THE USE OF ARTIFICIAL INTELLIGENCE IN EDISCOVERY

FEBRUARY 2021

This paper is dedicated to the memory of David Kittrell, a long-time pioneer and contributor in the field of eDiscovery and an early member of the drafting team for this paper. David passed away in September 2020 and will be sorely missed and fondly remembered.

EXECUTIVE SUMMARY

This brief survey provides a description of how artificial intelligence, or “AI,” is currently used in areas related to eDiscovery. Its primary purpose is to provide background information on the use of artificial intelligence in eDiscovery. Artificial intelligence and eDiscovery are both large, complex, rapidly changing fields and it is impossible to do justice to either of them in a paper of reasonable length. There are many books about AI and a few about eDiscovery. Our intention is to provide a preliminary overview of the field.

There is no universally accepted definition of “artificial intelligence.” We adopt a working definition that artificial intelligence refers to the capability of machines to mimic aspects of human intelligence, such as problem solving, reasoning, discovering meaning, generalizing, predicting, or learning from past experience. Of the different types of AI, machine learning is probably the most prominent and most familiar to eDiscovery practitioners.

Although machine learning receives the bulk of attention in the context of AI in eDiscovery, rulesbased systems also meet the working definition of artificial intelligence and can play an important role in eDiscovery and other legal activities.

Applications of AI in eDiscovery include document categorization (e.g., technology-assisted review (“TAR”) or predictive coding), identification of personally identifiable information (“PII”), investigations, and some forms of early case assessment. AI can offer significant cost and time savings relative to previous unautomated methods. AI should not be used by persons without an appropriate understanding of its capabilities and limitations. Doing so, or uncritical acceptance of its results, can lead to errors and even potential ethical issues.

1. INTRODUCTION

Artificial Intelligence, or “AI,” refers to the capability of machines to mimic aspects of human intelligence, such as problem solving, reasoning, discovering meaning, generalizing, predicting, or learning from past experience. AI includes both unsupervised and supervised machine learning, but it also includes a number of other processes, such as natural language processing (“NLP”). In the context of this paper, AI describes an automated process that is used to classify, categorize, summarize, makes predictions, or provide information regarding data or information using statistical, rule-based, or other algorithmic means.

Alan Turing generally is credited with the origin of the concept of AI when he speculated in 1950 that "thinking machines" could reason at the level of human beings. Turing proposed an "imitation game," which others have called a "Turing test," as a means of deciding whether a computer was intelligent.

Essentially in its simplest terms, the argument was that a computer that was capable of holding a conversation with a human observer would demonstrate intelligence if the observer could not discern whether he or she was conversing with another person or with a computer. A computer that could imitate a human in conversation would be said to pass the Turing test and therefore be intelligent.

A few years later, John McCarthy introduced the term "artificial intelligence" in a proposal for a 1956 workshop on building machines to emulate human intellectual capacity. This workshop was intended to investigate "the conjecture that 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. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves."[1]

Although McCarthy and the other participants in the 1956 workshop had high hopes for accomplishing with computers the full range of human intellectual capacity, computer scientists have succeeded instead in developing computer systems that each address a narrow range of accomplishments, such as playing games like chess, Go or *Jeopardy!*, identifying cancer in images of skin lesions, and driving vehicles. Artificial general intelligence—where a computer can perform all the tasks a human can—remains a future goal, and there is some debate as to whether it will ever be achieved.

Some methods of machine learning require human supervision and some do not. Unsupervised machine learning refers to AI systems that can operate independent of, or prior to, human intervention— such as concept clustering. These methods organize and classify documents or data by various features, which can include subject matter, without human training. Supervised machine learning refers to AI systems that are trained based on human decision making—such as technology-assisted review ("TAR") or predictive coding programs that work by having humans classify some documents as "relevant" or "not relevant" to a particular subject matter, and then having computer software learn to make those distinctions based on automated analysis of the human decisions. There are other technologies used in the law that may not use machine learning at all, but are sometimes referred to as "artificial intelligence" because they mimic human cognitive processes.

