



**HAL**  
open science

## Data-Driven Collaborative Human-AI Decision Making

Gregoris Mentzas, Katerina Lepenioti, Alexandros Bousdekis, Dimitris  
Apostolou

► **To cite this version:**

Gregoris Mentzas, Katerina Lepenioti, Alexandros Bousdekis, Dimitris Apostolou. Data-Driven Collaborative Human-AI Decision Making. 20th Conference on e-Business, e-Services and e-Society (I3E), Sep 2021, Galway, Ireland. pp.120-131, 10.1007/978-3-030-85447-8\_11 . hal-03648128

**HAL Id: hal-03648128**

**<https://inria.hal.science/hal-03648128v1>**

Submitted on 21 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# Data-driven Collaborative Human-AI Decision Making

Gregoris Mentzas<sup>1</sup>, Katerina Lepenioti<sup>1</sup>, Alexandros Bousdekis<sup>1</sup> and  
Dimitris Apostolou<sup>1,2</sup>

<sup>1</sup> Institute of Communication and Computer Systems, School of Electrical and Computer Engineering, National Technical University of Athens, Zografou 15780, Athens, Greece  
{gmentzas, klepenioti, albous}@mail.ntua.gr

<sup>2</sup> University of Piraeus, Department of Informatics, 18534 Piraeus, Greece  
dapost@unipi.gr

**Abstract.** Business analytics use advanced techniques that can analyze and process large and diverse data sets in order to generate valuable insights and lead to better business decisions. Of the three types of business analytics – descriptive, predictive, and prescriptive – only the latter focus on decision making. This paper aims to address two limitations of existing approaches in prescriptive analytics: (i) the lack of a transparent integration between predictive and prescriptive analytics and (ii) the incorporation of human knowledge and experience within the decision-making process. In order to address these points, the paper develops a framework that integrates data-driven predictions and the decision-making process by taking account human experience. The framework adopts interactive reinforcement learning algorithms and provides a concrete approach for data-driven human-AI collaboration. The main challenges and limitations of the approach are also discussed.

**Keywords:** Human-AI Interaction, Data Analytics, Reinforcement Learning.

## 1 Introduction

Decision making has been studied from normative and descriptive approaches [1]. Normative theories focus on how to make the best decisions by deriving algebraic representations of preferences from idealized behavioral axioms. For example, the principle of utility maximization in economics and the concept of equilibrium in game theory describe how self-interested rational agents should behave individually or in a group, respectively [2]. On the other hand, descriptive theories incorporate known limitations of human behavior in the decision-making process. For example, prospect theory can successfully account for the failures of expected utility theory in describing human decision making under uncertainty [3].

Recently, these two traditional approaches of decision-making research have merged with additional disciplines [4]. It is now increasingly appreciated that learning plays an important role in decision making, although this has been ignored in most economic theories. In particular, Reinforcement Learning (RL) provides a valuable framework to model how decision-making strategies are tuned by experience [5, 6, 7].

Turning from the human to the organizational level, a recent trend in organizational decision making is the exploitation of the staggering amounts of data that organizations have at their disposal [8]. These data, if analyzed and processed properly, can generate valuable insights and lead to better business decisions [9]. The use of advanced techniques that can analyze and process very large and diverse data sets that include structured, semi-structured and unstructured data, from different sources, and in different sizes from terabytes to zettabytes, has led to the field of business analytics [10].

Business analytics focuses on data-driven decision-making and consists of three phases: descriptive, predictive, and prescriptive. While descriptive and predictive analytics allow us to analyze past and predict future events, respectively, these activities do not provide any direct support for decision-making [11, 12]. Contrary, prescriptive analytics is a new type of data analytics which enable data-driven optimization for decision support and planning [13]. It has been recently argued that the full exploitation of prescriptive analytics for optimized business decision making requires the incorporation of human knowledge and experience in the decision-making process [11, 14].

In the present paper we investigate how to explicitly incorporate human knowledge and experience in the data-driven decision-making process. In order to accomplish this, we incorporate interactive RL since it presents many opportunities for an osmosis of decision-making research in both human and artificial agents [15, 16, 17]. In interactive RL, a human interacts with an RL agent in real-time [18]. This approach has been shown to considerably improve the RL agent's learning speed and can allow RL to scale to larger or more complex problems [19].

