



## informs. http://pubsonline.informs.org/journal/orsc

#### ORGANIZATION SCIENCE

Articles in Advance, pp. 1-25 ISSN 1047-7039 (print), ISSN 1526-5455 (online)

## Algorithm Supported Induction for Building Theory: How Can We Use Prediction Models to Theorize?

Yash Raj Shrestha, Vivianna Fang He, Phanish Puranam, Georg von Krogh\*

Received: April 30, 2019 Revised: July 15, 2019 Accepted: May 11, 2020

Published Online in Articles in Advance:

December 9, 2020

https://doi.org/10.1287/orsc.2020.1382

Copyright: © 2020 The Author(s)

Abstract. Across many fields of social science, machine learning (ML) algorithms are rapidly advancing research as tools to support traditional hypothesis testing research (e.g., through data reduction and automation of data coding or for improving matching on observable features of a phenomenon or constructing instrumental variables). In this paper, we argue that researchers are yet to recognize the value of ML techniques for theory building from data. This may be in part because of scholars' inherent distaste for predictions without explanations that ML algorithms are known to produce. However, precisely because of this property, we argue that ML techniques can be very useful in theory construction



Department of Management, Technology, and Economics, ETH Zürich, Zurich CH 8092, Switzerland; Management Department, École Supérieure des Sciences Economiques et Commerciales (ESSEC) Business School, 95021 Cergy-Pontoise Cedex, France;

Strategy Department, INSEAD, Singapore, Singapore 138676

#### RESEARCH ARTICLE





ORGANIZATION SCIENCE

#### Algorithm Supported Induction for Building Theory: How Can We Use Prediction Models to Theorize?

Task Rig Streetin," Volume Fung Inc." Phonos Purposes," Georg con Eroph"

"Department of Miningament, Switzerings, and Sciencesis, ESSE Farch, Farch (SS 400), Switzerland, "Miningament Department, Sand Suprimed the Science Sciences Sciences and Assessments (SSSE) Statement School, 1962 (1982) Suprimed Science, Science, Sciences Sciences Anni Sciences Sciences Sciences Sciences Sciences Sciences Sciences Sciences Sciences Sciences

hand plantation at the large contact with the little (1975) the following size the classic and any find that the little beauty processes and the large contact and any territorial and other processes and the large contact and any territorial and other processes and

Restrict Get 20, 2019 Revised, July S. 2019 Accepted No. 11, 2020 Accepted No. 11, 2020 Published Dates in Million in Advance. Discovery 1, 2022

Marchany's (Whoulder) Supply 1 DD for below Biotester, Access tomar Steller of sealed notices, smarkes bearing (MS) algorithms on appared advantage towards included by complete advantage to the property of the college of the property of the advantage and automation of data coding or for property, marketing or observed in the college of the property of the college of the college of the property of the college of the colleg

### Resolving governance disputes in communities: A study of software license decisions

Vivianna Fang He<sup>1</sup> | Phanish Puranam<sup>2</sup> Yash Raj Shrestha<sup>1</sup> | Georg von Krogh<sup>1</sup>

<sup>1</sup>ETH Zurich, Zurich, Switzerland <sup>2</sup>INSEAD, Singapore, Singapore

#### Correspondence

Vivianna Fang He, ETH Zurich, Weinbergstrasse 56-58, Zurich CH 8092, Switzerland.

Email: fhe@ethz.ch

#### **Funding information**

Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 169441

#### Abstract

Research summary: Resolving governance disputes is of vital importance for communities. Gathering data from GitHub communities, we employ hybrid inductive methods to study discussions around initiation and change of software licenses—a fundamental and potentially contentious governance issue. First, we apply machine learning algorithms to identify robust patterns in data: resolution is more likely in larger discussion groups and in projects without a license compared to those with a license. Second, we analyze textual data to explain the causal mechanisms underpinning these patterns. The resulting theory highlights the group process (reflective agency switches disputes from bargaining to problem solving) and group property (preference align-



### Agenda

What is Artificial Intelligence (AI)?

Al and Strategic Analysis

Al and Strategy Formulation and Implementation

**Nexus of Strategizing and AI: Future Research Imperatives** 

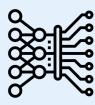
Challenges and Risks for Al and Strategizing

### What is Artificial Intelligence (AI)?

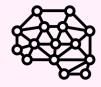


**Artificial Intelligence (AI)** «involves machines that can perform tasks that are characteristic of human intelligence» John McCarthy (1956).