Rules-based systems also can play an important role in eDiscovery. The main advantage of rulebased systems is their transparency or explainability; humans can often see how the computer makes its decisions. In supervised machine learning, humans provide the criteria for rendering a classification and examples of the different categories, so the computer can learn the distinction between the different categories. In rules-based systems, rather than the computer discerning the rules, humans provide the rules explicitly and transparently. Rules-based systems are more often used in legal contexts outside of eDiscovery, but they can have some use, for example, in anonymization and redaction. Rules-based

[1] McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

systems are used less often for categorization in eDiscovery because they require the expertise of both subject matter experts and rules-construction experts, including linguists and statisticians. Therefore, they tend to be more resource intensive and time consuming to develop as compared to machine learning approaches.

2. TYPES OF AI USED IN EDISCOVERY

a. CLUSTERING

Clustering is an example of unsupervised machine learning. The purpose of clustering is to group "similar" items together, which allows users to recognize the characteristics or topics that make them similar. This allows users to learn something about the composition of the data set or to take action on a whole group of similar items (*i.e.*, documents). Clustering is unsupervised in the sense that users do not control the dimension(s) along which "similarity" is defined and do not have to label examples of items in each cluster in order to train the system, but the designer of the system generally does have to specify the features along which item similarity is to be measured and how many clusters there should be. **b. EMAIL THREADING**

Most emails do not occur as single items but rather are part of an ongoing conversation. When one replies to an email, it is typical to include the original email as part of the reply. Email threading works to identify all of the emails in the same conversation so they do not have to be reviewed separately. Email threading reduces work and improves consistency by grouping the emails of a conversation together in a single unit.

c. CONCEPT SEARCH

Words tend to have different meanings because of their context. For example, "strike" in the context of "bat" and "ball," has a different meaning than "strike" in the context of "management" and "labor," or "smack" and "hit." When using keywords to search for documents, searchers are often challenged by the difficulty of guessing the exact words that are used in those documents. In addition, words can have multiple meanings, depending on their context (for example, "court" could be related to "judge" or to "tennis"). This ambiguity is called "polysemy." Various words have the same meaning, for example, "doctor" and "physician." These words are called "synonyms." In addition to synonyms, words can have related meaning. For example, if a document contains words like "lawyer," "attorney," or "judge," then that document is likely to be about something legal. Any one of those words could be sufficient to identify a legal topic in a document. Concept search is another unsupervised machine learning method in which the machine learns the context in which words are used and mathematically models the relationships among words. Users can then search by meaning rather than by individual terms. So, a search for "cups" would bring back documents about "mugs" and "glasses."

d. Technology-Assisted Review ("TAR")

Supervised machine learning is used extensively in the eDiscovery process in the U.S. and abroad to rank or classify electronically stored information ("ESI") to identify documents to produce, with "black letter law" supporting its use.[2] TAR or predictive coding is widely available in eDiscovery software products on the market today. While there is considerable variability in workflow and particulars, the core idea is that a computer learns to distinguish relevant from non-relevant documents based on the coding of human reviewers and can then classify unlabeled documents on its own. This technology is now well established. Many, but not all, TAR algorithms are "language agnostic" such that they can rank or classify documents regardless of language.

e. Entity Recognition

Entity recognition is a supervised machine learning process where the computer learns to identify entities such as names of people, places, or companies, dollar amounts, job titles, account numbers, case/matter numbers, or other things. Entities in a text are words, or numbers with patterns (*e.g.*, XXXXX-XXXX) but they often have properties of interest that go beyond the individual word or number. For example, one may want to search for all documents that contain a credit card number or other personally identifiable information ("PII"). Searching for numbers can bring back too many documents that contain numbers, but are not credit card numbers, so entity recognition can be used to discern particular numerical patterns.

f. Sentiment Analysis

Sentiment analysis is used to identify emotional content of data, which can include excitement, anger, or other positive or negative emotions. For example, a sentence like "I hated that movie," would be classified as having a negative sentiment. On the other hand, "I found a marvelous pair of shoes," would be analyzed as having a positive sentiment.

g. Machine Translation

It is becoming increasingly common to encounter foreign-language documents during discovery. This, coupled with the overall growth in data volume, can make the costly and time-consuming process of human translation particularly burdensome. When much of the translated content is irrelevant, the return on investment can seem quite low. AI tools can be used to reduce, though not eliminate, the need for human translation. As an initial step, these tools can quickly identify those documents that contain foreign language text and list the languages they contain, allowing for more comprehensive planning. Some tools can actually translate the text from one language to another. Such machine-translation tools have improved to a considerable degree in the past few years.