The paper is structured as follows: first we introduce the theoretical background of prescriptive analytics. Then, we present the reinforcement learning mechanisms and focus on the way human-augmented approaches can be implemented with interactive RL. We define our proposed framework and describe how it favors data-driven human-AI collaboration. We discuss the main challenges of the approach and, finally, present our conclusions and describe future works.

## 2 Data-Driven Decision Making: Prescriptive Analytics

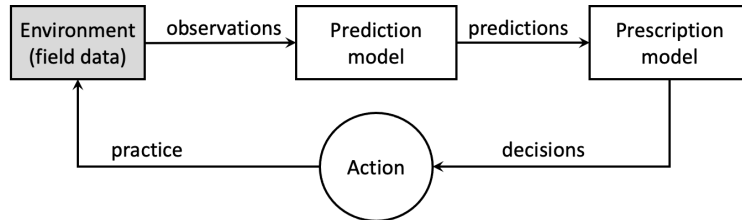
Business analytics refers to the extensive use of data, acquired by diverse sources, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions to proper stakeholders [20].

Business analytics is categorized to three main stages characterized by different levels of difficulty, value, and intelligence [11, 21]: (i) descriptive analytics, answering the questions “What has happened?”, “Why did it happen?”, but also “What is happening now?” (mainly in a streaming context); (ii) predictive analytics, answering the questions “What will happen?” and “Why will it happen?” in the future; (iii) prescriptive analytics, answering the questions “What should I do?” and “Why should I do it?” (see **Table 1**).

**Table 1.** Types of Business Analytics (adapted from [11], [13], [14])

Type of analytics	Descriptive	Predictive	Prescriptive
Questions answered	What happened? Why did it happen?	What will happen? Why will it happen?	What should we do? How can we make it happen?
Focus	Hindsight	Foresight	Decision
Main task	Analyze	Predict	Influence
Sample use cases	Performance management	Risk modelling	Portfolio optimisation
Indicative technologies	<ul style="list-style-type: none"> <li>• Dashboards</li> <li>• Statistics</li> <li>• Data mining</li> </ul>	<ul style="list-style-type: none"> <li>• Machine learning</li> <li>• Forecasting</li> <li>• Probabilistic models</li> </ul>	<ul style="list-style-type: none"> <li>• Mathematical programming</li> <li>• Logic-based models</li> <li>• Simulation</li> </ul>
Perspective	From Retrospective to Prospective		

Prescriptive analytics typically involves two aspects: (a) exploration of possible actions and (b) generation of the prescription. Typically, the decision space for prescriptive analytics tends to be large and there are multiple situations with many variables, options and constraints. Compared to descriptive and predictive, prescriptive analytics is still less mature. Recently, however, prescriptive analytics has been considered as the next step towards increasing data analytics maturity and leading to optimized decision making, ahead of time, for business performance improvement [14, 22].

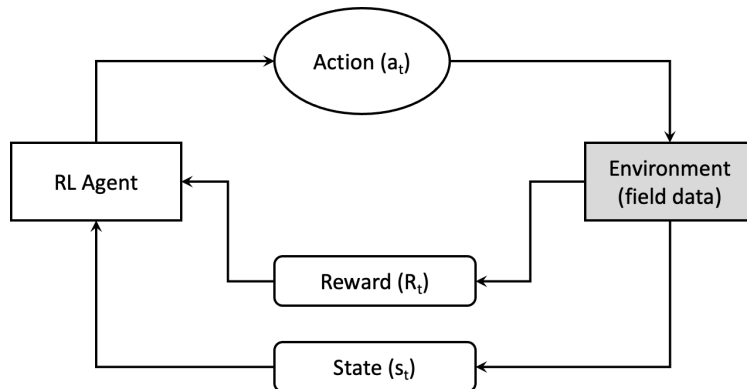
**Fig. 1.** Simplified diagram of prescriptive analytics

Integrating predictions and prescriptions is key for the extensive adoption as well as the exploitation of the values of prescriptive analytics [23, 24]. As for any method that relies on predictions, prescriptive analytics is stochastic by nature and may generate erroneous results. We argue that, similarly to control theory, an assessment of the validity of the prescriptions can be used as a feedback mechanism to generate a control action to improve the accuracy of the new predictions and the subsequent prescriptions (see Fig. 1). In the following section, we describe how such a feedback mechanism can be implemented using RL.

