is responsible for making the final arbitration determination; and (c) technology as the sole decision maker. As this thesis continues to develop its topic, it will specify which meaning applies.

## 2. Dispute

Because technology will become more embedded into the dispute resolution process and automation will undoubtedly increase, the definition of 'dispute' will become

---

<sup>797</sup> Katsh, *Digital Justice*, *supra* note 251, at 93.

essential in the overall dispute planning process. Automation and/or technology will be able to resolve some issues; issues that currently may be considered a dispute by many. For example, the Internet of Things (IoT) is widely deployed in supply chain transportation,<sup>798</sup> because IoT devices are prevalent, issues such as appropriate temperature control and regulation and similar issues may now be resolved by reconciling information disparities amongst IoT sensors through more frequent, and more widespread information gathering.<sup>799</sup> Issues such as this, that would previously require human interaction and dispute settlement, will now be reconciled in the technology. If multiple sources of information reveal a temperature variance greater than allowed, no dispute occurs, the technology will simply notify the part of the temperature variance and the part will be able to take action.<sup>800</sup> However, there will still be disputes in a technology driven world. For example, when a party receives goods that have suffered exposure to the elements resulting in the grading of the goods needing to be lowered. In this instance, the determination of the impact of the exposure will need to occur, most likely, outside of the technology environment. Such determination may require an expert to 'grade' the goods in their current, delivered condition. Technology, in general will not be able to do this type of activity and as such, it will need to be done outside the technology environment. Hence, disputes are more than activity done to merely reconcile information, a dispute is an issue that is a matter of disagreement, one that needs resolving through a process of human critical evaluation.

### 3. Automated Decision Systems

According to the United States Algorithmic Accountability Act an automated decision system is:

a computational process, including one derived from machine learning, statistics, or other data processing or artificial intelligence techniques, that

---

<sup>798</sup> See Daniel Newman, *How IoT Will Impact The Supply Chain*, Forbes, (Jan. 9, 2018).

<sup>799</sup> See *id.*

<sup>800</sup> See *id.*

makes a decision or facilitates human decision making, that impacts consumers.<sup>801</sup>

This definition will be adopted in our design, with the exception of the limiting term “consumer” being given a broader definition to encompass any entity using the system to resolve a dispute.

#### 4. Independence

As previously explored, in general the arbitration community considers dependence to arise from relationships between an arbitrator and one of the parties, or with someone closely connected with one of the parties.<sup>802</sup> In a hybrid decision making system dependence is measured by exploring the financial, business, and personal relationships of any human decision maker *and* the platform (technology based), which should include the platform owners, stakeholders, and auditors, amongst others. When evaluating the *appearance*<sup>803</sup> of dependency of the decision maker, the *reasonable person*<sup>804</sup> standard should be used, such that the standard for independence in an online environment becomes: would the facts and circumstances surrounding the participation of this particular individual or platform lead a reasonable person to believe there is an appearance of a dependent financial, business, or personal relationship?

Based on the definition of decision maker it is important to note, the decision maker can be either human or the technology. As such, dependence is measured by exploring the financial, business, and personal relationships of any human decision maker *and* the platform (technology based) decision maker, which should include the platform owners, stakeholders, and auditors, amongst others.

#### 5. Impartiality

---

<sup>801</sup> U.S. Algorithm Accountability Act of 2019, “S. 1108 — 116th Congress: Algorithmic Accountability Act of 2019 (Wyden Senate introduction text); H.R.2231 — 116th Congress (2019-2020).

<sup>802</sup> See IBA, *Ethics*, *supra* note 184 at Cannon 3.

<sup>803</sup> See *id.*

<sup>804</sup> See *id.*

Partiality arises when an arbitrator favors one of the parties, or where he is prejudiced in relation to the subject-matter of the dispute.<sup>805</sup> As with dependence, when evaluating the *appearance*<sup>806</sup> of partiality of the decision maker, the *reasonable person*<sup>807</sup> standard should be used. Partiality is relatively more difficult to identify and eliminate than dependence, however, creating an opportunity for technology-based improvement.

Some common types of partiality overlap with dependency issues.<sup>808</sup> For example, partiality can be presumed to exist when a decision maker is determining issues related to a family member. In these instances the decision maker is both dependent and presumed partial. It is instances of partiality that arise outside of dependency that are the more difficult to discover. Partiality is about a state of mind of the decision maker, and is often thought of as an emotion, predetermination, or the introduction of personal opinions that are not relevant to the determination of the issue at hand.<sup>809</sup> Because of the inherent difficulty of discovering partiality in the mind of an individual, arbitration institutions protect against it in two main ways: (1) requiring self-disclosure by the individual decision maker, and (2) allowing for after-the-fact challenges by one or both of the parties to the dispute alleging the partiality of an arbitrator as demonstrated by his/her behavior during the process.<sup>810</sup>

Both of these solutions can also be implemented in the online environment, and enhanced through the use of technology. In many ways, the term impartiality as used in the arbitration community to describe human decision makers can be considered equivalent to the term bias as used in the technology community.

## 6. Bias

Machine or algorithm bias is the effect of erroneous assumptions in machine learning processes. In technology research, partiality and dependence are aspects of bias. As

---

<sup>805</sup> IBA, *Ethics*, *supra* note 184 at Cannon 3.

<sup>806</sup> *Id.*

<sup>807</sup> *Id.*

<sup>808</sup> *See supra* Chapter 1, Section C and D; Chapter 2, Section A.

<sup>809</sup> *See supra* Chapter 1, Section C and D; Chapter 2, Section A; Chapter 5 Section B(4).

<sup>810</sup> *See supra* Chapter 2, Section B and Section C.

previously discussed in Chapter 4, machine/algorithm bias takes various forms.<sup>811</sup> One of the most prominent examples involves the use of machine learning systems to make judgments about individual people or groups of people. Within this chapter, we will explore machine/algorithm bias which arises within the decision making process or production of an outcome.

## 7. Negatively Impactful Bias/Dependence/Partiality

Negatively impactful bias occurs when the outcome is altered by the inclusion of a defined bias and the outcome has a defined negative impact on one or both of the parties.<sup>812</sup> Bias is considered to be problematic in the justice environment not merely when it exists but more narrowly when it leads to consideration of a factor or factors that are beyond the case and the factor or factors have an *actual* impact upon the outcome.<sup>813</sup> For example, while the introduction of race or socioeconomic status into a decision making process is a bias, it is not necessarily negatively impactful in all situations. Race, gender, or economic status could all be weighted as a positive factor, for example when it is treated as a positively weighted “adversity score” in college admissions.<sup>814</sup> As described in Chapter 4, because machine learning works from the already biased inputs it is given, similar biases can arise within a decision making process with or without AI.<sup>815</sup> In AI-assisted decision making processes, there must also be a determination of whether the bias has produced a negative impact.

### C. Addressing Bias in Technology-Driven Decision Making Models

As discussed in Chapter 4, “models specify how a problem is input and the type of [legal] result to output.”<sup>816</sup> In between, the *model builders* have constructed a computational mechanism to apply domain knowledge to perform the steps and

---

<sup>811</sup> See *supra* Chapter 4, Section A(3).

<sup>812</sup> See *supra* Chapter 5, Section B(5).

<sup>813</sup> See *id.*

<sup>814</sup> See Douglas Belkin, *SAT to Give Adversity Scores to Capture Social and Economic Background*, WALL STREET JOURNAL, (May 16, 2019)

<sup>815</sup> See *supra* Chapter 4, Section A(3).

<sup>816</sup> See *supra* Chapter 4, Section A(1).

transform the inputs to outputs.<sup>817</sup> Currently, as briefly discussed in Chapters 3 and 4, many different types of programs exist to model various legal processes.<sup>818</sup> For example, programs have been designed: (1) to model legal reasoning based on legal statutes and cases, (2) to integrate reasoning with legal rules, cases and underlying values, and (3) to predict outcomes of legal disputes.<sup>819</sup> And, of course, more models are being considered and tested every day. This section will consider general aspects of these types of models as relevant to the justice system, while the following section will present specific suggestions, in the form of best practices, for designing models of the neutral process and decision maker in the justice environment.

It is important to note, while the thesis has contained information and considerations of very advanced technology, as Chapter 4 concludes, we are likely years away from a “robot” decision maker.<sup>820</sup> Hence, the BEST PRACTICES that are presented at the end of each sub-section are based on currently existing possibilities. Moreover, these best practices will be general enough to capture many different technology deployments, not predicting any one technology as the future. The section that follows will put these best practices to work in a potential specific technology deployment designed to provide a neutral decision maker using arbitration as the final part of the ODR process.

## 1. The Model

The following sections and subsections will explain and then create BEST PRACTICES for (1) Model Selection, (2) Model Learning, (3) Weighting (if it were to occur), (4) Model Risk, and (5) Model Risk and Mitigation, including human actor mitigation.

### *a. The Reality of Legal Models*

At this time, there are portions of the justice system that can be, in certain terms, successfully captured by the kind of models discussed in Chapter 3, Section 4.<sup>821</sup> Yet,

---

<sup>817</sup> See Ashley, AI AND LEGAL ANALYTICS, *supra* note 135 at 4 (emphasis added).

<sup>818</sup> See *supra* Chapter 4, Section A(1).

<sup>819</sup> See *supra* Chapter 4, Section B.

<sup>820</sup> See *supra* Chapter 4, Section E.

<sup>821</sup> See *supra* Chapter 3, Section A and B; Chapter 4, Section A.

no existing model is able to combine all the components of a justice/arbitration process into a single stream of conclusions leading to a final determination. There are models that can predict outcomes, models that can use text analysis and other legal logic to answer simple legal enquiries, but no model (yet) which puts this all together to arrive at a legal conclusion.<sup>822</sup> Moreover, prediction-based models are simply not in the category of a functional justice replacement at all. This is because prediction by definition uses data from a justice system long criticized for various types of systemic bias in order to “predict” outcomes that replicate those patterns, warts and all.<sup>823</sup>

Keep in mind, there are systems currently being tested and deployed that seem to be incredibly good at prediction. In fact the creators of one such system boast “81% accuracy predicting whether claims relating to mortgage products will succeed (in the United Kingdom).”<sup>824</sup> Yet, as with most “disruptive” technology, few of the weightings, factors, or other key attributes used to train, improve, or monitor the process are publicly available, nor indeed available to the very human decision makers who use them. For example, in the instance of the mortgage prediction model the machine learning platform does what it is expected to do, and it seems to do it with a decent level of accuracy.<sup>825</sup> Nonetheless, to make a prediction it has to rely upon a data set, one that would undoubtedly contain biases that exist in the industry.<sup>826</sup> Why would we as a society want to further engrain a widespread problem, especially when the

---

<sup>822</sup> See *supra* Chapter 4(B).

<sup>823</sup> See e.g., Laurel Eckhouse, *Big Data May Be Reinforcing Racial Bias In The Criminal Justice System*, THE WASHINGTON POST, (Feb. 10, 2017)(discussing the reinforcement); Jeff Guo, *Researchers Have Discovered A New And Surprising Racial Bias In The Criminal Justice System*, THE WASHINGTON POST, (Feb. 24, 2016) (even in the face of existing protections); David Lammy, *The Racial Bias In Our Justice System Is Creating A Social Timebomb*, THE GUARDIAN, (Sept. 8, 2017) (considering the impact upon society of the bias).

<sup>824</sup> Author, *Meet Sibyl AI — The New Claims Prediction System*, ARTIFICIAL LAWYER, (July 20, 2018).

<sup>825</sup> See e.g., Kehinde Andrews, *UK Banks Have A Racial Discrimination Problem. It's Time They Admitted It*, THE GUARDIAN (Jan. 13, 2017) (revealing wide spread bias in banking, including in instances of fraud)

<sup>826</sup> See *id*; Patrick Wintour, *Nick Clegg To Accuse Banks Of Racism*, THE GUARDIAN (Nov. 23, 2011) (discussing the outcomes of a banking study and the political impacts) And in the U.S.: See Camila Domonoske, *Interactive Redlining Map Zooms In On America's History Of Discrimination*, NPR (the Two Way)(Oct, 19, 2016)(mapping redlining practices).



engraining will occur in a single, non-descript, black box? Thus, there is no machine learning model in current use that we should be comfortable deploying without a human in the decision making loop.

Thus, in this instance, BEST PRACTICE would be (a) to continue to develop and test models within the justice environment, but to insist upon specific design protections, including that a human must make the final decision. (b) The model should be viewed by the human as suggestive and outcomes should be displayed in a manner that fully informs the human decision maker of the nature of the predictions being made and the accuracy/percentage certainty of the prediction. The manner in which this could be deployed will be described more fully below in Section C.