[2] *Rio Tinto PLS v. Vale S.A.*, 306 F.R.D. 125 (S.D.N.Y 2015).

h. Anonymization and Identity Masking

AI tools are used to improve the speed and accuracy of the anonymization and identity masking of personal, confidential, and/or privileged information contained in electronic records. Some programs are able to systematically identify PII that then can be automatically redacted or "pseudonymized" (replaced with other information) to protect confidentiality. For example, identifying all instances of a person's name and replacing it with "Jane Doe 1." This can help, for example, in complying with the requirements of the European General Data Protection Regulation ("GDPR"). AI redaction tools can also copy redactions across duplicate documents, thereby ensuring consistency.

3. Uses of AI in eDiscovery

a. Ranking, Classification, and Review of ESI

In litigation, the volume and diversity of text-based data and the cost of reviewing that data manually has quickly accelerated the use of cutting-edge techniques.

One common use of AI is to segregate potentially relevant information that needs to be reviewed for litigation or investigations from likely irrelevant information. Another common use of AI by parties facing document requests is to prioritize and/or confirm relevancy decisions by human reviewers during the quality control ("QC") process.

AI technology is also used by receiving parties to prioritize incoming data to find the most important evidence to build their case. Products that offer nuanced entity extraction, such as the identification of people, places, and organizations, also bring AI to the ESI search realm. All these techniques coexist with more traditional keyword and metadata search, along with visualization and other analytic tools, to provide comprehensive data analysis capabilities.

The above applications most commonly assume that the input data is textual. Of course, there are also other data types. Vendors are increasingly finding better ways to handle video, audio, images, and structured data (such as databases or GPS data). Such solutions often incorporate AI technologies.

The quality of the ranking or classification resulting from a supervised machine-learning system typically depends more on the consistency and accuracy of the human input than on the precise machine learning method used to produce the ranking or classification.

b. Document Review and Quality Control

There are several different types of AI-driven technologies that can assist with document culling and review. First, some information can be culled from review with the assistance of unsupervised machine learning tools, such as clustering, email threading, and other classification tools, that can group information in ways that allow certain categories of documents or data to be eliminated from further consideration.

After initial culling is completed, TAR can be used to prioritize documents for human review and further eliminate irrelevant information from manual review. There are a variety of TAR tools available, some of which are calibrated based on training sets of documents and some of which continue to be calibrated throughout the course of human review.

AI can also augment the capabilities of document reviewers or provide quality control of their work. Software may auto-suggest tagging choices, highlight key text in a document, or provide context for references to people or key terms. While some of these capabilities can be built without employing AI, more advanced techniques tend to involve AI. Quality control is perhaps the simplest and least controversial place to apply the classification algorithms being adopted for prioritization or identification of relevant documents. For example, looking for high-scoring documents that have been coded as nonrelevant by a human reviewer and low scoring documents coded as relevant by a human. Even a review conducted entirely by humans can benefit from using AI to look for inconsistencies. In fact, algorithms tend to see features or patterns that are often distinct from those identified by people, lending greater confidence to situations where the algorithms agree with the humans. Cases where the algorithms and humans disagree can be a good place to focus quality control efforts.

c. Privilege Determinations

Privilege review is one of the greatest pain points in eDiscovery where AI may also helpful. There has been a gradual adoption of AI by some lawyers to assist with identification of privileged documents and the verification of privilege coding, as the tools have become more sophisticated and the magnitude of privilege review has grown. Privilege-review technology has not evolved to the point where it should be the only source used to categorize privileged documents or to make final decisions on privilege. This is due in part to the complexity of privilege, for example distinguishing legal from business advice, and the fact that privilege calls can differ from document to document based on subtle changes in language or additions to the recipient list. Two emails from the same email thread could have a different privilege status simply through the addition of a single person that waives the privilege. Not all classification tools use metadata and factor sender, recipients, and date into the equation when classifying documents. A human reviewer with knowledge of the elements that comprise privilege (or its waiver) may still have to check machine-made privilege calls. In addition, because there are documents that fall into “gray” areas where reasonable reviewers can reach opposite conclusions, it can be difficult to train AI tools to make such distinctions. Therefore, most AI tools using privilege-review technology must work hand in hand with human reviewers. However, the technology can speed up and improve the privilege identification and review process and can provide a quality control element.