### 3 Reinforcement Learning for Decision Making

Reinforcement Learning (RL) is one of the three types of machine learning (the other two being supervised and unsupervised learning) [25]. In reinforcement learning the problem is represented by an environment consisting of states and actions and learning agents with a defined goal state. The agents aim to reach the goal state while maximizing the rewards by selecting actions and moving to different states [5, 6].

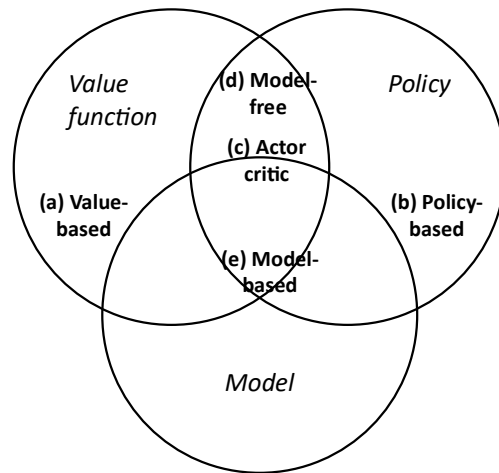
In RL, an agent learns how to perform a sequential decision task, i.e., a policy that decides which action to take in a state of the environment the agent encounters. A sequential decision task is modelled as a Markov Decision Process (MDP), denoted as  $\{S, A, T, R, \gamma\}$ . In MDP,  $S$  represents a set of all possible states and  $A$  represents a set of all possible actions. Time is divided into discrete time steps. At each time step  $t$ , the agent receives a representation of the environmental state,  $s_t \in S$ , takes an action  $a_t \in A$  that results in next state of the environment  $s_{t+1}$ . One time step later, as a consequence of the action at taken based on the current state  $s_t$ , the agent will receive a numerical reward,  $R_{t+1}$  specified by a reward function which decides a numeric reward value at each time step based on the current state, action chosen, and the resultant next state. The probability of next state  $s_{t+1}$  that the agent will experience is decided by a transition function which describes the probability of transitioning from one state to another given a specific action (see **Fig. 2** for the typical RL approach).



**Fig. 2.** The typical reinforcement learning approach.

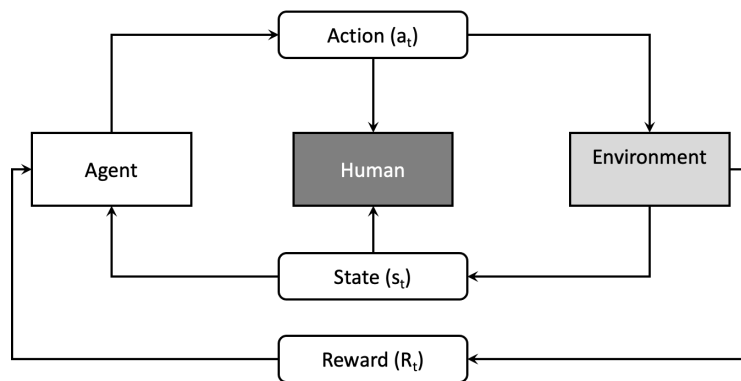
What makes RL different from other machine learning paradigms is that there is no supervisor, only a reward signal. Feedback is delayed, not instantaneous and time really matters. The agent's actions affect the subsequent data it receives. An RL agent may include one or more of these components: (i) a *policy*: it is the agent's behavior function, a map from states to actions – policies may be deterministic or stochastic; (ii) a *value function*, which models how good is each state, acts as a predictor of future reward and is therefore used to select between actions; and/or (iii) an *action model*, which is the agent's representation of the environment and predicts what the environment will do next [26].