*b. Model Selection*

The selection of the appropriate model depends upon a number of factors, the most primary in this instance being the intended use of the particular model.<sup>827</sup> And, it is in this decision that the difficulties of legal environment become readily apparent. For example, if the objective is to predict the outcome of a particular type of case, one would select a model designed to accomplish this objective, yet this does nothing to ensure accuracy or reliability of the outcome. Consequently, to make an appropriate selection of a model, the person making the selection *should* also understand how the model is designed;<sup>828</sup> currently, this is frequently not the case.<sup>829</sup> As previously discussed, Predictive Outcome Models are built upon a set of *features*.<sup>830</sup> Features are essentially brackets designed to hold particular data points,<sup>831</sup> for example the judge's name, the legal firm representing a client, the type of case/issue, and various attributes of the persons in the case, such as gender, race, or prior criminal record.<sup>832</sup> Machine learning can continue to add features into the model over time, and can adjust weighting within the model.<sup>833</sup> For example, the computer program may notice a pattern of

---

<sup>827</sup> *See supra* Chapter 4, Section A(1).

<sup>828</sup> *See id.*

<sup>829</sup> *See id.*

<sup>830</sup> *See id.*

<sup>831</sup> *See id.*

<sup>832</sup> *See id.*

<sup>833</sup> *See id.* at Section A(2).

correlating data points such as black male, prior conviction for a violent offense, under the age of 25. These data points, considered as unit, could be weighted within the process<sup>834</sup> so that individuals considered to be member of that unit are scored more likely to committed another violent offense. Or, the pattern may emerge<sup>835</sup> that a particular judge, or a particular court, is more likely to hand down guilty convictions later in the afternoon on Thursdays, or immediately after lunch. In most instances, no individual would notice these patterns. Machine learning notices them, and incorporates them into the model as new features that are used to build its future predictions.<sup>836</sup> Do note, such a model is performing exactly as expected — and may be ridiculously *accurate* in its predictions. Yet, that accuracy may be built upon features, unit identification, or weightings, that we as a society may not support. Thus, it is not merely enough to test the accuracy of the outcomes. The model selection, design, and outcomes MUST stand up to the scrutiny of the principles embedded within the RoL.

How, then, to perform initial and ongoing evaluation of the selected model? As previously described in Chapter 4, there are five errors the technology community identifies: Framing Trap, Portability Trap, Formalism Trap, Ripple Effect Trap, and Solutionism Trap.<sup>837</sup> Each requires consideration of the complexity of the environment and the human factor in the use. One idea that can be used to bridge some of the questions arising in model selection was also discussed in Chapter 4, that being model cards.<sup>838</sup> As Mitchell, Wu, Zaldivar, Barnes, Vasserman, Hutchinson, Spitzer, Raji, and Gebru recommend, model cards that document, in a fashion legible to the end user, key aspects, design considerations and performance characteristics should be issued with every model.<sup>839</sup> The author made the following minimal suggestions as to relevant information to be captured in the model card, for example: (1) model details, such as basic information about the model (person or organization developing model, model date and model type);<sup>840</sup> (2) information about training algorithms, parameters, fairness

---

<sup>834</sup> See *supra* Chapter 4, Section A

<sup>835</sup> See *id.*

<sup>836</sup> See *id.* at Section A(2).

<sup>837</sup> See *id.* at Section A(3).

<sup>838</sup> See Mitchell, *Model Cards*, *supra* note 517 at 22.

<sup>839</sup> See *id.*

<sup>840</sup> See *id.*

constraints or other applied approaches, and features;<sup>841</sup> (3) the intended use;<sup>842</sup> (4) the evaluation data, such as details on the dataset(s) used for the quantitative analyses in the card and training data;<sup>843</sup> and finally, (5) any ethical considerations or recommendations.<sup>844</sup> Based on this set of specific and detailed information released, a model can be evaluated for purposes both of initial selection and long-term ability to achieve its purpose. Maybe more importantly, model cards would be a good first step in improving transparency and holding designers more accountable for the model choices they make. The U.S. Algorithm Accountability Act of 2019 requires “a detailed description of the automated decision making system, its design, its training data, and its purpose.”<sup>845</sup> The Act, although new and yet to be passed, captures ideas described in this thesis, including transparency and accountability of automation and algorithms, essential considerations in an alternative justice environment.

Thus, in this instance, BEST PRACTICE would be (a) transparency of model selection, model deployment, the features used to train the model (if any), the features discovered and added to the model after training, and the features currently used in the prediction, and (b) some level of accountability for the deployment of the automation/AI, especially if it is to occur in a justice environment.

### *b. Model Learning*

An emerging area of research involves proxy discrimination. As described in Chapter 4, proxies are attributes that stand in place of other attributes.<sup>846</sup> In general, these are created via a learning process, when a particular attribute is unavailable in the data or restricted from use.<sup>847</sup> In these situations proxies are created that functionally stand in the place of other attributes.<sup>848</sup> The rejection of particular attributes or proxies often fails to fully eliminate discrimination as the machine learning algorithm merely

---

<sup>841</sup> *See id.*

<sup>842</sup> *See id.*

<sup>843</sup> *See id.*

<sup>844</sup> *See Mitchell, Model Cards, supra note 517 at 22.*

<sup>845</sup> *Algorithmic Accountability Act, supra note 714 at Sec. 2 (2).*

<sup>846</sup> *See supra* Chapter 4, Section A(2) and (3).

<sup>847</sup> *See id.*

<sup>848</sup> *See id.*

discovers a new proxy.<sup>849</sup> Thus, fighting proxy discrimination requires more than merely rejecting the use of certain attributes and proxies.

First, it is important to understand proxy discrimination within this model design. Proxy discrimination is a specific phenomenon that produces a disparate impact, and has been captured previously in the thesis by the term “negatively impactful bias”<sup>850</sup> to avoid limiting the discussion to U.S. law and to encourage a broader conceptualization of the issue. However, proxy discrimination is a specific U.S. constitutional legal concept that has been prominent in recent months (early 2019), so the narrower conceptualization is important to consider. For proxy discrimination as defined in the U.S. to occur requires two conditions: (1) a facially-neutral attribute that is relevant to achieving the objectives of a given model must be correlated with membership in a protected class, and (2) the predictive power of this facially-neutral characteristic must be at least partially attributable to its correlation with a suspect classifier.<sup>851</sup> While this thesis will not consider the full breadth of this U.S. specific test case, the technology-driven mechanisms of reducing the impact of proxy discrimination are important to consider as the thesis has consistently advanced the RoL expectation of a neutral decision maker, which demands taking measure of potential impactful bias.

In this instance, Prince and Schwarcz argue “One of the only ways to truly squash proxy discrimination is to remove access to potential substitute variables.”<sup>852</sup> This is because, as described in Chapter 4, machine learning algorithms that are denied access to particular features, such as ones arising in a protected class, will merely naturally find proxies for those features.<sup>853</sup> Thus, proxy discrimination can only be reduced if the platform designers are aware of both protected variables and potential substitute

---

<sup>849</sup> *See id.*

<sup>850</sup> *See* Chapter 5(B)(5).

<sup>851</sup> *See* Anya Prince & Daniel Schwarcz, *Proxy Discrimination In The Age Of Artificial Intelligence And Big Data*, (SSRN August 5, 2019), IOWA LAW REVIEW, (Forthcoming); Darcy Steeg-Morris, Daniel Schwarcz & Joshua Teitelbaum, *Do Credit Based Insurance Scores Proxy for Income in Predicting Policyholder Risk*, 14 J. EMPIRICAL LEGAL STUD. 397 (2017).

<sup>852</sup> Prince, *Proxy Discrimination*, *supra* note 736, at IIIB(1).

<sup>853</sup> *See supra* Chapter 4, Section A(2) and (3).

variables.<sup>854</sup> Of course, such widespread removal of variables or features comes at a cost, as models will be less effective and more costly.<sup>855</sup> Thus, only the most impactful algorithms should be subject to “restricted from use” lists. However, there is little doubt that any system proposed as an alternative to the courts must be considered as meeting this threshold of impact, and the RoL demands mitigations be taken to control discrimination.

Thus, in this instance, BEST PRACTICE would be: To ensure transparency of the features used to train the model (including but not limited to features that are on the “restricted from use” list, the features discovered and added to the model after training, and the features currently used in the prediction should be readily accessible). AND, to ensure some level of accountability for model designers and institutions for implementing steps to reduce bias and discrimination (including but not limited to restricted-from-use lists).

*c. ADR/ODR Decision Model*

As explained in Chapter 3, ODR models envision a three-part process, (1) negotiation, (2) mediation, (3) arbitration.<sup>856</sup> While the bulk of this thesis has focused on the human decision maker in the third stage, *i.e.* arbitration, it is the three-part process as a whole that must involve safeguards against the reduction of impartiality and dependence. As mentioned in Chapter 3, it is estimated that only 10 to 15% of any potential arbitration cases are actually finalized in arbitration,<sup>857</sup> the final stage of the process. Thus, if the negotiation and mediation stages are designed correctly, the dispute will rarely need to proceed to arbitration. That is not to write that the final stage is not important; instead it should be considered highly important as it serves as a ultimate assurance of compliance with the RoL.

---

<sup>854</sup> *See id.*

<sup>855</sup> *See id.*

<sup>856</sup> *See supra* Chapter 3, Section A.

<sup>857</sup> *See* Colin Rule, *eBay Resolution Center Up for Dutch Innovating Justice Awards - Needs Your Vote!* [MEDIATE.COM](http://MEDIATE.COM), (June 2011).

The model thus, should consider the design of stages one and two, negotiation and/or mediation. And, in this area, as shown in Chapter 3, some amazing AI-based work is already occurring without a human in the process.<sup>858</sup> For example, Smartsettle ONE<sup>859</sup> is a form of blind bidding that incorporates eight patented algorithms. The bidding process uses a bidding bar which “enables the parties to make offers by moving a flag along a horizontal bar comprising a range of numbers from 0 to the amount claimed in the proceedings.”<sup>860</sup> Each party has two flags, one green and one yellow, which represent the current offer and the range the party is willing to accept. While the other party only sees the green flag, the system deploys:

various algorithms that enable the system to “learn” the tactics and priorities of the parties and, with that knowledge, assist the parties into an agreed settlement without at any time revealing their secret offers.<sup>861</sup>

The technology assists the parties in coming to a settlement point through the deployment of a set of machine learning algorithms, with no human involvement.<sup>862</sup> And, while this description is based on assisting parties resolving a single issue, Smartsettle has also developed blind bidding systems for more complex, multi-issue disputes as well.<sup>863</sup>

Technology such as this is not futuristic, it is now being deployed and used.<sup>864</sup> As such, it can and should be deployed within the early stages of the ODR process. The placement of these technologies within the process will assist parties in understanding the true value and expected cost of the damages and provide the parties the opportunity to resolve issues without the need to enter a more adversarial process. However, if the

---

<sup>858</sup> See *supra* Chapter 3, Section A.

<sup>859</sup> Graham Ross, *Smart Settlement: An Opportunity For Mediators*, InfoLaw Newsletter (March 2019) available at <https://www.infolaw.co.uk/newsletter/2019/03/smart-settlement-opportunity-mediators/> (last viewed Aug. 13, 2019).

<sup>860</sup> *Id.*

<sup>861</sup> *Id.*

<sup>862</sup> See *id.*

<sup>863</sup> Known as Smartsettle Infinity. See Ross, *Smart Settlement*, *supra* note 763.

<sup>864</sup> See *supra* Chapter 3, Table D, Table G, Table H. See also Ross, *Smart Settlement*, *supra* note 763.

mediation process fails, arbitration is the perfect final stage as explained in Chapter 2.<sup>865</sup>

The arbitration stage should be considered to include, the institutional structure in which decision makers are considered for inclusion in the dispute process, the conflict/non-appointment process, the appointment process, the process of challenging the appointment, and the auditing of the appointment process and the outcomes.<sup>866</sup> Each will be discussed in greater detail within the risk mitigation section,<sup>867</sup> taking the view that the overall design of this stage is an opportunity to mitigate risk associated with a biased process and/or a partial or dependent decision maker.

## 2. Model Risk in the Decision Making Process

As described in Chapter 4, model risk is “the potential for adverse consequences from decisions based on incorrect or misused model outputs and reports.”<sup>868</sup> In general, model risk in situations involving decision making arises in two key areas:

- (1) A model may have been built as it was intended, but could have fundamental errors and produce inaccurate outputs when compared to its design objective and intended use; or,
- (2) A model may be used incorrectly or inappropriately, or its limitations or assumptions may not be fully understood.<sup>869</sup>

In the particular instance of model risk as it relates to decision making, these two reasons can be alternately described as, (1) model risk associated with the process, and (2) model risk arising from incorrect use or lack of understanding. Each will briefly be explained, with the following sections setting out specific solutions.

### *a. Model Risk Arising from Errors*

---

<sup>865</sup> See *supra* Chapter 2, Section Introduction.

<sup>866</sup> See *id.*

<sup>867</sup> See *infra* Chapter 5, Section C.

<sup>868</sup> See *supra* Chapter 4, Section A(3).