AI tools can also be used to identify inconsistent privilege determinations across similar or identical documents, to refine privilege search terms, and to identify unknown parties involved in privileged communications.

d. Investigations

Investigations may be conducted for various purposes. Sometimes they take the form of Early Case Assessment(“ECA”) to quickly get a handle on key facts, projected costs, and likely outcomes. Sometimes they are part of an internal investigation into employee conduct or compliance issues that may lead to later employment-related or legal action. Sometimes they involve responding to government investigations that could include second requests, whistleblower complaints, regulatory inquiries or subpoenas, or general government oversight.

By using techniques such as data visualization, interactive dashboards, communication network analysis, clustering, email threading, concept search, and TAR, just to name a few, legal professionals can quickly analyze the document collection, and gain an understanding of the key underlying facts. Below we discuss some of the specific ways that AI can be used with regard to Early Case Assessment, internal investigations, and government investigations.

i. Early Case Assessment

AI is making significant inroads on the ECA process. Identifying the potential cost, risks, and related issues earlier in a case has always been a goal for legal professionals who are responsible for deciding on case strategies. Over the years, this process has become more complicated because data volumes have continued to grow, and datasets are becoming more heterogeneous and complex. To combat these conditions, legal professionals are employing AI during ECA to help them quickly analyze, assess facts, determine case strategies, and identify relevant parties involved in a case.

Even before a complaint is filed, legal teams are implementing AI on a subset of the document collection to learn the "players" and facts of a case. Legal teams can examine relationships between concepts and custodians. AI can help the legal team identify additional custodians who should be interviewed or placed on legal hold. Techniques such as query expansion and concept search can help identify new, unknown keywords and search phrases to find relevant ESI. Many types of AI software also help narrow the focus to highly relevant custodians and concepts. AI can also help legal teams identify large chunks of data that are likely non-relevant, allowing those documents to be quickly removed from search and review. These initial steps can save time and cost by reducing data volumes, as well as helping legal teams to focus on what is important and to set case strategy earlier.

Keyword search has not yet disappeared. In fact, search terms are sometimes used to "jumpstart" AI software by identifying good seed or training documents. Legal professionals are using search terms to identify "low hanging fruit" – documents that are easily accessible and relevant to the case issues. Once these relevant documents are identified, they are used to locate additional, conceptually similar documents and to train supervised machine-learning algorithms. AI tools can use an entire record or group of documents to bring up likely relevant ESI. Concept-search tools can also help them to identify terms or concepts in the collection that may not previously have been known to them.

Communication network analysis tools provide visualizations of communication patterns among individuals or email domains. Legal teams can quickly see the quantity, frequency, and timing of communication traffic; how people are self-organizing; and when new or unexpected parties enter the conversation. These tools allow legal teams to prepare and refine custodian lists and streamline custodian interviews. Once legal teams have identified the custodians, they can concentrate on specific individuals to further investigate the topics being discussed.

ii. Internal Investigations

Internal investigations may be initiated for a variety of reasons. Companies may receive whistleblower complaints about specific conduct; they may learn of other legal or compliance issues impacting other companies in their industry; or internal investigations may be undertaken as part of the process of ensuring regulatory compliance or to confirm that employees are behaving in accordance with legal requirements and company rules.

As in the Early Case Assessment example, AI can be used to quickly hone in on key facts and circumstances. Indeed, some AI may be operating in the background on an ongoing basis to identify potential employee misconduct or even external threats, such as data breaches. Similarly, AI can be used during mergers and acquisitions as part of "due diligence" to evaluate potential risks.

iii. Government Investigations

There are several aspects of government investigations that may be different from eDiscovery for typical litigation or even for internal investigations. For example, in some investigations, government agencies may not want to fully disclose what they are investigating, but targets of the investigation may still be compelled to respond to broad Civil Investigative Demands ("CIDs"), subpoenas, or other voluntary requests. In such circumstances, AI tools that can provide the recipient of such a demand an overall picture of the data set resulting from particular search parameters. This can help the subject or target to hone in on what the government may be investigating. A second difference from ordinary eDiscovery in the context of litigation is that there typically is no judge to referee discovery disputes, and the subject or target of the investigation may be entirely at the mercy of the investigating agency when seeking to limit the scope of the demands. Here, some of the AI tools may provide data that will help in negotiations with investigating authorities. Government investigations may also have shorter or less flexible deadlines for responses, therefore increasing the need for AI tools to help prepare a timely response. Note that when responding to government agency demands, it may be important to discuss in advance any AI tools that the responding party proposes to use for identification of responsive records, to avoid surprise and any adverse reaction from the investigating agency.