The combinations of these three components can be used to categorize RL agents into five types: (a) value based (there is no policy, or the policy is implicit and there is a value function); (b) policy-based (there is a policy and no value function); (c) actor critic (there is both a policy and a value function); (d) model-free (there may be a policy and or a value function but no model); and (e) model-based (there may be a policy and or a value function and there is a model - see **Fig. 3**). As we will see later, our proposed framework adopts the actor-critic type of RL.



**Fig. 3.** Venn diagram of the space of agents in reinforcement learning (source [27])

In real-world applications humans may change the RL agents' optimal behavior by teaching them interactively according to their likings. In this case, standard RL cannot be applied, since the optimal behavior is usually preprogrammed via a reward function and most human users are laymen in agent design and programming. Interactive RL has been developed and proven to be a powerful method for facilitating humans to teach agents in a natural way [18,28].

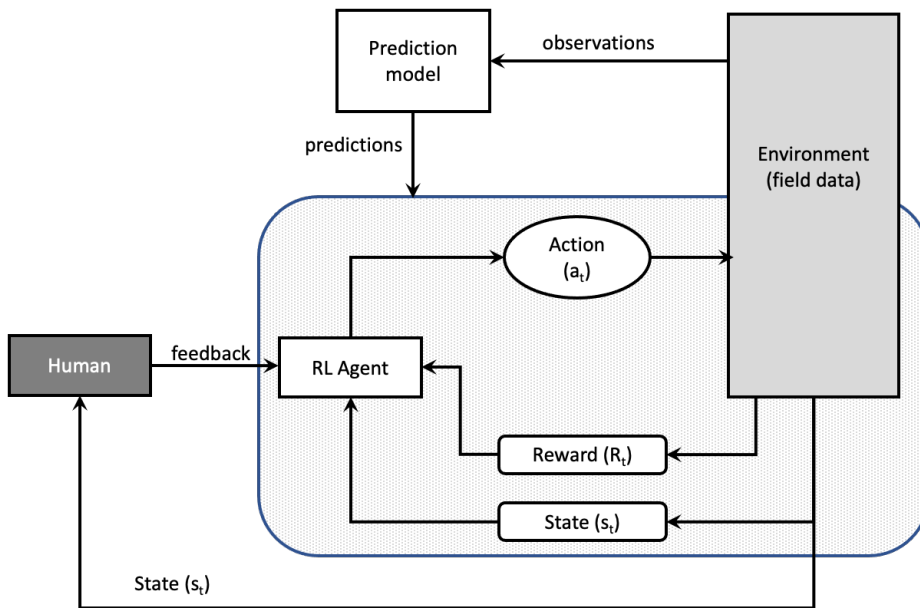


**Fig. 4.** Interactive reinforcement learning framework.

A human user might not be an expert in programming but has knowledge about how to perform the decision task, which will reduce the agent's exploration time and speed up its learning. In interactive RL, every time the agent takes an action in a state, the observing human can provide feedback which tells the quality of the selected action based on the human's knowledge, as shown in **Fig. 4**. The agent then uses the feedback to update its policy. Therefore, the agent learns how to perform the decision task by interacting with a human, and it is the human feedback that decides the agent's behaviour [29, 30].

#### 4 Framework for Collaborative Human-AI Decision Making

The proposed framework for collaborative human-AI decision uses the principles of interactive RL to implement the prescriptive model. **Fig. 5** shows how the integrated RL mechanism uses predictions generated by the prediction model as well as human feedback to guide the RL agent to provide actions for the environment. The RL agent is triggered by prediction events about future states of the environment which are generated by the prediction model. Subsequently, the agent prescribes the appropriate actions or set of actions among the alternatives so that an undesired future state is avoided, or a desired future state is sought after. Finally, the agent adjusts its policy based on both the rewards and the human feedback received.



**Fig. 5.** Proposed framework for data analytics and interactive reinforcement learning.

To gain a better understanding of the proposed framework function, in the following we examine three critical design choices and their implications: (i) the choice of the



predictive model to be used; (ii) the choice of the type of RL agent; and (iii) the type of human-AI interactions.