<sup>869</sup> See *id.*

A model may be well designed overall, yet fundamental errors may have been captured in such a way as to produce inaccurate outputs.<sup>870</sup> In these situations, technology built and deployed by the arbitration institutions can assist in monitoring decisions (outcomes) for patterns that suggest the presence of an error in the process.<sup>871</sup> The benefit of this monitoring will emerge, however, only if all the algorithmic operations going into the decision making process are transparent and freely available. For example, technology can be used to weed out inappropriate selections for a human decision maker, through arbitration institutions managing conflict lists. This is a technology-assisted way to make the decision to exclude certain individuals from the institution consideration for the role of arbitrator. This decision of the arbitration institution, to exclude, is one that should be available for review as part of the error discovery process; that much can hardly merit questioning. What level of detail should be made available for audit, however, requires more thought. Some pieces of information will need to be removed, creating a balance crafted by the arbitration institution to protect both the principle of transparency and private information related to the candidates as individuals.

Arbitrator selection, of course, is not the only decision that will need to be reviewed. ALL decisions throughout the process should be reviewed by the arbitration institutions and any patterns that are potentially attributable to the human in the decision loop should be fed back to the human to assist in improving their self-awareness of their own human cognitive bias.<sup>872</sup> Supporting the emotionally intelligent decision maker<sup>873</sup> is one of the best reasons to use technology — it presents the opportunity to find patterns and reveal those patterns to otherwise well-intentioned individuals.

Any identified errors or risks discovered during the auditing process will require careful mitigation of the risk by the arbitration institutions which should always be done with a strict adherence to RoL. For example, it is possible to reduce algorithmic bias through

---

<sup>870</sup> *See id.*

<sup>871</sup> *See supra* Chapter 4, Section C(2).

<sup>872</sup> *See supra* Chapter 4, Section C(4).

<sup>873</sup> *See supra* Chapter 5, Section A(6)(c).



the use of proxies, as described above and discussed more fully in Chapter 4.<sup>874</sup> However, proxies must be carefully monitored so as to not introduce proxy-based bias.<sup>875</sup> Sometimes the technology-driven solutions are not sufficient alone. Any mitigation must be done with the same adherence to best practices as the original design. And, in the case of the human decision maker failing to mitigate risk-associated cognitive errors once they have been identified, it may be the individual will need to be removed from the roster of decision makers by the arbitration institution.

BEST PRACTICES in this area entail arbitration institutions: (1) establishing a transparent process, including the presentation of information about the human decision makers, their conflicts, prior outcomes and party involvement in the appointment process; (2) the ability to audit *all final determinations*, including final award decisions. The auditing information should be used as both a measure to ensure appropriate process and as a mechanism to provide feedback to the human decision maker; (3) the ability to evaluate, identify and mitigate any identified risk, provided the manner of mitigation is in line with RoL and best practices; (4) the creation of a robust, multi-discipline review team to audit the entire process and outcomes.

*b. Incorrect Use Or Lack Of Understanding*

A long line of research has shown that people follow technology with an almost blind belief in its ability to be “right” or “based in math and therefore presumptively absent of bias” and similar misperceptions.<sup>876</sup> As discussed at length in Chapter 4, model designers must be aware of the ways in which technology will have effects upon the individuals using the technology, and even create feedback loops between human and technology.<sup>877</sup> As such, arbitration institution model designers must design a technology-driven system that allows for the individual in the system to understand the basic facts about the technology they interact with. We’ve already mentioned the importance of model cards;<sup>878</sup> in addition, humans must recognize that many cases

---

<sup>874</sup> See *supra* Chapter 4, Section C(2).

<sup>875</sup> See *supra* Chapter 4, Section (2).

<sup>876</sup> See *supra* Chapter 4, Section 3 and corresponding notes.

<sup>877</sup> See *id.*

<sup>878</sup> See Mitchell, *Model Cards*, *supra* note 517 at 22.

require outcomes that are less than “complete” or do not appear as the final product of an equation.<sup>879</sup> As in the current system used to predict weather, the decision makers must be presented with multiple sources of information, all clearly presenting outputs together with key details of the model that produced them. The human decision maker is then allowed and nudged to be attentive to the final selection from the various presentations given to him/her.

BEST PRACTICES in this area, under the assumption that there will be a human in the decision making loop, are: (1) outcomes must be transparent; (2) evaluation and identification of issues within the process, including issues arising from incorrect use, must be identified and the discovery must be part of a feedback loop to the human decision maker, (3) the outcomes, predictions and/or assessments must be presented to the human decision maker in a manner that clearly explains the technological process is not intended to be a complete process, but rather to provide key information and assessments; (4) the outcomes, predictions and/or assessments must be accompanied by a model card including full details of the model and its capacity and design; the outcomes, predictions and/or assessments must be presented in a format that allows for the presentation of multiple model outcomes.

### 3. Risk Mitigation In The Decision Making Eco-System

In this thesis, risk mitigation is a process of developing options and actions to reduce impacts arising from model risk.<sup>880</sup> In general, engineers needing to handle model risk turn to several common risk mitigation options, including:

- (1) Assume/Accept the risk by acknowledging the existence of a particular risk, and making a deliberate decision to accept it without engaging in special efforts to control it.
- (2) Avoid the risk by adjust program requirements or constraints to eliminate or reduce the risk.
- (3) Control by implementing actions to minimize the impact or likelihood of the risk.
- (4) Transfer by reassigning organizational accountability,

---

<sup>879</sup> See *supra* Chapter 4, Section 3.

<sup>880</sup> This should be contrasted with other areas definition of risk mitigations. For example, systems engineers define “risk mitigation planning is the process of developing options and actions to enhance opportunities and reduce threats to project objectives.” Project Management Institute, *A Guide to the Project Management Body of Knowledge*, (PMBOK Guide), Fourth Edition, ANSI/PMI 99-001-2008, pp. 273-312 (2008). See also, *supra* Chapter 4, Section 3.

responsibility, and authority to another stakeholder willing to accept the risk. And, finally (5) Watch/Monitor.<sup>881</sup>

The design strategies allow engineers to manage risk and to consider the solution to the identified risk. For example, if model design engineers are aware that the machine learning component is being trained on data that is incomplete or reflective of discrimination, engineers can seek to mitigate against the negative impacts from that data.<sup>882</sup> They have many options for mitigating the risk, some technology-driven, some located in the socio-technical design;<sup>883</sup> the decision of which one to deploy should be guided by the stakeholder choices as they reflect upon the suggestions above.<sup>884</sup> In the case of incomplete or skewed training data, the stakeholders may decide that the appropriate mitigation is to monitor the system, to determine if actual impact is occurring.<sup>885</sup> Or they may decide the appropriate mitigation is to accept the risk from the use of incomplete or skewed data and to acknowledge that limitation when reporting outputs to human decision makers.<sup>886</sup>

The author's suggested model (below) is built with an eye toward identified risk and mitigation principles. As such, each of these strategies will be deployed in the decision making model to mitigate negative impacts to the fullest extent possible. It is important to note, first, that the author has focused on upholding the RoL, and second that the Rules of Action for this model were created with an understanding of the risk associated with data. In this particular instance, data that is incomplete, skewed and/or disproportionately impactful upon those captured in the justice data eco-system simply cannot be mitigated against in a manner that eliminates the negatively impactful effects. This incredibly large issue coupled with the inability of the model to capture context and discretion has led to the author insisting upon a human in the decision making loop. This is an example of an avoidance mitigation, that is eliminating the potential risk by

---

<sup>881</sup> MITREA Author, *Implementation, and Progress Monitoring*, MITRE's Systems Engineering Guide, available at <https://www.mitre.org/publications/systems-engineering-guide/acquisition-systems-engineering/risk-management/risk-mitigation-planning-implementation-and-progress-monitoring> (last viewed Aug. 13, 2019).

<sup>882</sup> *See id.*

<sup>883</sup> Selbst, *Fairness and Abstraction in Sociotechnical System*, *supra* 503 at 59.

<sup>884</sup> *See* MITREA, *supra* note 881.

<sup>885</sup> *See id.*

<sup>886</sup> *See id.*

avoiding it. In the case of a decision maker in a justice environment, this is simply not open for any other mitigation, because the consequences to the commitment to the RoL are too great and the importance of creating a trustworthy system with a high level of integrity is too important to embrace any other option. As such, the model is built with a human in the decision making, with technology that can augment, assist and facilitate their process. But ultimately, the human makes the final decision. Of course, the human is also a source of risk. Yet, in an innovation upon the existing justice environment, the system is designed to deploy technology in a manner which reduces the impact of the human-in-the-loop bias. With these two important key points again emphasized, the remaining sections seek to address risk issues arising in the decision eco-system via a mitigation process.

*a. Managing Model Risk Arising within the Decision Process*

Model design has to be carried out in a manner which accurately accounts for traditional technology design biases.<sup>887</sup> In the alternative justice environment the main issue is that the data that is presumptively to be used to build algorithms will come from the existing justice environment. The age-old axiom states “garbage in, garbage out.”<sup>888</sup> Hence, as described in Chapter 4, model designers must recognize the limitations of their starting data sets.<sup>889</sup> One potential solution to “bad data” or the lack of data is to create a knowledge commons focused on the law. Knowledge commons refers to cultural and intellectual resources, including data and information, that are created, shared, and passed to future generations through experiences, education, science and interactions.<sup>890</sup> Through the use of a knowledge commons framework, data can be gathered, shared and sustained with a common best practices approach.<sup>891</sup> Knowledge commons have been used in other, similar areas such as open educational resources, free digital media such as Wikipedia, and open scientific collections such as the Public

---

<sup>887</sup> See also, *supra* Chapter 4, Section 3.

<sup>888</sup> Carolyn Healey, *Garbage In, Garbage Out – How Bad Data Hurts Your Business*, Service Objects Website, (Oct. 9, 2018).

<sup>889</sup> See *supra* Chapter 4, Section 3.

<sup>890</sup> Michael Veale, *Can People Trust the Automated Decisions Made by Algorithms?*, Info Q, (June 21, 2018).

<sup>891</sup> See e.g. Brett M. Frischmann, Michael J. Madison, Katherine J. Strandburg, *GOVERNING KNOWLEDGE COMMONS*, Oxford University Press; (September 2, 2014)

Library of Science.<sup>892</sup> Open frameworks such as these encourage open research environments focused on the processes of information research, gathering and production.<sup>893</sup> Criticism can be leveled against these environments, but in the case of building a justice environment encouraging broader diversity and cultural coverage in data sets is a priority, so the potential benefits are expansive. Professional access to process data should be provided, with the objective of minimizing discriminatory effects.

In addition, as discussed in Chapter 4,<sup>894</sup> technology deployed in the justice environment should be held to certain standards of transparency and subjected to auditing of outcomes for potential RoL violations:

Computer systems can be designed to prove to oversight authorities and the public that decisions were made under an announced set of rules consistently applied in each case, a condition we call *procedural regularity*.<sup>895</sup>

As explained in Chapter 1, RoL demands that the rules are consistently applied and that outcomes are consistent across types of cases and/or issues.<sup>896</sup> Arbitration institutions auditing of the system, or system oversight, provides such protections and assurances. No ODR system or decision maker should be allowed to operate outside the expectation of procedural regularity. Thus, BEST PRACTICES include: (a) outcomes must be transparent, (b) compliance with rules and process must be audited, and (c) case decisions must be measured against other outcomes to ensure a high level of predictability and the absence of bias in the system. And, as described in Chapter 1, this does not prohibit the inclusion of a human decision maker who exercises discretion; in fact as described in Chapter 1, it is important to understand that the exercise of discretion, while essential, is also bounded. Thus, the system should allow discretion, but should audit the process and produce transparent outcomes so discretion is part of the system within the correct bounds.

---

<sup>892</sup> *See id.*

<sup>893</sup> *See id.*

<sup>894</sup> *See supra* Chapter 4, Section C.

<sup>895</sup> Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, *Accountable Algorithims*, UNIVERSITY OF PENNSYLVANIA LAW REVIEW, Vol. 165: 633, 637, (2017).

<sup>896</sup> *See supra* Chapter 1, Section A.

*b. Managing Model Risk in the Decision Maker*

The decision making portion is the most important aspect of the model for the purposes of the thesis. This section attempts to address risk associated with the process of adjudicating a final decision. This will require arbitration institutions to create mitigation activity as it relates to the creation of the roster of decision makers, the appointment, the decision process itself, and the human interaction with the technology (i.e., socio-technical aspects). The large sections will explore and suggest appropriate design as it relates to: (1) the role of institutions and/or platform designer; (2) the roster of decision makers; (3) nomination and appointment of the human decision maker; (4) challenges to the appointment; (5) transparency of the appointment process and the humans in the process, including: (a) risk arising from the non-disclosure/rule compliance; (b) risk arising from implicit/hidden bias of the decision maker; (c) user risk; (d) risk not already considered arising from the introduction of discretion. It will conclude by introducing the concept of risk mitigation via the use of impact assessments.

*i. Managing Risk by the Introduction of a Human in the Decision Making Process*

The introduction of a human into the decision making process is an obvious potential risk as it reintroduces the partiality and dependence issues that technology was intended to reduce. However, as was discussed above and in Chapter 4, there is currently no historic data set from a justice environment that does not contain the historic biases and prejudices that are part of our complex history.<sup>897</sup> Without such a data set, machine learning algorithms will be trained on historically biased data, creating a further engraining of historic prejudice, but covering it under a sea of trust in technology and math.<sup>898</sup> As such, it is difficult for this author to support a machine learning based system as an advancement in the delivery of justice.<sup>899</sup> Without a doubt, we can design a model that can accurately predict the outcomes of a narrow band of cases, such as

---

<sup>897</sup> See *supra* Chapter 4, Section A (2) and (3).