There is general and narrow AI



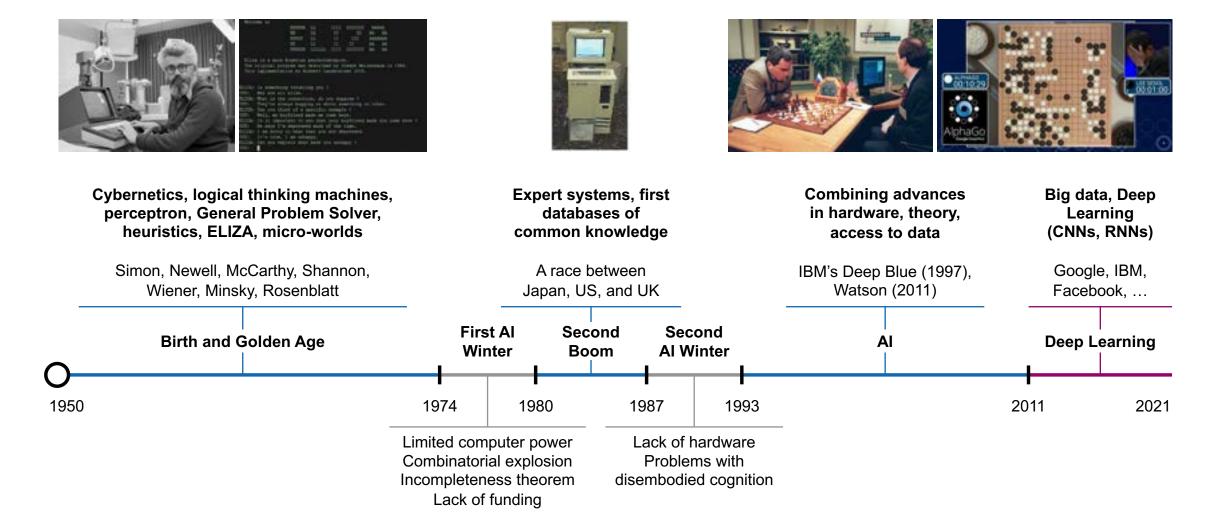
Machine Learning (ML) is «the ability to learning without being explicitly programmed» Arthur Samuel (1959).