4. Benefits of using AI in eDiscovery

AI can offer significant advantages over the manual execution of tasks, including cost and time savings, increased efficiency, and greater consistency. Cost and time savings can be achieved by reducing the need for time-consuming human decision making. In addition, computers are not prone to human foibles such as fatigue, carelessness, and random inconsistencies. In certain circumstances, AI can also reduce or eliminate biased decision making.

5. Key considerations in using AI in eDiscovery

A key consideration when selecting AI technology for use in eDiscovery is to ensure that the purpose of the AI tool matches the desired results. All artificial intelligence tools are narrowly tailored to solve specific challenges. For example, most existing TAR technology is based on text classification and currently will not work well on documents that are strictly or primarily images (jpg, gif, etc). Many images with relevant content could be missed if TAR were to be applied. Steps should be taken to normalize the data to ensure it is properly prepared for the use of the selected AI tool. For example, optical character recognition ("OCR)" can be used to extract textual content from images. A consideration to bear in mind is: *Garbage in, Garbage out.* Without appropriate inputs, AI tools will not generate the desired results.

Another consideration when utilizing AI, as with all technology, is to ensure that sensitive data is secure during the entire process. A security team should be involved in auditing the entire AI data lifecycle as any unsecure links in the chain can compromise sensitive data.

Highly scalable solutions allow for increased potential use in legal analysis. They also allow for mistakes to rapidly scale across an entire data set. If the tool is not well-suited for the task, the data is not properly prepared, or there are too many coding inaccuracies in the human coding, such inaccuracies can promulgate across the universe. As a result, sufficient QC of human input is critical because AI has the power to amplify bad decisions, as well as good decisions.

6. Ethical Considerations

The use of AI in eDiscovery implicates several ethical concerns. The ABA Model Rules of Professional Conduct, versions of which have been adopted by 49 of the 50 U.S. states, highlight some of these issues.

Model Rule 1.1 and its accompanying comments require attorneys to provide competent representation to their client(s). As of the date of this writing, ethics committees in 37 U.S. states have agreed that this duty includes a requirement that attorneys stay informed and up-to-date on current technology and how that technology can be leveraged to provide the best results for their clients. Attorneys must, therefore, have a general understanding of AI technologies in the context of legal practice.

Attorneys also owe their clients duties of communication and confidentiality, as described by Model Rules 1.4 and 1.6. Attorneys must keep their clients informed about the technology being used in their cases and the involvement of any third parties. Attorneys must take affirmative steps to ensure the confidentiality of their client's' information, including examining the security measures of any third parties who have access to that information and how client information may be accessed on the thirdparty's system.

Attorneys must make reasonable efforts to expedite litigation consistent with the interests of their clients under Model Rule 3.2, as well as Fed. R. Civ. P. 1, requiring that attorneys participate in bringing about the just, speedy, and inexpensive resolution of disputes. As discussed above, AI has the potential to save significant time and money in the discovery process. Similar to the duty of competence, in order to provide effective representation, attorneys must know when and how to employ such technology to best serve their clients.

Model Rule 3.4 discusses attorneys' duty of fairness to opposing parties and to counsel, which has interesting applications when applied to eDiscovery and the use of AI technologies in particular.

Unscrupulous attorneys can hide evidence and otherwise behave unethically with or without the assistance of technology, but some have argued that AI may expand the ability of individuals misbehaving or for negligent attorneys to inflict harm absent adequate safeguards. Whether or not that is true, those responsible for employing AI technology should take extra precautions to ensure that ethical standards are met. Sampling and other QC measures can be powerful tools to help ensure that relevant evidence is not being improperly withheld.

Finally, the Model Rules hold attorneys responsible for the actions of others, including their attorney and non-attorney subordinates (see Model Rules 5.1, 5.3, 5.4, and 5.5). Licensed in-house and law firm attorneys maintain ultimate responsibility over every aspect of their legal representation, even those performed by non-attorneys or attorneys who are employed by non-attorney owned companies. In the realm of eDiscovery, this becomes particularly important because, as stated in Model Rule 5.4, any provider that is not 100% attorney owned and any employees that are not licensed in the jurisdiction in which they are employed are not authorized to practice law. Therefore, law firm and in-house attorneys must provide adequate oversight over any work done by litigation support and discovery service providers.