<sup>898</sup> See *id.*

<sup>899</sup> See *supra* Chapter 4, Section C(4).

failure to deliver a parcel.<sup>900</sup> There is little doubt in the author's mind that we could design a model that predicts outcomes based on the assignment of a case to a particular court or judge. These models do nothing more than predict the prejudice that already exists in the system, something this author is unwilling to embrace. In fact, as of June 2019, France has passed a law making it illegal to build models based on the outcomes of particular judges, so-called judge analytics.<sup>901</sup> While this may seem an extreme response, it is one that arises from the fear associated with further engraining historic prejudices that permeate the justice system. As such, the author desired to expect more from a newly created system and to deploy technology *only* if it can improve the existing system. In this case, it means we need a human in the decision making loop, but we must monitor the human and the outcomes to mitigate against partiality and bias. In this way, we deploy technology to do better.

#### (1) The Role of Institutions and/or Platform Designer

Institutions play an important part of risk mitigation in protecting the ODR process because they are in a position to create institutional rules, to insist upon a robust commitment to the essential aspects of RoL, and to monitor and enforce rules related to the nomination, conflict declarations and behavior of decision makers. While some commentators insist the ODR process can exist outside of institutions, the author believes that a commitment to RoL demands some type of central authority that protects both the process and guarantees unbiased delivery of justice.<sup>902</sup> It is not essential that the authority mirror or be named like the institutions of traditional arbitration; rather the essential point is that some authority must be responsible for protecting the integrity of the alternative justice environment — and key to this is a commitment to the RoL.<sup>903</sup>

Institutions, as well as platform creators/designers, must create and uphold rules of process and appointment.<sup>904</sup> And all of the rules relating to process, appointment,

---

<sup>900</sup> In fact, smart contracts are capable of that very thing.

<sup>901</sup> Author, *France Bans Judge Analytics, 5 Years In Prison For Rule Breakers*, ARTIFICIAL LAWYER (June 4, 2019).

<sup>902</sup> *See supra* Chapter 4, Section C(3).

<sup>903</sup> *See id.*

<sup>904</sup> *See id.*

auditing, transparency and the like must be publicly available to everyone.<sup>905</sup> Individuals must be fully informed of process and rules in user-friendly language. Full disclosure allows parties to make determinations as to the appropriateness of a process and rules for their situation and issue. As discussed in Chapter 2, arbitration is built upon foundations of contract and voluntary consent,<sup>906</sup> neither of which can be protected if institutions/platforms are allowed to create a black box environment.

Enacting its commitment to ensuring neutral decision makers, the institution/platform designer must insist upon a system of required disclosures, conflict rules, mandatory training, and must insist on a strong enforcement mechanism that refuses appointment to anyone who is not compliant.<sup>907</sup> Moreover, to assist in the reduction of certain types of bias that will be discussed in more detail later, the institution must also create rules that require a specific type of training for any decision maker that will use the system.<sup>908</sup> Training in the process and the technical aspects of the system, as well as understanding of the impact that technology-driven prediction systems can have on human decision makers, is essential.

In addition, further consideration should be given to the creation of specific legal requirements for the institution/platform designer,<sup>909</sup> including the possibility of strong penalties or individual liability for failure to comply with cyber security laws, disclosure laws, and the protection of RoL principles, such as the presence of a neutral decision maker.

## (2) The Roster of Decision Makers

The creation by arbitration institutions of institutional rosters of those available for selection within the resolution process is also an important step in the reduction of partiality and dependence.<sup>910</sup> As previously described, institutions should require

---

<sup>905</sup> *See id.*

<sup>906</sup> *See supra* Chapter 2, Section Introduction.

<sup>907</sup> *See supra* Chapter 4, Section C(3).

<sup>908</sup> *See supra* Chapter 4, Section C(4).

<sup>909</sup> *See supra* Chapter 4, Section C(3).

<sup>910</sup> *See supra* Chapter 4, Section C(4).



decision makers to create a public and a private disclosure page.<sup>911</sup> The public page should contain any information that could create an appearance of bias relative to the types of cases that are heard by the platform.<sup>912</sup> For example, being a co-owner in a large transport firm creates the appearance of bias if the platform handles transport-related disputes, so it must be publicly disclosed. It might, however, be appropriate for other disclosures to remain private if they could only be relevant to partiality under a narrow and unlikely set of circumstances.<sup>913</sup> For example, if I occasionally drive for Uber and the platform handles family disputes, my interest in the privacy of my information likely outweighs the requirement of transparency.

A well-defined set of institutional rules that create clear disclosure requirements could reduce bias and thus the number of challenges to the participation of individual decision makers. Moreover, further reductions can be made when the public and the private disclosure lists are used in tandem to create an “excluded from participating in this decision” list. The list would be used to exclude individuals on said roster from participating in a particular decision.

The technology can also accommodate a weighting system, such that individual factors and certain clusters of factors can be weighted more heavily. For example, a 2019 ICSID case asked for a determination on the dependence of a decision maker.<sup>914</sup> Using the applicable UNCITRAL Arbitration Rules the institution enumerated a four-part test to assess whether a connection between the arbitrator and one of the parties was likely to impact the arbitrator’s impartiality and independence.<sup>915</sup> These factors are (1) the proximity of the connection, (2) the intensity of the interactions, (3) the arbitrator’s financial dependence on the party, and (4) the materiality of any benefits derived.<sup>916</sup> Technology could be used to identify such a cluster of factors, or any cluster of recorded activities or behaviors that reasonably lead to the conclusion that a financial dependency exists.

---

<sup>911</sup> See *supra* Chapter 5, Section A.

<sup>912</sup> Similar to disclosures in arbitration. See *supra* Chapter 2, Table C.

<sup>913</sup> See *supra* Chapter 2, Section B(1).

<sup>914</sup> See ICSID Case No. ARB/03/19 (2019).

<sup>915</sup> See *id.*

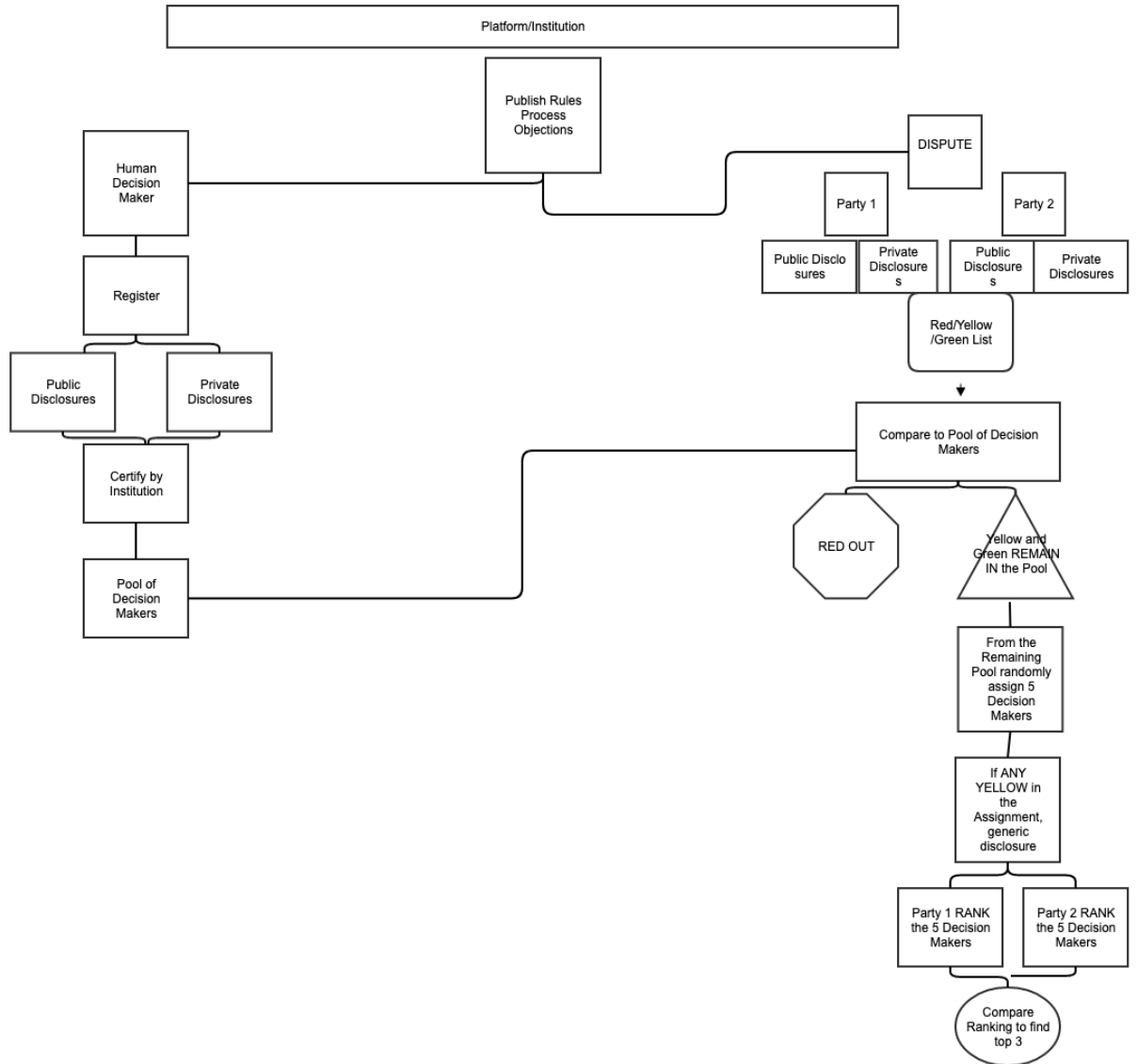
<sup>916</sup> See *id.*

Once a “non-excluded” list is created, selection of the five decision makers presented to the parties could be done by random sample. The five randomly selected potential decision makers are presented to each of the parties; who in turn are asked to rank them, with access to the public disclosure list. This check and balance allows the parties to serve as a secondary mechanism to preclude the presence or appearance of bias.

(3). Nomination and Appointment of the Human Decision Maker

The portion of the model in which the decision maker is selected (from the roster of arbitration institutionally approved decision makers) can assist in greatly reducing dependence and partiality. The following schematic/workflow is a visualization of how the appointment process could occur within a smart contract, after a breach has occurred and mediation/negotiations have failed:

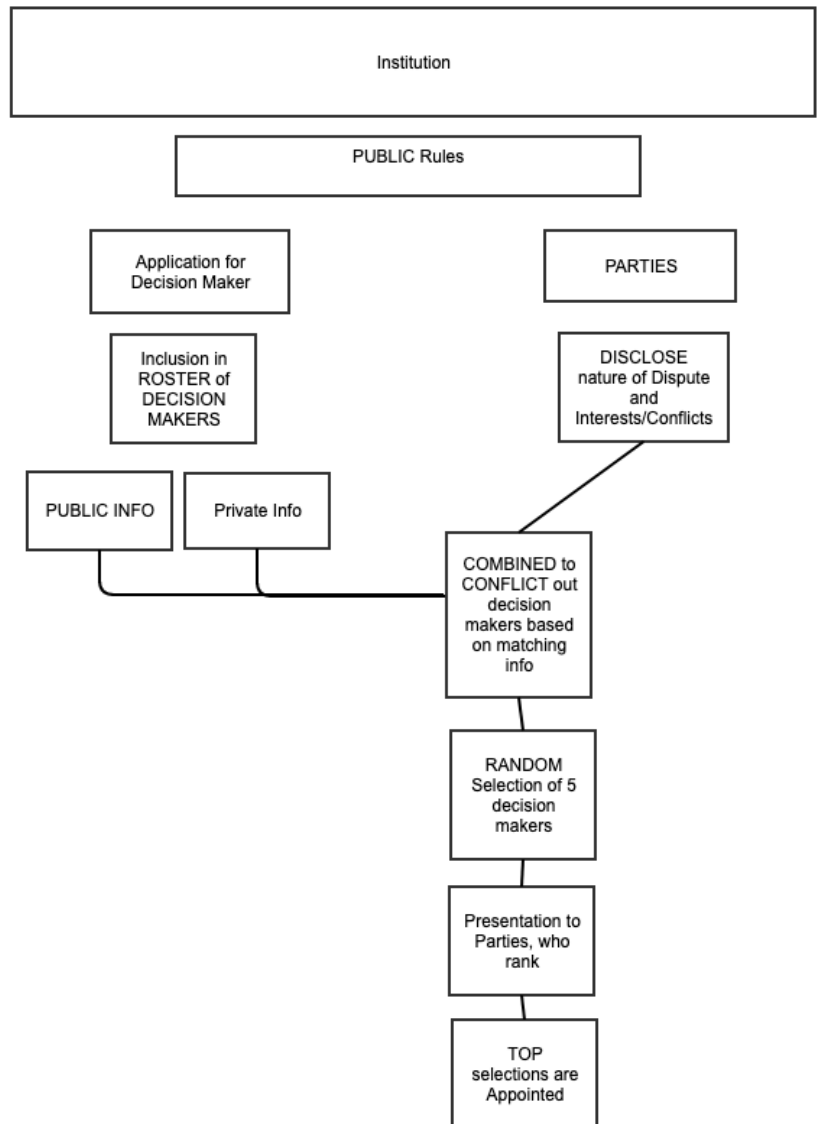
*Schematic 2: The Suggested full Process*



Design created and schematic created by Raymond (2019)

As can be seen, the workflow begins after negotiations have failed. In this instances, the parties progress into an online arbitration process, which envisions the use of a human decision maker. Key to this model is the appointment process.