**Deep Learning (DL)** was inspired by the structure of the brain, namely the interconnecting of many neurons.



### A brief history of Artificial Intelligence





6

### Types of Machine Learning Algorithms

#### **Machine Learning**



#### **Supervised ML**

Task Driven

**Predicts Next Value** 

Applications: Demand forecasting and forecasting investment returns



### **Unsupervised ML**

Data Driven

Identifies Patterns in Data

Applications: Strategic group analysis



### **Reinforcement Learning**

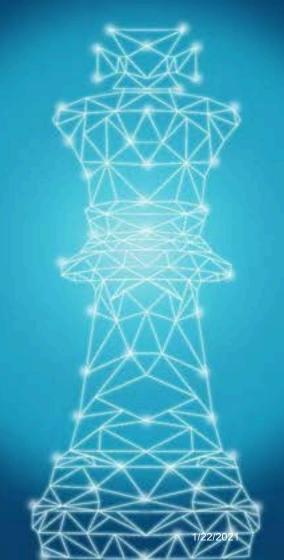
Simulation Driven

Learns from Mistakes

Applications: Evaluating various alternatives and their counterfactuals



# Al and Strategizing

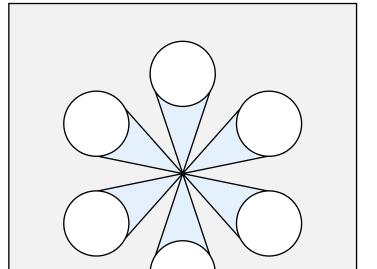




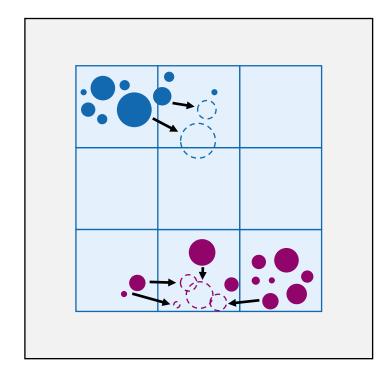
### ! "#\$%&# ' ()\$(\*+,-#! %\$./0,0

! "#\$.+3),(260#6\$/#2\$7\*#(2\*#@3(\*%(,\$.#43)#&)\$6\$(,-\$../#!"#\$%&'%()\*+!),-\$.!/),-0#!/)\$%&),#!!&)01)
,\*2\$\*!('-)\$%\$.3,',(2)3A+2#,(0#-\$@\$-,(/#43)#,%(\*)\$-(,%+#8,(2#(2\*#\*%7,)3%6\*%(#\$%&#+\*%\*)\$(,%+#&\$(\$9#

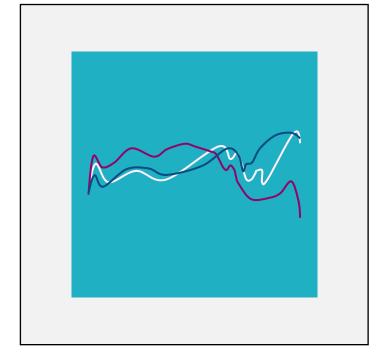
9"\*!2%\$.)5%\$.3,',



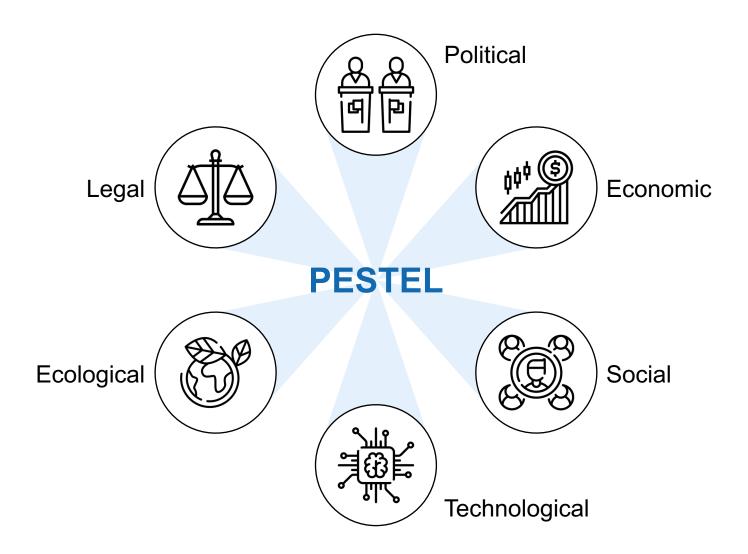
607#!\*'\*'8!)5%\$.3,',



4%\*!2%\$.)5%\$.3,',



## **External Analysis**





### External Analysis – Political Factors



#### Machine Learning Algorithms are used to ...

Infer "political climate" across regions through Al-based text analysis tools

Identify patterns in news outlets, legislative debates and online political discourse

Predict outcomes of elections (e.g., Coletto et al., 2015), policy changes (e.g., Chan & Zhong, 2018), and political bias and conflicts (Biessmann et al., 2016)



## External Analysis – Economic Factors



#### Machine Learning Algorithms are used to ...

Measure economic trends, such as economic growth, the onset of economic recessions (e.g., Wu et al., 2020), increasing poverty (Kshirsagar et al., 2017), and bankruptcies (Cielen, Peeters, & Vanhoof, 2004)

Predict stock returns and thereby make better investment decisions (Avramov, Cheng, & Metzker, 2019)

Estimate systemic financial risks (e.g., Kou et al., 2019)



### External Analysis – Social Factors



### Machine Learning Algorithms are used to ...

"map the contours of cultural fields, classify cultural elements and trace the evolution of culture over time" (Bail, 2014)

Enable the systematic measurement of culture and modeling of its evolution within organizations and social groups (Doyle et al., 2017)



## External Analysis – Technological Factors



#### Machine Learning Algorithms are used to ...

Assist companies in monitoring technological developments and anticipate any relevant technological changes

Identify patterns of technology development by sifting through massive amounts of patent or publication data (e.g., Lee et al., 2018)

Construct "knowledge profiles" of their industry and major competitors (Suominen, Toivanen, Seppänen, 2017).