The first design choice refers to the type of predictive model to be used and the associated algorithms. Clearly the selection of the predictive model and algorithms depends largely on the problem at hand. Among prominent methods, Deep Learning (DL) is increasingly being used to make accurate predictions based on learned representations of data with multiple levels of abstraction [32]. An example of DL is Long Short-Term Memory (LSTM) networks. LSTM are a type of recurrent neural networks capable of learning order dependence in sequential prediction problems [33]. LSTM can learn to bridge minimal time lags and, hence, have proven to be effective when the input events are real-time streaming data, a case which is relatively typical in data-driven problems [34].

The second design choice refers to the type of the RL agent. In our framework we adopt actor-critic reinforcement learning [35]. The idea of the critic-only approach is to learn a value function based on which the agent can compare ("criticize") the expected outcomes of different actions. In the actor-only approach the agent senses the state of the environment and acts directly, i.e., without computing and comparing the expected outcomes of different actions, hence the agent learns a direct mapping (a policy) from states to actions. In contrast to these two approaches, the key idea of the actor-critic approach is to simultaneously use an actor, which determines the agent's action given the current state of the environment, and a critic, which judges the selected action. The actor learns a parameterized policy, and the critic learns a value function to evaluate state-action pairs [35]. Although learning the policy depends on the quality of the value function estimate (which is learned simultaneously by the critic), our preference for the actor-critic algorithm is due to the fact that it provides the policy necessary for the policy shaping method that consists our choice in the third design issue.

The third design issue refers to the type of human-agent interaction. Human interaction in RL algorithms has been classified in five categories: (i) standard imitation learning, in which the human trainer observes the state information and demonstrates action to the agent; (ii) learning from evaluative feedback, where the human trainer watches the agent performing the task and provides instant feedback on each agent decision; (iii) imitation from observation, which is similar to standard imitation learning except that the agent does not have access to human demonstrated action; (iv) learning attention from human, which requires the trainer to provide attention map to the learning agent; and (v) learning from human preference, in which the human watches two behaviors generated by the learning agent simultaneously and decides which is more preferable [36, 37].

Our framework adopts the learning from evaluative feedback approach. In this approach the agent adjusts its policy based on the feedback received. The simplest form of evaluative feedback is a scalar value indicating how desirable an observed action is. This approach does not require the human trainer to be an expert at performing the task - it only requires the trainer to accurately judge agent behaviors. One of the main challenges in this approach is to interpret human feedback correctly. Methods that assume different interpretations of human feedback include: (a) policy shaping, which interprets human feedback as direct policy labels; (b) reward shaping that interprets human

feedback as the value function or a human-specified reward function; (c) human intervention, in which a human supervises the training process of an RL agent and blocks catastrophic actions; and (d) policy-dependent feedback, which posits that feedback should be interpreted as an advantage function that specifies how much better or worse when deviating from the agent’s current policy.

## 5 Discussion

The recent increase of the use of AI technologies in our everyday lives has generated the need for AI systems to work synergistically with humans in an effective, transparent and ethical way [38]. However, if we are to treat AI systems in such a manner as to allow them to augment our abilities and compensate for our weaknesses, we need a new understanding of AI that takes humans explicitly into account. We need to change our view from AI systems as “thinking machines” and treat them as “cognitive prostheses” that can help humans think and act better [39, 40].

Our proposal of a framework for collaborative human-AI decision making is towards this direction. It aims to exploit the predictive power of data-driven predictive analytics techniques with the decision-making approach of reinforcement learning with human feedback. Our framework can be considered a specific instance of the hybrid structure of the human and AI-based decision making combinations described in [41]; actually it can be categorized in the “Hybrid 1: AI to human” category, in which the AI agent works in tandem with the human by providing algorithmic prediction results, which are then used as input to a collaborative human-AI decision-making process.

The proposed framework outlines an approach towards the combination of predictive and prescriptive analytics, which – according to analysts’ reports [42] – will allow organizations to reap significant business benefits. However, there are also critical costs to be considered, chief among them the cost of acquiring and annotating data. The appropriate datasets for predictive tasks have recently become key corporate assets. The issue of sample efficiency is especially important when reinforcement learning is used; if the data samples are limited in numbers the learning algorithm would have very low exploration of the environment from training data as the information and states are not fully covered. To handle this issue, the efficiency of data should be good enough or the RL approach should be sample efficient to learn the environment from limited amount of data [26]. An additional characteristic, that may increase the cost of datasets, is the need to guarantee and continuously assess data quality [43, 44].