*Schematic 3: Appointment*

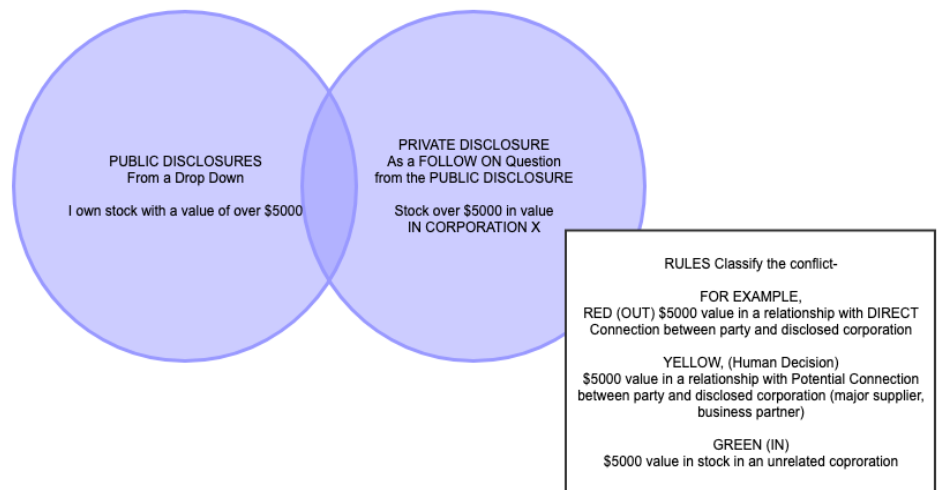


Design created and schematic created by Raymond (2019)

As we covered in an earlier chapter, dependency is relatively easy to identify compared to partiality, but this is nonetheless an area where AI can provide improvements. In this instance, the arbitral institution will have a roster of human decision makers to select from. Having an institution supervise and create a roster of human decision makers ensures compliance with the RoL as the institution can ensure the decision makers have

the appropriate qualifications.<sup>917</sup> Moreover, the institution can create rules of disclosure that also protect the RoL, as the decision maker should be ensured to be independent from the parties and from the potential outcome.<sup>918</sup> The institution should ensure this by creating public informational pages about each decision maker on the roster, in addition to maintaining a private database of more detailed information about decision makers that is not widely released but can be consulted when appropriate. The institution can then use an algorithm to do a comparison between the issues and parties in the dispute and all available decision makers. ANY decision maker with either a publicly revealed conflict or a private issue that gives rise to a conflict is conflicted out of participating. This process ensures that the RoL expectation of an independent decision maker is protected as fully as possible.

*Schematic 4: Disclosures*



Design created and schematic created by Raymond (2019)

As previously described, after the institution generates the list of all human decision makers who have not been conflicted out, a new algorithm could make a random selection of decision makers and create a list of five options. Similarly to the process laid out in the AAA rules and the JAMS rules presented in Chapter 2,<sup>919</sup> the parties

<sup>917</sup> See *supra* Chapter 1, Section C.

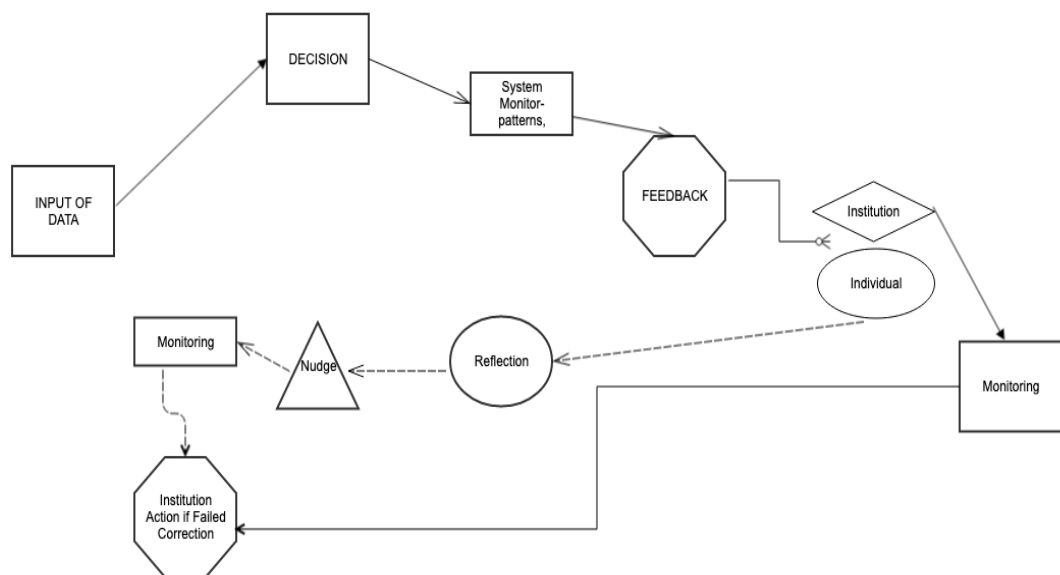
<sup>918</sup> See *id.*

<sup>919</sup> See *supra* Chapter 2, Section B.

would then receive the five selections and be asked to rank the potential decision makers based on personal preference. It is important to note, the human decision makers will have a public information/conflict declaration page, and the parties can (and should) make determinations based upon this information. This portion of the process can greatly reduce the impact of dependence.

The reduction of potential partiality can also be impacted at this stage as well, as the arbitrator roster has both a public and a private portion. Within the private portion of the system could be the results of prior decisions made by the human decision makers, and key markers/features about each case.

*Schematic 5: Auditing Outcomes and the Feedback loop*



An algorithm could be used to conflict out individuals who display particular patterns of outcomes that suggest partiality. It is important to note, the author does NOT think this information should be released to the public or contained within the public disclosures. However, noting patterns in outcomes is an important step in assisting individuals to be emotionally intelligent decision makers. The ability to take on board feedback and consider whether personal bias is impacting decision making is an important attribute of an emotionally intelligent decision maker as described by Justice

Sotomayor. To encourage self-reflection, this level of feedback should be private, not used as an incentive or a punishment. Instead, the platform can encourage self-reflection and can feedback suggestions in instances where similar self-reflection should be noted by the decision maker. The patterns should, also be monitored and recorded by the platform, as it is the platform's obligation to adhere to the RoL, including the expectation of protecting parties from a partial or dependent decision maker.<sup>920</sup> Moreover, there are other methods for publicly ranking decision maker behavior: as described in Chapters 2 and 3,<sup>921</sup> systems are currently being created to "rank" arbitrators based on parties' satisfaction with the arbitrator. This type information can and should be made public as it assists in the transparency of the process and the selection.

Allowing the parties to obtain public information, such as qualifications, conflict declarations and satisfaction survey outcomes, is an additional step that evidences a strong commitment to several RoL principles, including the commitment to having safeguards in the appointment process.<sup>922</sup>

The ranking of each party would then be used to create a final selection of the decision maker for the particular case. The process described in not necessarily new, *per se*. Rather, it combines several approaches to appointment described in Chapter 2<sup>923</sup> and some of the newest technology deployments described in Chapter 3.<sup>924</sup> It does ensure that the RoL principles relating to the decision maker that were emphasized in Chapter 1, such as appropriate qualifications, independence, impartiality and protections from conflicts of interest,<sup>925</sup> are enshrined within the ODR appointment process.

#### (4). Challenges to the Appointment

---

<sup>920</sup> See *supra* Chapter 1, Sections C and D.

<sup>921</sup> See *supra* Chapter 2, Section A and Chapter 3, Section B.

<sup>922</sup> See *supra* Chapter 1, Section C and Section D.

<sup>923</sup> See *supra* Chapter 2, Section B.

<sup>924</sup> See *supra* Chapter 1, Section C.

<sup>925</sup> See *supra* Chapter 3, Section C.

Both parties must have the right to challenge the selection of any decision maker, and this is true regardless of the process of selection.<sup>926</sup> While this may seem overly stringent or time-consuming, it is important to recall that the standard is “appearance of bias,” as presented in Chapter 1,<sup>927</sup> and as such parties must be allowed to ask for decision makers to be eliminated from selection. However, a balance can be struck between the right of contestation and the efficiency of the process by setting bounds on the length of time for making a challenge and reasons allowed for a party to challenge. Remember, in this process parties begin by ranking potential decision makers, which should greatly limit the likelihood of a challenge. It is hard to imagine a scenario where a party ranks a decision maker last, but the selection still proceeds. As such, only new information, or an objection of the highest order based on the initial information, should allow a party to challenge an appointment. To limit procedural delays, the challenge process must take place in a predetermined delimited timeframe and not allowed to hold up the entire process. Moreover, the challenging party must specify the reason for the challenge, which must meet the standard of casting doubt “in the eyes of a reasonable person.”<sup>928</sup> These conditions safeguard the essential principle that any appearance of partiality or dependence must be eliminated, without making the process easy to for either party to delay or derail.

#### (5). Transparency of the Appointment Process and the Humans in the Process

During the above described process, transparency is not handled in the manner discussed in Chapter 2, that being the presumptive “disclose everything,” but is instead limited to disclosures made for a specific purpose.<sup>929</sup> For example, the institutional rules, the process, and key information about the institution, including the use of technology, should be fully disclosed and completely transparent. Of course, as discussed in Chapter 2, this is already the case in the majority of, if not all, arbitration institutions.<sup>930</sup> Moreover, as discussed above at Chapter 4, transparency requires any automated or other technology that is involved within the decision making process to

---

<sup>926</sup> See *supra* Chapter 2, Section B.

<sup>927</sup> See *supra* Chapter 1, Section D.

<sup>928</sup> See *supra* Chapter 1, Section D.

<sup>929</sup> See *supra* Chapter 2, Section B(1).

<sup>930</sup> See *id.*



be fully described in a publicly available environment, via the use of model cards and other public disclosures.<sup>931</sup>

As it relates to the neutral decision maker, the balanced release of some information concerning the human decision maker is essential to maintaining the protections of the RoL.<sup>932</sup> However, in this instance, because it is a human at the heart of the process, transparency must be balanced against the right of an individual to not have his/her whole life disclosed. For example, a potential arbitrator should not necessarily be prohibited from hearing issues involving cases that tangentially touch upon stock in her retirement portfolio. Requiring over-disclosure of personal information could deter qualified individuals from serving as decision makers, shrinking the pool, or be used to game the system by eliminating a large group of qualified decision makers on frivolous grounds. As such, as discussed in Chapters 1 and 2, the RoL is protected by publicly releasing information that is deemed to be more likely than not considered by a reasonable person to create an appearance of bias.<sup>933</sup> However, information that may create the appearance of bias only in narrow and specific cases should be protected from full disclosure. These disclosures are then processed in two distinct ways, maximizing their value for identifying possible partiality or bias: (1) they inform the parties and therefore allow humans to level a reasonable objection, and (2) they feed into a conflict list managed by the institution that can identify possible problems through automation.

In situations related to potential conflicts of interest, the system can remain compliant with the RoL, while still encouraging participation and preventing crafty challenges to the decision makers by greatly reducing the number of times the dilemma arises and by creating rules to reduce the impact on all concerned. The RoL demands transparency and disclosure of the mechanism of making the decision and the rules that will be used should be transparent.<sup>934</sup> RoL demands the system work to reduce the impacts, which in this case can be done by a well-defined process, clear rules, ranking by the parties,

---

<sup>931</sup> See *supra* Chapter 4, Section C(3).

<sup>932</sup> See *supra* Chapter 1, Section C and Section D.

<sup>933</sup> See *supra* Chapter 1, Section C and Section D and Chapter 2, Section A.

<sup>934</sup> See *supra* Chapter 1, Section A.

and random final selection.<sup>935</sup> The definitions, the conflict rules and the criteria used to make selection determinations should be disclosed and widely available on the platform and any corresponding websites.

When a decision is made to exclude, the reasons for that decision should be made available, modified as necessary to prevent unnecessary disclosure of private details. This type of transparency allows both the rules and the outcomes to be reviewed.<sup>936</sup> Thus, when a decision maker is excluded from selection, both parties as well as researchers should be able to request a list of all (de-identified) individuals or platforms excluded from selection along with information surrounding the circumstances for the non-selection. However, to encourage potential decision makers to feel comfortable making the most capacious possible disclosure without fear, this report should not disclose the specific names of those prevented from participating, but should instead create a disclosure list based on the stated issue that caused the exclusion.