### External Analysis – Ecological Factors



#### Machine Learning Algorithms are used to ...

Research into whether hydrogen electrical vehicles may become the dominant means of transportation for consumers (Ranaei et al., 2016)



## External Analysis – Legal Factors



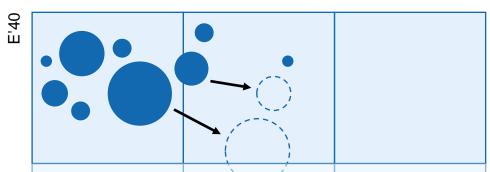
### Machine Learning Algorithms are used to ...

Support firm internal collection and processing of legal data, potentially reducing overall legal expenditures (e.g., Yousfi-Monod, Farzindar, and Lapalme 2010)

Automate financial compliance monitoring and regulation (e.g., Treleaven & Batrinca, 2017)

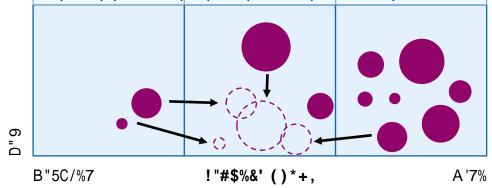


## ! "#\$%&'&'(%)\*+,-./'/



6 D), -4 "1'&0 # /)5, +)F%)C/%7)&")-\$, \* '-./01'")', +-&0+"#%21 5"+/'/&'+4)"2)2'1 # /)&0, &)2"--"9)/'#'-, 1)
/&1, &%4'%/6), +7), -/")2", \$-&'03 #4, 3, \* '10#.0.-"3105, '6, , \*0'7, 1, 0+"#%210H3O1%/&0,)%&), -:)>?>?I:

\*-4"1'&0#/)5,+)\$%12"1#)5"#\$%&'&"1),+,-./'/)"+),)5"+&'+C"C/)F,/'/),+7)0%-\$)/&1,&%4'/&/)7.+,#'5,--.)
4,C4%)&O%)(,-'7'&.)"2)&O%)2'1#J/)5C11%+&)/&1,&%4.:



### Internal Analysis – Al in Human Resources

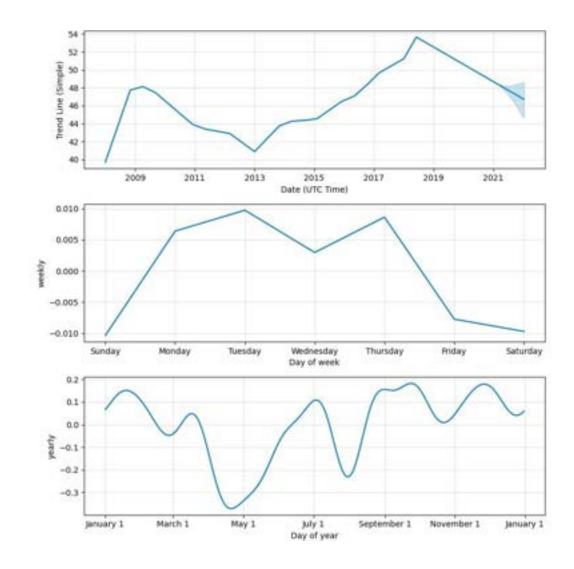
Al may also augment strategic analysis of internal organizational factors, including human, financial, and auxiliary resources, such as supply chains.

Al assist managers in identifying employee performance, predicting career trajectories, and revealing patterns of compensation and inequality, among others (Strohmeier & Piazza, 2013).



### Internal Analysis – Accounting and Finance

Algorithms may boost the interpretation of what causes certain financial resource conditions (e.g., fluctuations in liquidity, exchange rate premiums at corporate treasuries) relevant to the type and timing of the firm's strategic commitments (Fethi & Pasiouras, 2010).