The design and development of human-AI systems may demonstrate unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous. There is therefore a clear need to address the human computer interaction (HCI) issues of human-AI systems in a systematic manner. The recent advances generate a stream of challenges and opportunities for the HCI community. Although recent efforts have already developed reusable design guidelines [45], many challenges still persist in designing and innovating valuable human-AI interactions. Topics such as the uncertainty surrounding AI’s capabilities and the complexity of AI’s output create yet unsolved design challenges [46].

The way human users are involved in a direct collaboration with RL agents is an additional topic that needs special attention. Interactive RL approaches need to be designed in such a manner as to produce behaviors that align with the user’s intention and allow clear communication between the human and the RL algorithm. For example, in order to motivate users to give feedback several studies suggest the use of gamification strategies [18].

Furthermore, the fatigue of users and its effects on the quantity and quality of feedback should be considered; for example, it has been observed that humans tend to reduce the quantity of feedback they give over time while the quality also diminishes. Developing appropriate design strategies to address the engagement of humans becomes paramount. Finally, the knowledge level of the human involved in the interaction is an important feature of the human-AI interactions. This level has an impact on the quality and quantity of feedback and could limit the use of some feedback types that require more precise information, such as demonstrations or action advice [18].

The practical real-world challenges of implementing reinforcement learning algorithm constitute another area of concern. Topics such as the ones related to the exact definition of reward functions, the number of actions and state spaces, as well as the data sample efficiency (already mentioned above) have significant impact on the performance of RL algorithms in real-world applications [47]. In addition, the environment in real-world application domains may allow only partial observation. Conventional RL methods assume that environments are fully observable Markov environments. Partial observability requires the generalization of Markov decision processes to partially observable ones, which increases computational complexity.

Of course, the more general issues of human-AI interactions also apply in our case. Topics like the need for governance structures that address AI failures; the need for explaining AI recommendations; and the issues of trust, transparency and reliability are still open research questions in the organization science, information systems and artificial intelligence fields [48, 49].

## 6 Conclusions and further work

In the present paper we developed a conceptual framework that integrates the data-driven predictions of predictive analytics with the decision-making process of prescriptive analytics. The framework explicitly takes into account human experience using the mechanisms of interactive reinforcement learning. Early implementations of the framework in application domains like financial management (stock trading) and industry 4.0 (predictive maintenance in steel industry) provide encouraging results.

In our further work we intend to address problems that need to solve several tasks with different rewards simultaneously. For such problems, we aim to experiment with multi-objective RL (MORL). MORL can be viewed as the combination of multi-objective optimization (MOO) and RL to solve decision-making problems with multiple conflicting objectives [50]. Moreover, in order to facilitate learning the reward function from demonstrations by human experts, we aim to experiment with inverse RL [51], which attempts to extract the reward function from the observed behavior of an agent.

**Acknowledgements.** This work is partly funded by the European Union's Horizon 2020 project COALA (Grant agreement No 957296). The work presented here reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