And, of course, the final decision, the basic information about the parties, the length of the dispute, number of submissions, or documents and similar details should all be disclosed publicly, as should the name of the human decision maker and the final outcome. Transparency of outcome is the only true way the system should ever be trusted to be compliant with the RoL.<sup>937</sup> Thus, there should be no burdensome requirements placed upon requesting such information, nor exceptions allowed.

i. Risk Arising from the NON Disclosure/Rule Compliance

It is assumed that the majority of individual decision makers and platform designers/owners will be rule compliant, especially as it relates to disclosures on questions of independence and impartiality. However, it is important to assume, and protect against, the possibility that a small minority of institutions, platforms designers and human decision makers will be non-compliant; thus monitoring and verification are essential.

---

<sup>935</sup> See *supra* Chapter 1, Section C and Section D.

<sup>936</sup> See *supra* Chapter 4, Section C.

<sup>937</sup> See *supra* Chapter 1, Section C and Section D.

Fortunately, this is one area where technology can be of great assistance. Technology is already being used in multiple settings to verify key pieces of information. For example, several well-known services are used to verify information contained in resumes using commonly available public data. In a similar fashion, in the justice environment, decision makers' public and private profiles could be matched against information provided in other settings to verify the information disclosed. The institution could insist upon such verification as a term of joining the roster of available decision makers and could remove individuals from the roster should discrepancies be discovered.

ii. Risk Arising from Implicit/Hidden Bias of the Decision Maker

Despite all best efforts and risk reduction techniques, model designers must be aware that the introduction of a human into the decision making process may re-introduce bias that had been previously eliminated. As such, risk arising from human implicit bias must be monitored. In general, there are two mechanisms that could be deployed to reduce implicit/hidden human bias. As we have covered, the first is to be transparent and to publish all decisions, not just outcomes but all decisions along the way. The second is to monitor and audit all decisions for patterns of bias. To maximize self-reflection, learning and the ability to correct partiality, the author recommends that the details of such audits be kept private when they implicate an individual decision maker, so that individual can be given an opportunity to learn from the feedback and adjust future behavior.

There is a third potential use of technology, that being the use of technology designed specifically to measure human reactions and behaviors for signals of deceit and discomfort. I have previously written about this topic in *The Angel on Your Shoulder: Prompting Employees to Do the Right Thing Through the Use of Wearables*,<sup>938</sup> in which we suggest that technology currently exists which could measure individuals' heart rates and eye movements, use this data to identify times when the individual is engaged

---

<sup>938</sup> Fort, Raymond, and Shackelford, *The Angel on Your Shoulder supra* note 685, at 139-170.

in deception or stressed by a moral dilemma, and could nudge people into becoming aware of their dilemma and suggest possible moral frameworks for resolving it.<sup>939</sup> Researchers are already using this process to detect deception in security clearance situations.<sup>940</sup> By applying it in a user-directed context, it could instead be deployed to increase self-awareness and growth. While the author is not yet supportive of the use of such technology within this particular setting, further research could be done to consider if deployments such as this would be beneficial to the human decision maker within the justice environment.

In addition, there is the possibility that bias will arise out of the human decision maker's attitude toward the technology or faulty perceptions of the technology, the manner in which technology is deployed, the way in which information is displayed and all other kinds of bias that arise in human–technology interactions.<sup>941</sup> One such concern would be that a human might fail to fully review recommendations in a meaningful and active manner.<sup>942</sup> However, these are bias patterns that can be discovered, evaluated and mitigated through automated monitoring. In situations such as this, several of the already discussed mitigations, such as training and monitoring of human decision makers, is essential.

### iii. Managing User Risk

Of course, the newest research is exploring in greater detail how technology impacts human experience. For example, it is reasonably clear that individuals are now interacting with their world in a different manner than in the past based upon the cell

---

<sup>939</sup> *See id.*

<sup>940</sup> *See* Marisa Taylor – McClatchy, *FBI Turns Away Many Applicants Who Fail Lie-Detector Tests*, McClatchy Washington Bureau, (May 20, 2013).

<sup>941</sup> Known as automation-induced complacency and lack of interpretability. *See* Information Commissioners Office, *Automated Decision Making: The Role Of Meaningful Human Reviews*, AI Auditing Framework Blogspot, (2019) available at <https://ai-auditingframework.blogspot.com/2019/04/automated-decision-making-role-of.html> (last visited Aug. 13, 2019).

<sup>942</sup> A concern that has been raised, within E.U. as “human input must be meaningful, and individuals must have the authority and competence to challenge the decision.” Moreover, “[t]o qualify as human involvement, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture.” Center For Data Innovation, *Comments To The Article 29 Working Party*, *supra* note 691 at 3.

phone in their pocket. For example, few individuals use a paper map and are now more likely to follow paths suggested by their technology. The ability of technology to nudge us to behave in certain ways, through advice, prediction and other means, is of growing concern in many areas.

As previously detailed, Green and Yen performed a study designed to review how risk assessments influence human decisions within criminal justice adjudication<sup>943</sup> and found that participants made decisions that were less accurate than the advice provided. Despite being presented with the technology-driven predictions,<sup>944</sup> participants could not effectively evaluate the accuracy of their own or the risk assessment's predictions,<sup>945</sup> and participant interactions with the risk assessment tool introduced new forms of bias into decision making.<sup>946</sup> This new form of bias is very concerning as the risk assessment itself was able to influence the behavior of the participants.<sup>947</sup> In this way, the presence of technology may be altering the way individuals view the world and their interactions with the world. For example, Frischmann and Selinger in their book *Re-Engineering Humanity* argues “we shape our tools, fall in love with them, and thereafter, our tools shape us.”<sup>948</sup> Consequently, it is imperative to monitor the manner in which the humans within the environment are reacting, adjusting and internalizing the interactions, nudges, and outputs that arise from the system.

Until more direct research can be done, all systems must be designed in a manner that clearly presents the full picture of the outcome produced. For example, if the machine is predicting that a given event has caused a given amount in damages, alongside that prediction the system must also present HOW it arrived at that decision (e.g., data used, weights given), and a “confidence” or accuracy score. The information displayed, the manner of display, the human interaction with the system and other key metrics should

---

<sup>943</sup> See Green, *Disparate Interaction*, *supra* note 507, at 90.

<sup>944</sup> See *id.*

<sup>945</sup> See *id.*

<sup>946</sup> See *id.*

<sup>947</sup> See *id.*

<sup>948</sup> Brett Frischmann and Evan Selinger, *Re-Engineering Humanity*, CAMBRIDGE UNIVERSITY PRESS, (2018) at 52, *relying upon* John Culkin, *A Schoolman's Guide to Marshall McLuhan*, THE SATURDAY REVIEW (March 18,1967).

also be kept and analyzed to evaluate if the manner or content of this is in fact influencing decision making.

In addition, there is growing evidence to suggest that as part of the outcome monitoring process individuals should be required to explain any departures they make from the technology's recommendations. The requirement to explain departure from the technology should be used both to monitor the decision maker and the machine learning system for new forms of bias. As Green and Yen explain, "disparate interactions, whereby their use of risk assessments leads to higher risk predictions about black defendants and lower risk predictions about white defendants"<sup>949</sup> is a phenomenon that occurs because of the introduction of a human in the decision making process. Ongoing research is needed to further understand how individuals process and interact with technology within a prediction-based environment, and, because of the high impact, this is especially true of the human decision maker within the ODR environment. Limited research exists in this very narrow area. Any system created for use in a justice environment should be required to participate in research on human-computer interaction issues.

#### iv. Risk not already considered arising from the introduction of discretion

As discussed in Chapter 1 and reviewed earlier in Chapter 5, discretion is an important aspect of the RoL and our justice environment.<sup>950</sup> RoL, however, demands that discretion be bounded by rules for when it is allowed and how much discretion is permissible.<sup>951</sup> Yet, as Green and Chen argue,

Our results highlight a significant but often overlooked aspect of algorithmic decision-making aids: introducing risk assessments to the criminal justice system does not eliminate discretion to create "objective" judgments, as many have argued. . . Instead, risk assessments merely shift discretion to different places, which include the judge's interpretation of the assessment and decision about how strongly to rely on it.<sup>952</sup>

---

<sup>949</sup> Green, *Disparate Interaction*, *supra* note 507, at 90.

<sup>950</sup> See *supra* Chapter 1, Section A and Chapter 5, Section 2.

<sup>951</sup> See *supra* Chapter 1, Section A.

<sup>952</sup> Green, *Disparate Interaction*, *supra* note 507, at 96.

The introduction of a human into the decision making loop does not eliminate discretion to create an objective outcome; instead the human using technology likely shifts discretion into new areas.

It follows that system design must allow discretion in well-defined areas, with bounded choices for its exercise. Monitoring and auditing of outcomes and interactions with the technology will hopefully reveal areas in which new processes must be deployed. For example, when choosing to depart from the technology's recommendation, decision makers could be limited to a pre-proscribed list of allowable reasons. Moreover, while machine learning has the ability to update its own processes over time, the technology can be blocked from making automatic algorithmic adjustments in these key areas. In this way, rather than learning on its own, the technology includes the human decision maker and the system designer in its feedback loop.

#### 4. Risk Mitigation: Impact Assessments

Impact assessments are nothing new; in fact they have been implemented in scientific and policy domains from environmental protection to privacy. They should be considered in the area of AI, especially in situations where the government is involved the deployment of the technology, as is the case in several contexts. AI impact assessments are already being developed in a wide variety of environments, including governments.

As Special Advisor on Artificial Intelligence–Policy and Implementation Noel Corriveau explains,

an impact assessment is a tool used for the analysis of possible consequences of an initiative with a view to provide recommendations as to how to deploy the initiative and under what conditions.<sup>953</sup>

---

<sup>953</sup> Noel Corriveau, *The Government of Canada's Algorithmic Impact Assessment: Towards Safer and More Responsible AI*, AI for Good Social GitHub publication, (2018) available at [https://aiforsocialgood.github.io/2018/pdfs/track2/83\\_aisg\\_neurips2018.pdf](https://aiforsocialgood.github.io/2018/pdfs/track2/83_aisg_neurips2018.pdf) (last viewed Aug. 14, 2019).

While the exact form and questions of the assessment vary across the various promulgated developers, most contain several key focus areas: (1) develop the description of the project: the goals that are pursued by using AI, the data that are used, the actors involved, such as the end users and other stakeholders; (2) formulate project goals, not only at the level of the end user, who experiences the consequences of the service, but also at the level of the organization offering the service and of the society; (3) discover, explore and consider the relevant ethical and legal frameworks to be mapped onto the framework of the application; (4) make strategic and operational choices that consider and carry out activities in relation to all stakeholders; (5) apply different ethical and legal considerations to define the permissible deployment of AI; (6) document the previous steps and justify all decisions taken; and finally, (8) monitor and evaluate the impact of AI.<sup>954</sup>

The importance of the impact assessments cannot be overstated. Algorithms are being deployed in a multitude of environments, with both positive and negative consequences. In the case of algorithm deployment in a justice-based environment, the need to monitor, evaluate and analyze technology is especially important as negative impacts can be so great that RoL principles are affected. Impacts such as this not only harm individuals, but can be felt by the justice system as an institution. Impact assessment must become part of an ongoing commitment to the responsible deployment of technology, especially algorithms, in the justice environment.

Despite the growing demand for increasing use of impact assessments, impacts are often still an afterthought in terms of design and deployment. Consider the very recent example, in which bias was discovered in a very specific set of decisions within the

---

<sup>954</sup> Summarized from Platform for the Information Society, *Artificial Intelligence Impact Assessment*, Platform for the Information Society Publication (2019) available at <https://ecp.nl/wp-content/uploads/2019/01/Artificial-Intelligence-Impact-Assessment-English.pdf> (last viewed Aug. 14, 2019). This is not the only framework in use, for example AINow, *Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability* (2018). While others contain similar considerations in a broader discussion. See e.g., Information Commissioner's Office, *AI Auditing Framework, Call for Input*, *supra* note 697.



French judicial system.<sup>955</sup> The response by the French government was to implement the Justice Reform Act which provides in Article 33:

The identity data of magistrates and members of the judiciary cannot be reused with the purpose or effect of evaluating, analyzing, comparing or predicting their actual or alleged professional practices.<sup>956</sup>

The response, originally widely criticized, is one example of growing concern about careless deployment of AI undertaken without due consideration of its possible impact. And, while this may seem (and likely is) an over-reaction to the use of judicial data in algorithmic processes, one needs to understand the reasoning behind the law. Commentator Michaël Benesty explains that the legislation was written in response to disturbing patterns discovered amongst the various judges making asylum decisions.<sup>957</sup> He further asserts that the main issue that arose was releasing the names of the particular judges.<sup>958</sup> How did a deployment of technology that resulted in revealing bias not envision bias discovery as a potential outcome? In that situation, how can anyone imagine or justify releasing the names as an appropriate first step in correcting the issue discovered?