### Internal Analysis – Planning and Forecasting

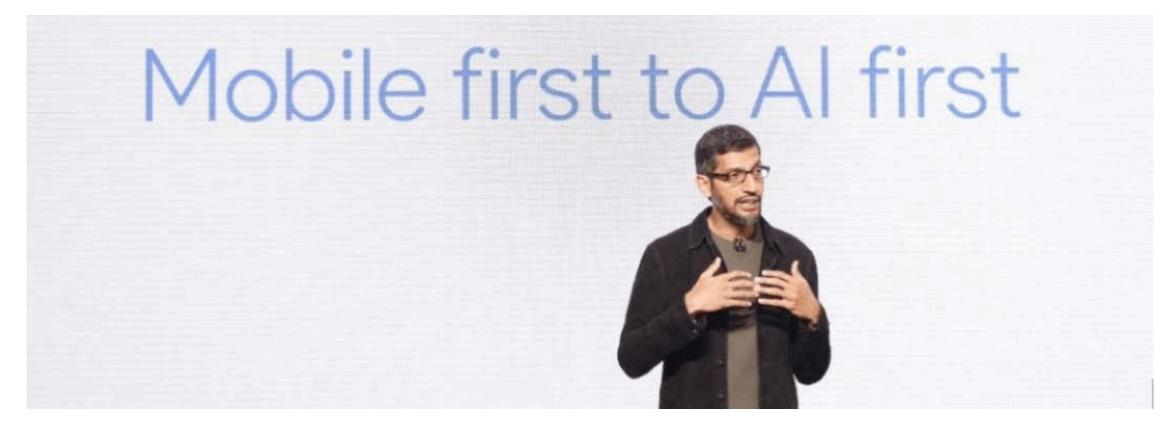
ML algorithms are increasingly providing efficient analysis on other resources relevant to the strategy process, such as demand forecasting, production planning, resource allocation, and logistics.

ML can augment production planning decisions by automating the process of searching for potential suppliers by mining data from online catalogues and other repositories (Nissen & Sengupta, 2006), providing predictions on performance of prospective suppliers (Humphreys, McIvor, & Huang, 2002), and even estimating valuation and evaluating online bids (Cheung et al., 2004).



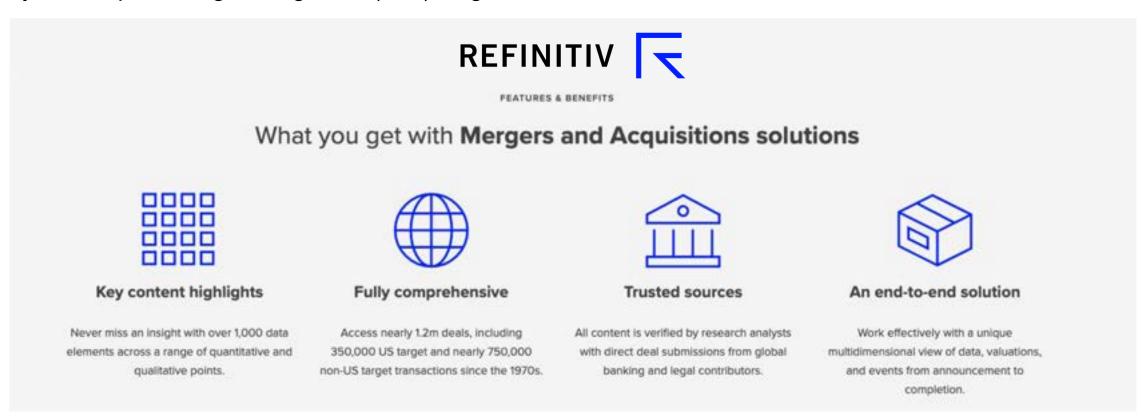
#### Al at the center of high-level strategy

Al offers applications in the generation, evaluation, and selection of strategic options and choices.



#### **Al and Corporate Strategy**

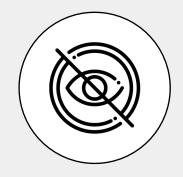
Refinitiv uses AI to identify and screen *merger and acquisition* opportunities, as well as manage the deal cycle and post-merger integration (PMI) stage.