## References

1. Johnson, J. G., & Busemeyer, J. R. "Decision making under risk and uncertainty", Wiley Interdisciplinary Reviews: Cognitive Science, 1(5), 736-749 (2010).
2. Von Neumann, John, and O. Morgenstern. "Theory of Games and Economic Behaviour Princeton Univ." Press, Princeton (1944).
3. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–292 (1979).
4. Lee, Daeyeol. "Decision making: from neuroscience to psychiatry." *Neuron* 78, no. 2, 233-248 (2013).
5. Sutton, Richard S., and Andrew G. Barto. "Reinforcement learning: an introduction" Cambridge, MA, MIT Press (2018).
6. Kaelbling, Leslie Pack, Michael L. Littman, and Andrew W. Moore. "Reinforcement learning: A survey." *Journal of artificial intelligence research* 4 (1996): 237-285.
7. Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves et al. "Human-level control through deep reinforcement learning." *Nature* 518, no. 7540, 529-533 (2015).
8. Davenport, Thomas H. "Competing on analytics." *Harvard business review* 84, no. 1 (2006).
9. LaValle, Steve, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, and Nina Kruschwitz. "Big data, analytics and the path from insights to value." *MIT sloan management review* 52, no. 2, 21-32 (2011).
10. Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. "Business intelligence and analytics: From big data to big impact." *MIS quarterly*, 1165-1188 (2012).
11. Lepenioti, Katerina, Alexandros Bousdekis, Dimitris Apostolou, and Gregoris Mentzas. "Prescriptive analytics: Literature review and research challenges." *International Journal of Information Management* 50, 57-70 (2020).
12. Bertsimas, Dimitris, and Nathan Kallus. "From predictive to prescriptive analytics." *Management Science* 66, no. 3, 1025-1044 (2020).
13. Frazzetto, Davide, Thomas Dyhre Nielsen, Torben Bach Pedersen, and Laurynas Šikšnys. "Prescriptive analytics: a survey of emerging trends and technologies." *The VLDB Journal* 28, no. 4, 575-595 (2019).
14. Sappelli, Maya, Maaike HT de Boer, Selmar K. Smit, and Freek Bomhof. "A Vision on Prescriptive Analytics." *ALLDATA 2017*, 54 (2017).
15. Neftci, Emre O., and Bruno B. Averbeck. "Reinforcement learning in artificial and biological systems." *Nature Machine Intelligence* 1, no. 3, 133-143 (2019).
16. Collins, Anne Gabrielle Eva. "Reinforcement learning: bringing together computation and cognition." *Current Opinion in Behavioral Sciences* 29, 63-68 (2019).
17. Amershi, Saleema, Maya Cakmak, William Bradley Knox, and Todd Kulesza. "Power to the people: The role of humans in interactive machine learning." *Ai Magazine* 35, no. 4, 105-120 (2014).
18. Arzate Cruz, Christian, and Takeo Igarashi. "A survey on interactive reinforcement learning: Design principles and open challenges." In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, pp. 1195-1209 (2020).

19. Li, Guangliang, Randy Gomez, Keisuke Nakamura, and Bo He. "Human-centered reinforcement learning: a survey." *IEEE Transactions on Human-Machine Systems* 49, no. 4, 337-349 (2019)
20. Holsapple, Clyde, Anita Lee-Post, and Ram Pakath. "A unified foundation for business analytics." *Decision Support Systems* 64, 130-141 (2014).
21. Delen, Dursun: *Prescriptive Analytics: The Final Frontier for Evidence-Based Management and Optimal Decision Making*. FT Press (2019).
22. Larson, Deanne, and Victor Chang. "A review and future direction of agile, business intelligence, analytics and data science." *International Journal of Information Management* 36, no. 5, 700-710 (2016).
23. den Hertog, Dick, and Krzysztof Postek. "Bridging the gap between predictive and prescriptive analytics-new optimization methodology needed." Tilburg Univ., Tilburg, The Netherlands (2016) [http://www.optimization-online.org/DB\\_FILE/2016/12/5779.pdf](http://www.optimization-online.org/DB_FILE/2016/12/5779.pdf), last accessed 2020/12/28.
24. Mundru, Nishanth. "Predictive and prescriptive methods in operations research and machine learning: an optimization approach." PhD diss., Massachusetts Institute of Technology (2019) <https://dspace.mit.edu/handle/1721.1/122099>, last accessed 2020/12/28.
25. Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." *Science* 349, no. 6245, 255-260 (2015).
26. Naeem, Muddasar, S. Tahir H. Rizvi, and Antonio Coronato. "A Gentle Introduction to Reinforcement Learning and its Application in Different Fields." *IEEE Access* (2020).
27. Silver, D. "Lectures on Reinforcement Learning", <https://www.davidsilver.uk/teaching/> (2015), last accessed 2020/12/28.
28. Griffith, Shane, Kaushik Subramanian, Jonathan Scholz, Charles L. Isbell, and Andrea L. Thomaz. "Policy shaping: Integrating human feedback with reinforcement learning." *Advances in neural information processing systems* 26: 2625-2633 (2013).
29. Li, Guangliang, Randy Gomez, Keisuke Nakamura, and Bo He. "Human-centered reinforcement learning: a survey." *IEEE Transactions on Human-Machine Systems* 49, no. 4, 337-349 (2019).
30. Lin, Jinying, Zhen Ma, Randy Gomez, Keisuke Nakamura, Bo He, and Guangliang Li: "A Review on Interactive Reinforcement Learning from Human Social Feedback." *IEEE Access* 8, 120757-120765 (2020).
31. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521, no. 7553 (2015): 436-444.
32. Najafabadi, Maryam M., Flavio Villanustre, Taghi M. Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic. "Deep learning applications and challenges in big data analytics." *Journal of Big Data* 2, no. 1 (2015)
33. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9, no. 8, 1735-1780 (1997).
34. Yu, Yong, Xiaosheng Si, Changhua Hu, and Jianxun Zhang. "A review of recurrent neural networks: LSTM cells and network architectures." *Neural computation* 31, no. 7, 1235-1270 (2019).
35. Konda, Vijay R., and John N. Tsitsiklis. "Actor-critic algorithms." In *Advances in neural information processing systems*, pp. 1008-1014 (2000).
36. Grondman, Ivo, Lucian Busoniu, Gabriel AD Lopes, and Robert Babuska. "A survey of actor-critic reinforcement learning: Standard and natural policy gradients." *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 42, no. 6, 1291-1307 (2012).