The process proposed within this thesis is designed to monitor decisions but also intentionally designed to provide feedback without publicly shaming decision makers. The same intentional approach is taken with some personal data that may be gathered during the roster/conflict process. During early research into designing a prediction algorithm the author was involved in, I asked “what if what we assume is true is right, what if we discover a judge is in fact biased and we publish that determination, openly?” Ultimately the team changed course on the research, determining that our goal was to predict and to help demonstrate bias, which can be done without public shaming. And, in the process we decided to use the knowledge as a tool to create better decision makers — through awareness and education. This is (hopefully) a perfect example of examining deployments with impact assessments as a key component. It’s a decision I am proud of still today. Not fully thinking through the impact upon *all parties and*

---

<sup>955</sup> See Author, *France Bans Judge Analytics*, *supra* note 793.

<sup>956</sup> *Id.*

<sup>957</sup> Author, *The Judge Statistical Data Ban – My Story – Michaël Benesty*, ARTIFICIAL LAWYER, (June 7, 2019)

<sup>958</sup> *Id.*

*groups* in the system, including the individuals providing the data, is a fundamental flaw in the manner in which technology has been historically deployed. Releasing data, finding patterns, predicting behavior, these are all old technology being deployed in a highly surveilled, massively connected, ultra-high speed, digital world — we simply have to consider the impacts beyond our own deployments and deliverable promises. This is a new adventure for everyone.

#### D. Design Guidance Conclusions

This final chapter has provided eight rules of action: (1) the Rule of Law must be protected within the online justice environment; (2) an independent and impartial decision maker is an essential RoL principle, such that no departure from the principle should be allowed; (3) it is the method of ensuring the independence and impartiality of the decision maker that is open for greater debate; (4) it is the model design that mitigates dependence and bias; (5) the reduction of negatively impactful bias is the current “best practice” standard; (6) there must be a human within the final step of the technology driven decision making model, built upon the ideal of an emotionally intelligent decision maker; (7) accountability is essential as ODR is part of the justice eco-system institution; and finally, (8) best practices demand a wider conversation about all of these issues.

The author recommends that the justice community continue to develop and test models within the justice environment, but to insist upon specific design protections, including that a human must make the final decision. The model should be viewed by the human as suggestive and outcomes should be displayed in a manner that fully informs the human decision maker of the nature of the predictions being made and the accuracy/percentage of certainty of the prediction.

In the area of Model Selection, best practices would be to insist upon transparency of model selection, model deployment, the features used to train the model (if any), the features discovered and added to the model after training, and the features currently used in the prediction, and some level of accountability for the deployment of the automation/AI, especially if it is to occur in a justice environment.

In the area of Model Learning, best practice would be established by focusing on transparency of the features used to train the model. Those features that are on the “restricted from use” list, the features discovered and added to the model after training, and the features currently used in the prediction. Consideration should be given to restricting automated machine learning around complex and impactful concepts that are not easy to fully capture in algorithmic terms (e.g., race). And, again some level of accountability for implementing steps to reduce bias and discrimination, including (but not limited to) restricted-from-use lists.

Model risk best practices entail establishing a transparent process, including the presentation of information about the human decision makers, their conflicts, prior outcomes, and party involvement in the appointment process and the ability to audit *all final determinations*, including final award decisions. The Auditing information should be used as both a measure to ensure appropriate process and as a mechanism to provide feedback to the human decision maker; with a built-in ability to evaluate, identify and mitigate risk — so long as the manner of mitigation is in line with RoL and best practices. Finally, the creation of a robust, multi-discipline review team to audit the entire process and outcomes.

The risk of incorrect use or lack of understanding of the technology gives rise to best practices built around commitments to the outcomes being transparent; a robust evaluation and identification of issues within the process, including issues arising from incorrect use, must be part of a feedback loop to the human decision maker. Moreover, focus must be placed on the manner in which the human decision maker will interact with and be influenced by the technology; the outcomes, predictions and/or assessments must be presented to the human decision maker in a manner that clearly explains the technological process is not intended to be a complete process, but rather to provide key information and assessments, and that displays the potential for multiple outcomes. Outcomes, predictions and/or assessments accompanied by a model card, which must include full details of the model and its capacity and design. Decision makers must undergo thorough training on the technology they are to use, its risks, and appropriate ways to use it.

This chapter has proposed a model that may be used in the near future and one that could serve as an initial point of discussion for addressing risk-based issues and the intended mitigations to those risks. The author identified key solutions that could be deployed in the areas of: (1) the role of institutions and/or platform designer; (2) the roster of decision makers; (3) nomination and appointment of the human decision maker; (4) challenges to the appointment; (5) transparency of the appointment process and the humans in the process, including: (a) risk arising from the non-disclosure/rule compliance; (b) risk arising for implicit/hidden bias of the decision maker; (c) managing user risk; (d) risk not already considered arising from the introduction of discretion.

## VI. Thesis Conclusions

As Online Dispute Resolution (ODR) begins to enter the wider market, it promises to substantively change the way arbitration works in the justice eco-system. The justice community must determine the fundamental values to be incorporated into the online justice environment so that it can address ethical and regulatory issues in the design of ODR. This thesis has engaged in an in-depth examination of the arbitration eco-system as a mechanism to ensure the Rule of Law in the emerging Online Justice Environment. Chapter 1 set out that a commitment to the RoL, including the principle of judicial independence and impartiality, demands that final decision makers in the ODR system must be *both* impartial *and* independent, without exception, even if the adherence to this rule of law principle conflicts with maximizing the efficiencies of the overall ODR system.

To accomplish successfully meeting this commitment, the ODR system must be examined through the lens of risk identification and mitigation principles being deployed to reduce risk in technology-driven deployments. As it relates to the decision maker, this is especially true in instances where algorithms are deployed within the decision making process. In these instances, technology-driven bias must also be identified and mitigated against.

Chapter 3 sets out to examine the existing ODR platforms, institutional governance efforts, and industry practices to consider the presence of RoL commitments in both the ODR community and the academic commentary. It is in this chapter that differences are clearly emerging between those that seek ODR as an ancillary to the court systems (public) and those that attempt to craft ODR as a private mechanism tangentially attached via enforcement regimes to the existing justice environment. The thesis takes a neutral perspective on the best approach to long term governance, instead focusing on the issue of the decision makers in the process, regardless of public or private built system.

This thesis directly addresses the following questions that were posed in the introduction: What is the ideal relationship between technological design and fundamental principles of the rule of law? What are the essential features a judicial decision maker must embody in order to uphold the rule of law? As they establish guiding principles and best practices, model designers will need to consider the legally proscribed parameters and protections associated with the legal conceptualizations of the key terms *independence* and *impartiality* as well as the mechanisms needed to ensure a neutral decision maker. It is also important to ask what risk model should be used to evaluate the new systems being designed, and what mitigation can be deployed to reduce identified risks?

To address these issues, the thesis argues for adherence to seven rules of action: (1) the rule of law (RoL) must be protected within the online justice environment; (2) an independent and impartial decision maker is an essential RoL principle, such that no departure from the principle should be allowed; (3) it is the method of ensuring the independence and impartiality of the decision maker that is open for greater debate; (4) it is the model design that mitigates dependence and bias; (5) the reduction of negatively impactful bias is the current “best practice” standard; (6) there must be a human within the final step of the technology-driven decision making model; (7) accountability is essential as ODR is part of the justice eco-system. Finally, best practices include (8) wider ongoing conversations about all of these issues.

Based on these Rules of Action, the author has proposed specific model design expectations necessary to ensure commitments to the RoL and the seven rules within the design of the ODR eco-system. First, the thesis calls for institutions to continue to develop and test models within the justice environment, but to insist upon specific design protections, including that a human must make the final decision. Designers must actively mitigate against human machine bias by ensuring the human in the loop is viewing results as suggestive, with outcomes displayed in a manner that fully informs the human decision maker of the nature of the predictions being made and the accuracy/percentage of certainty of the prediction.

Concerning model selection and the manner of learning deployed, transparency and accountability are key. Transparency of model selection, model deployment, the features used to train the model (if any), the features discovered and added to the model after training, the features currently used in the prediction, and a level of accountability for the *design and deployment* of the automation/AI, are all essential points that merit transparency — especially if used in a justice environment.

Risk arising from model design and deployment is one of the more challenging endeavors; transparency, auditing, feedback, mitigation and multi-stakeholder review are key within this area. Accordingly, the thesis calls for the establishment of a transparent process, including the presentation of information about the human decision makers, their conflicts, prior outcomes and party involvement in the appointment process with the ability to audit *all final determinations*, including final award decisions. The auditing information should be used as both a measure to ensure appropriate process and as a mechanism to provide feedback to the human decision maker, providing the decision maker the opportunity to evaluate, identify and mitigate any identified risk — so long as the manner of mitigation is in line with RoL and best practices. All of these processes must be done under the watchful eye of a robust, multi-discipline review team to audit the entire process and outcomes.

There is growing research about risk that arises when the human lacks understanding of the technology they are using or use it in incorrect ways. This risk can be partially mitigating by requiring that the outcomes, predictions and/or assessments must be

presented to the human decision maker in a manner that clearly explains the technological process is not intended to be a complete process but instead must be presented in a format that allows for the presentation of multiple model outcomes. Any issues or patterns identified in the system or the individual must be considered part of a feedback loop to the human decision maker, allowing the human to receive the feedback, reflect upon it and react as the human sees appropriate. As a final mitigation safeguard, the human in the loop must be considered part of the auditing process. Thus, outcomes must be transparent and must be considered part of an evaluative process in which issues within the process, including issues arising from incorrect use, are identified.

The author identified keys solutions that could be deployed to reduce negatively impactful bias by focusing on several keys areas of design. First, as was explored in great detail in Chapter 2, the international arbitration community has considered the role of the institution within the alternative justice environment. Drawing from the lessons learned in this area, the thesis supports the continuing role of an institution, or similarly situated entity, as a mechanism to ensure decision makers can be evaluated under rules establishing parameters of independence and impartiality. Moreover, institutional rules can establish disclosure expectations and can serve as a gatekeeper for the dissemination of information surrounding the decision makers. The institution can also serve as a caretaker of appropriate design, by being held accountable for the deployment and serving as a mechanism to sanction poor design by those that design the technology.

Chapter 5 sets out some basic requirements for how the institution should be expected to create a roster of decision makers, establish a nomination and appointment process and create rules concerning the transparency of the entire process and outcome. Most importantly, the institution in the ODR eco-system is allowed to create a shorter process of appointment, because the technology can, and should be expected to be used to, eliminate some decision makers from selection, based upon a set of criterion designed to reduce partiality and dependence even in privately disclosed activities. The significantly shorter list of available decision makers could then be used to create a random list of potential decision maker, which the parties could then rank based

Notably, this is an area where technology, through automatically managing one tier of conflict checks, can add efficiency with no damage to RoL.

Finally, the commitment to include arbitration as the final stage of alternative dispute resolution builds in a level of governance, one that needs to be specifically expanded in the ODR eco-system. The current need to comply with enforcement expectations embodied within the New York Convention makes the RoL a legal requirement and not merely aspirational. The enforcement mechanism offers final assurance of compliance with RoL, an independent decision making process, and the absence of negatively impactful dependence and partiality in the decision maker. Through the courts, those who build and deploy ODR models could be held accountable by demonstrating appropriate identification of risk, using impact assessments, with the requirement to explain the trade-offs that occur in the mitigation of those risks.

However, this proposal assumes several things, first that arbitration is in fact the final stage of the process and second that courts are willing to review RoL, due process and human rights within the enforcement of an award. Neither of these are currently true assumptions. First, smart contracts are in vogue right now and many entities are considering how to use smart contracts in many traditional contracting arrangements. If smart contracts do become more mainstream, the enforcement mechanism will likely be less based in courts and more based in execution of code. Second, courts seeking to refuse enforcement of awards under the New York Convention frown upon expanding the potential grounds of refusal. There is no clear application of technology considerations within the convention, and some argue it would be well outside its drafters' intentions. As such, stakeholders in the justice eco-system must begin to envision a world of technology-driven dispute resolution and must insist upon a commitment to basic RoL protections. A commitment such as this was undertaken by the international arbitration community as arbitration began to grow in popularity as an alternative to the justice system — something outside the traditional justice environment, and not without critics. Opportunistic disruption should not be allowed to overwhelm commitments to RoL. Thus, there should be a commitment to wider conversations amongst stakeholders from divergent, interested communities to create a



system of technology deployment that upholds justice as an essential consideration in the justice eco-system.

This thesis reflects the law, jurisprudence, and doctrine until 14 August 2019. Thus subsequent developments are not included.

### **\*Epilogue**

As Online Dispute Resolution (ODR) is now entering the justice environment in many new ways, especially in the Covid-19 world, the use of technology in the justice environment is growing, exponentially in many areas. In many instances, technology is deployed to mimic existing systems, thereby using technology as nothing more than a simple digital communication device. In other deployments however, the technology community is seen as a vendor, selling a product- often without anyone considering the broader considerations necessary when any change is made to a justice environment. As such, the justice community must determine the fundamental attributes to be incorporated into the online justice environment and must determine how to advance, support, and engrain those attributes into the technology deployments. This thesis is the first attempt of any ODR scholar (at the time of writing) to identify the fundamental attributes that must be embedded in the decision-making portion of the ODR process.