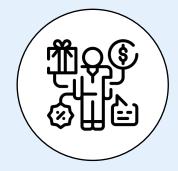
22

#### Strategic Innovation, Entrepreneurship and Renewal

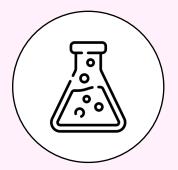
Cockburn, Henderson, and Stern (2018) argue that AI serve as a new "method of inventing" with the potential to re-shape the nature of the innovation process. Chalmers, MacKenzie, and Carter (2020) identify three ways in which AI can enhance information search and idea generation activities:



Deep-learning allows to search for and experiment with previously unobservable opportunities (Repurposing)



Al can support new venture activity through the identification and exploitation of consumer needs

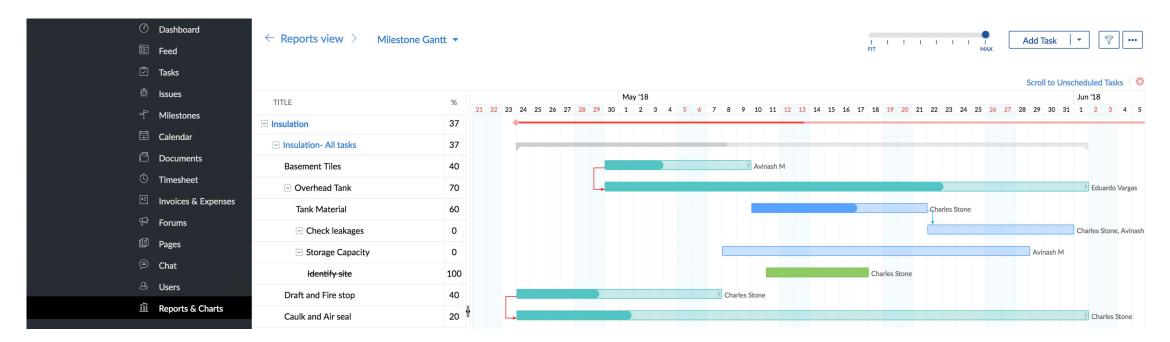


Al-based simulations and promise the ability to test innovations and new venture ideas

#### **Strategy Control**

Strategy control concerns the efforts of strategists to ensure that the implementation of strategic plans unfolds as intended and to measure progress on strategic goals.

Research on AI in strategy control has been limited thus far. Nonetheless, promising areas can be identified – particularly compelling is the **use of AI project management and internal communication**.



### Nexus of Strategizing and AI: Future Research Imperatives

#### Al is no Magic Bullet

Al may offer many attractive functions for strategists in orchestrating strategic processes. At the same time, there is a particular need for scholars to take a prudent approach to Al and attempt to identify boundaries as well as challenges in implementing such systems in organizations.



It is key for scholars and strategists alike to understand that there is **no magic bullet**, **single solution**, **or best practice for Al applications in strategy processes**. The introduction of Al needs to be an iterative process that is based on trial-and-error learning



#### **Data Privacy and Security**



#### Reduced search costs

The application of AI requires vast amounts of personal and interactional data, which often interferes with individual privacy and data security. Firms should be highly cautious and aware of potential risks when using such data for strategic analysis – in particular, when using customer, employee, or client data for strategic analysis.



#### Machine learning bias







#### **Data Privacy and Security**



#### Reduced search costs



### Machine learning bias

With the reduced search costs in information acquisition, organizations have relatively similar access to information, reducing information asymmetry (Goldfarb & Tucker, 2019). Such lowered costs could have **potentially mixed**effects. On the one hand, they make it easier to rapidly find rare and potentially valuable strategic decision alternatives (Yang, 2013; Zhang et al., 2018). On the other, search algorithms represent popularity bias, resulting in lack of variety with respect to information acquisition (Fleder & Hosanagar, 2009).





#### **Data Privacy and Security**



#### Reduced search costs

ML algorithms have been shown to produce results that are **systematically prejudiced with respect to gender and ethnicity**, among other factors. There are valid concerns that use of algorithms amplifies and perpetuates social inequalities (Barocas, Selbst, 2016; Starr, 2014).



#### Machine learning bias







#### **Data Privacy and Security**



#### Reduced search costs

Algorithms often feed off each other's behavior.

Output from one algorithm often impacts the learning of another. This interconnectedness may sometimes lead to erratic behavior as we have seen in the case of high-frequency trading, where algorithms have put significant financial assets at risk.



#### Machine learning bias







Professor Georg von Krogh gvkrogh@ethz.ch

ETH Zurich
Chair of Strategic Management and Innovation
WEV J 411
Weinbergstr. 56/58
8092 Zürich, Switzerland

www.smi.ethz.ch