37. Ruohan Zhang, Faraz Torabi, Lin Guan, Dana H. Ballard, Peter Stone: "Leveraging Human Guidance for Deep Reinforcement Learning Tasks", Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, Survey track. 6339-6346 (2019).
38. Kambhampati, Subbarao. "Challenges of Human-Aware AI Systems." *AI Magazine* 41, no. 3 (2020).
39. Akata, Zeynep, Dan Balliet, Maarten De Rijke, Frank Dignum, Virginia Dignum, Gusztai Eiben, Antske Fokkens et al. "A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence." *Computer* 53, no. 8, 18-28 (2020).
40. Trunk, Anna, Hendrik Birkel, and Evi Hartmann. "On the current state of combining human and artificial intelligence for strategic organizational decision making." *Business Research*, 1-45 (2020).
41. Shrestha, Yash Raj, Shiko M. Ben-Menahem, and Georg Von Krogh. "Organizational decision-making structures in the age of artificial intelligence." *California Management Review* 61, no. 4, 66-83 (2019)
42. Gartner Inc.: "When and How to Combine Predictive and Prescriptive Techniques to Solve Business Problems", Published: 25 October 2018 ID: G00368423 (2018)
43. Wang, Richard Y., and Diane M. Strong. "Beyond accuracy: What data quality means to data consumers." *Journal of management information systems* 12, no. 4 , 5-33 (1996).
44. Cai, Li, and Yangyong Zhu. "The challenges of data quality and data quality assessment in the big data era." *Data science journal* 14 (2015).
45. Amershi, Saleema, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh et al. "Guidelines for human-AI interaction." In Proceedings of the 2019 chi conference on human factors in computing systems, pp. 1-13. (2019).
46. Yang, Qian, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. "Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design." In Proceedings of the 2020 CHI conference on human factors in computing systems, 1-13 (2020).
47. Gabriel Dulac-Arnold, Daniel Mankowitz, Todd Hester: "Challenges of Real-World Reinforcement Learning", ICML 2019 Workshop, Reinforcement Learning for Real Life, June 14, 2019, Long Beach, CA, USA (2019).
48. Glikson, Ella, and Anita Williams Woolley. "Human trust in Artificial Intelligence: Review of empirical research." *Academy of Management Annals* ja (2020).
49. Samek, Wojciech, Grégoire Montavon, Andrea Vedaldi, Lars Kai Hansen, and Klaus-Robert Müller, eds. *Explainable AI: interpreting, explaining and visualizing deep learning*. Vol. 11700. Springer Nature (2019).
50. Liu, Chunming, Xin Xu, and Dewen Hu. "Multiobjective reinforcement learning: A comprehensive overview." *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, no. 3, 385-398 (2014).
51. Zhifei, Shao, and Er Meng Joo. "A review of inverse reinforcement learning theory and recent advances." In 2012 IEEE congress on evolutionary computation, pp. 1-8. IEEE, 2012.