Although the thesis demanded a narrowing of issues to a topic that was capable of being resolved within the confines of a thesis, the narrowing of the topic does not diminish its importance. Instead, the thesis tackles one of the key aspects of any alternative to brick-and-mortar justice environments, the principle of a decision maker's independence and impartiality. The issue is even more complex when technology is considered an actor in the decision making. As such, the thesis takes on the additional concern that arises in the deployment of technology, concerns of bias.

Although narrowed to a specific aspect of potentially a much larger topic, and undoubtedly a subject of future debate, the exercise of identifying the issue, finding various potential approaches, identifying commonality and specifying a definition, and

then operationalizing those aspects into the design process of the decision making portion of the ODR world, is essential first step in advancing broader conversations in the area of online justice. It is in fact, the next in a series of conversations for many systems considering the use of technology in their justice environment. Court systems (around the world), APEC, even the World Economic Forum are working on these vary issues and that is no accident. Scholars, lawyers, and court technology adopters are growing aware of the need to consider design principles, models, and essential attributes as key parts of initial conversations when considering alternative between providers and technology systems.

The author makes no claim to this thesis being perfect, in fact- as an epilogue, it is fair to say the world is much different from the time of writing, as the global community is now in the midst of a Covid-19 online reality. Time and thesis limitations prevent the author from adding the new world to the thesis. However, for the first time the justice community- world-wide, is considering how technology can be used to improve access to justice. Yet, as the author writes early in the paper, efficiency is no substitute for key attributes of the justice process, instead efficiency is a goal after ensuring essential attributes exist in the system.

As such, the thesis contributes to the academic and justice community in ways unforeseen at the time of initially launching the thesis. This thesis has consistently advocated for the principle of judicial independence and impartiality: decision makers in the system must be *both* impartial *and* independent, without exception, even if the adherence to this rule of law principle conflicts with maximizing the efficiencies of the overall system. As it relates to the decision-making process that incorporates technology within the model, technology driven bias must be identified and mitigated against. To accomplish these goals, the ODR system must be examined through the lens of risk identification and mitigation principles. This thesis is the first, to examine the ODR model within the technology driven aspects of model design. And, the new reality of the growing support for wider use of technology and online alternative decision making does not change the argument, in any way.

Rule of law guides us to the broad essential principles, in this case, as explained in Chapter 1, the independence and impartiality of the decision maker is essential. Chapter 2 uses the well regarded, and widely supported alternative of arbitration to help define the specific parameter of those two key terms. And, while the terms are not capable of operationalizing based on the arbitration institutional rules, as Chapter 2 also explains, court cases specific to arbitration lend us the parameters- and limitations of the terminology. The chapter clearly defines the terms- terms that are then brought into Chapter 5.

To operationalize these terms, the thesis sets out, explains and supports the assertion that any system – or model- must adhere to the following seven rules of action: (1) the rule of law (RoL) must be protected within the online justice environment; (2) an independent and impartial decision maker is an essential RoL principle, such that no departure from the principle should be allowed; (3) it is the method of ensuring the independence and impartiality of the decision maker that is open for greater debate; (4) it is the model design that mitigates dependence and bias; (5) the reduction of negatively impactful bias is the current “best practice” standard; (6) there must be a human within the final step of the technology driven decision making model; (7) accountability is essential as ODR is part of the justice eco-system.

Each key section within the thesis ends with a list of specific attributes that must be contained within the model and those attributes are then operationalized within the model. Each of these attributes are then incorporated in the model with Chapter 5 as explained in the Model conclusion, where the author cut and pastes the attributes from the end of each section and then highlights how those are brought into the model.

To date, there is no existing research that considers, in such specific and supported detail, the area of independence and impartiality in the ODR deployments. In fact, as Chapter 2 reveals, arbitration institutions themselves lack this level of specificity and operationalization. Simply put, the matter of definitions could have been a thesis, unto itself, as could an examination of rule of law in the area, but that would have left an emerging system without the benefit of academic scholarship demanding a robust examination and creation of clear, operationalized terms, and model design.





## Appendix A: Details of collaboration and publications:

The Accord Project, Dispute Resolution Working Group, Co-Organizer (2018-Current).

APEC Workshop for Developing a Collaborative Framework for Online Dispute Resolution (ODR) Japan International Dispute Resolution Center, Osaka, Japan, US State Department (2018).

GOVERNANCE OF ALGORITHMS: RETHINKING PUBLIC SECTOR USE OF ALGORITHMS FOR PREDICTIVE PURPOSES (with Ciabhan Connelly) in CAMBRIDGE HANDBOOK ON LAW AND ALGORITHMS, Cambridge Press, (expected late 2019).

*Information and The Regulatory Landscape: A Growing Need To Reconsider Existing Legal Frameworks*, 24 WASHINGTON AND LEE SCHOOL OF LAW JOURNAL OF CIVIL RIGHTS & SOCIAL JUSTICE, 357 (2018).

*Building a Better HAL 9000: Algorithms, the Market, and the Need to Prevent the Engraining of Bias*, (with Emma Arrington Stone Young and Scott J. Shackelford), 15 NORTHWESTERN JOURNAL OF TECHNOLOGY & INTELLECTUAL PROPERTY 215 (2018).

*The Consumer As Sisyphus: Should We Be Happy With 'Why Bother' Consent?*, JOURNAL OF LEGAL STUDIES IN BUSINESS, Vol. 20, 1-26 (2017).

*A Meeting Of The Minds: Online Dispute Resolution Regulations Should Be Opportunity Focused*, U.C. DAVIS BUSINESS LAW JOURNAL, Vol 16(2) (2016).

*The Angel on Your Shoulder: Prompting Employees to Do the Right Thing Through the Use of Wearables*, (with Timothy L. Fort & Scott J. Shackelford) 14 NORTHWESTERN UNIVERSITY JOURNAL OF TECHNOLOGY AND INTELLECTUAL PROPERTY 139, 139-170 (2016).

*Jury Glasses: Wearable Technology And Its Role In Crowdsourcing Justice*, (with Scott J. Shackelford) CARDOZO JOURNAL OF CONFLICT RESOLUTION, 17:115, 115-153 (2015).

*Trusting Strangers: Dispute Resolution in the Crowd*, (with Abbey Stemler) 16 CARDOZO JOURNAL OF CONFLICT RESOLUTION 357, 357-394 (2015).

*The Dilemma of Private Justice Systems: Big Data Sources, the Cloud and Predictive Analytics*, NORTHWESTERN JOURNAL OF INTERNATIONAL LAW & BUSINESS (Ambassador-online\*)(2015)

*Technology, Ethics And Access To Justice: Should An Algorithm Be Deciding Your Case?* 35 MICHIGAN JOURNAL OF INTERNATIONAL LAW 485, 485-524 (2014) (with Scott Shackelford).

*Yeah, But Did You See the Gorilla? Creating and Protecting an 'Informed' Consumer In Cross Border Online Dispute Resolution,* 19 HARVARD NEGOTIATION LAW REVIEW 129, 129-171 (Spring 2014).

*Building the Virtual Courthouse: Ethical Considerations for Design, Implementation, and Regulation in the World of ODR* (with Scott J. Shackelford) 2014 Wisconsin Law Review 615 (2014).

*Confidentiality, In a Forum of Last Resort? Is the Use of Confidential Arbitration a Good Idea for Business and Society?*, 16 AMERICAN REVIEW OF INTERNATIONAL ARBITRATION, 479 (2005) ISSN:1050-4109.

*Ethics for Commercial Arbitrators: Basic Principles and Emerging Standards* (Co-author with H. Gabriel), 5 WYOMING LAW REVIEW, 453-470 (2005) ISSN: 0023-7612.

## Appendix B: Influences and Influencers

Shortly after I started developing this thesis, I was selected to participate in one of the first “Ethics of Predictive Algorithms” Workshops, hosted by Janine Hiller. At the time, I had written around the edges of this emerging area, but neither of us could have imagined how important this topic would quickly become. Shortly after the Workshop there were several books published considering similar areas, including Cathy O’Neil, Weapons of Math Destruction; Virginia Eubanks, Automating Inequality; Brett Frischmann, Re-Engineering Humanity; and some very important journal papers written by Andrew Selbst; Solon Barocas; Margaret Hu; Nizan Geslevich Packin; Kevin Werbach; John Bagby; Philip Nichols; David Nersessian; Yafit Lev-Aretz; Anne Washington; Jessica Eaglin. Each has influenced my thinking in ways that could never be fully captured in a footnote. They have wonderful writings in important areas, some have challenged me in person to think more creatively about solutions, and all have influenced my ability to digest and engage in these incredibly important areas of research.

Around the same time that prediction was emerging as a topic, ODR was seeing renewed interest as well. Vikki Rogers first got me involved with the U.S. efforts at the U.N. Working Group, and brought my attention to focus on a clear area of interest. Those important meetings led me to a wonderful ODR family: Darrin Thompson; Mike Dennis; Jeff Aresty; Colin Rule; Amy Schmidt; Nancy Welsh; Ethan Katsh; Nevena Jevremovic; Larry Bridgesmith; Sharon Salter; each has had an influence on me in immeasurable ways.

Several of my co-authors have also greatly contributed to my thinking, Dr. Scott Shackelford, Abbey Stemler and Emma Young have all influenced my examination of various aspects of the digital world. As have those I have been privileged to workshop papers with, such as: Deven Desai; Ben Green; Gregg Bloom; Simon Boehme; Milton Mueller; Doc and Joyce Searls; and Jody Blanke, to name a few. And, those that I have been fortunate to hear speak, including Ryan Calo and Bruce Schneider.

None of this would have been possible without my background in international commercial law and international commercial arbitration. The chapter on arbitration was the easiest to write thanks to my experiences with CCLS and those who fill its hallways, especially Dr. Loukas Mistelis, Dr. Julian Lew (QC) and Dr. Stavros Brekoulakis. And, of course, I have received guidance and immeasurable feedback Rosa Lastra.

As of August 14, 2019 I declared this thesis closed to additions. Nonetheless a groundswell of interest is building in the area of AI Governance, and new frameworks, white papers, initiatives, legislation and working groups are formed every day. Momentum is building and it's exciting to be contributing to these advancing conversations.

I am incredibly proud to have been thinking about widespread use of algorithms early in the process. Proud of the people we have brought to campus at Indiana University, the trips I have done, being part of a growing effort to have people move beyond, design-and-dash, deliverable-focused technology and expand the conversation to include the humanities, pro-sociology, and socio-technical considerations.



## Appendix C: Style Guide

The following is a citation guide, based on a modified U.S. Law Review citation system, known as THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION. The style has been modified to accommodate a longer document, an international audience, and an online heavy research environment. As such, a style guide is included which has been specifically designed to assist the examiners in finding sources.

### Journal and Similar Citation

Author, *article title*, JOURNAL CITATION, (year).

-These citations may also include, page number, and a (descriptive summary of the point the source is being relied upon to support).

### Book Citation

Author, BOOK TITLE, Publisher, (year).

-These citations may also include, page number, and a (descriptive summary of the point the source is being relied upon to support).

Author, *Chapter title, cited in* Author, BOOK TITLE, Publisher, (year).

### Materials, in general

Author/Institution/Agency, *title*, institutional notation/number/etc., article number/section/ page, (year).

### Cases

U.S. cases- Name of Parties, *citation*, (year).

Non-U.S. cases, in general- Parties Names, [year] *citation*.

### Online materials

Author, *article title*, JOURNAL CITATION, (year).

-These citations may also include, page number, and a (descriptive summary of the point the source is being relied upon to support).

-These citations may also include a link to a webpage and a (Last viewed) date. The extra information was included due to (1) the

entrepreneurial nature of the environment, and/or (2) to citations that are less well known that the author believed might be more difficult for an examiner to discover.

### Short Citations

*Supra* (previously) and *infra* (coming later in the text) are used to assist readers in finding prior discussions and to prevent repetitive lengthy footnotes.

Journals and similar documents

Author, *short title*, *supra* note number.

Books

Author, SHORT TILE, *supra* note number.

Cases

*First party listed name*, short court number, at location.

### Signals

Are italicized internal notations to help the reader better understand the type of citation that is being put forward.

*See* – is used when the citation is being done to give attribution to a source, but the citation is not a direct quote. Thus, the absence of a see signal means a direct quote is in the thesis.

*See also-* is used when authors have written on the subject and may be useful to the reader to consider within the context of a broader discussion.

*See discussion-* is a signal used to direct the reader to a section, footnote, or table in the thesis that contains additional or more expansive information and support.

*Id.* is a signal used to prevent repetitive footnotes. The reader should look to the last citation in the prior footnotes. *Id.* is used when the material is a direct quote.

*See id.* is a signal used to prevent repetitive footnotes. The reader should look to the last citation in the prior footnotes. *See id.* is used when the material is a paraphrasing or summary of the material.

*See e.g.*, is a signal used to direct the reader to other available material that is an example of the point made in the sentence. It can be read as ‘for example.’

*See generally*, is a signal used to direct the reader to other available material that is highlighted for a general proposition. It is not used to support a specific point, but a general area being supported.

[hereinafter] is a signal used to create a short cite for a frequently used authority. The signal is at the end of the first footnote in which the authority appears.