7. Future of AI in eDiscovery

Since the beginning of the rise of AI in the legal industry, there have been concerns of computers replacing, or diminishing the need for, attorneys and other litigation support professionals. While there is not universal agreement about the potential impact of AI on future legal employment, the increased use of AI and adoption by courts and legal entities across the world certainly brings to light the question of what effects AI may have on the legal industry. With the efficiency gained through leveraging AI technology, the questions must be asked: Will first-pass review by contract attorneys be a thing of the past? Will billable hours diminish or even disappear with the need for a new law firm billing structure? Will more automation of eDiscovery and reduced burden lead to courts allowing broader discovery?

While the answers to these questions are yet to be determined, there is no doubt that we can expect increasing reliance on AI to assist with a number of basic and time-consuming tasks, such as highly targeted collections, early data assessment, relevance and privilege review, issue categorization, and quality control. As the use of AI continues to expand, attorneys and their clients will increasingly have to decide: embrace the technology or be left behind.

A Brief Bibliography on Artificial Intelligence

Bostrom, N. (2014). Superintelligence: Paths, dangers, strategies. Oxford, UK: Oxford University Press.

Brown, T., Mann, B., et al. (2020) Language Models are Few-Shot Learners. arXiv:2005.14165v4 [cs.CL] 22 Jul 2020. https://arxiv.org/abs/2005.14165

Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM, 55(10), 78–87. doi:10.1145/2347736.2347755; http://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf

Fernández-Delgado, M., Cernadas, E., Barro, S., and Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems. Journal of Machine Learning Research, 15, 3133–3181. http://jmlr.org/papers/volume15/delgado14a/delgado14a.pdf

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., Lally, A., Murdock, J. W., Nyberg, E., Prager, J., Schlaefer, N. & Welty, C. (2010). Building Watson: An Overview of the DeepQA Project. AI Magazine, 31, 59--79.

Grossman, Maura R., and Gordon V. Cormack. "Technology-assisted review in e-discovery can be more effective and more efficient than exhaustive manual review." Rich. JL & Tech. 17 (2010): 1.

Markoff, J. (2016). Machines of Loving Grace: The Quest for Common Ground Between Humans and Robots. Ecco; Reprint Edition.

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Roitblat, H. L. (2020). Algorithms Are Not Enough: Creating General Artificial Intelligence. Cambridge, MA. MIT Press. https://mitpress.mit.edu/books/algorithms-are-not-enough

Roitblat, H. L., Kershaw, A. & Oot, P. (2010). Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review. Journal of the American Society for Information Science and Technology, 61(1):70-80.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review, 65, 386–408. doi:10.1037/h0042519

Russell, S., & Norvig, P. (2020) Artificial Intelligence: A Modern Approach, 4th Edition. Pearson.

Samuel, Arthur L. (1959). Some studies in machine learning using the game of checkers. Computation & intelligence: collected readings. American Association for Artificial Intelligence, USA, 391–414. http://www2.stat.duke.edu/~sayan/R_stuff/Datamatters.key/Data/samuel_1959_B-95.pdf

Shafer, G. (1976). A mathematical theory of evidence. Princeton, NJ: Princeton University Press.

Turing, A. M. (1965). On computable numbers with an application to the Entscheidungsproblem. In M. Davis (Ed.), The undecidable (pp. 116–154). New York, NY: Raven Press. (Original work published in Proceedings of the London Mathematical Society, Ser. 2, Vol. 42, 1936–7, pp. 230–265; corrections ibid., Vol. 43, 1937, pp. 544–546).

Turing, A. M. (1947). Lecture to the London Mathematical Society on 20 February 1947. Reprinted in D. C. Ince (Ed.) (1992), Collected works of A. M. Turing: Mechanical intelligence (pp. 87–105). Amsterdam, the Netherlands: North Holland.

Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59, 433–460.

CONTRIBUTORS

Mallory Acheson
Nelson Mullins

Kelly Atherton Driven, Inc.

David Cohen
Reed Smith

Derek Duarte
Blackstone Discovery

Tara Emory Driven, Inc.

Jeff Gilles
Relativity

Maura R. Grossman
University of Waterloo and Maura Grossman Law

Angela Lindstrom
Amazon

Wilzette Louis
Hogan Lovells

Herbert Roitblat
Mimecast

Matthew Sinner
HaystackID

Bradley Whitecap
Reed Smith