RAIS RESEARCH ASSOCIATION for INTERDISCIPLINARY STUDIES
June 6-7, 2024

10.5281/zenodo.12684418

Knowledge in the 21st Century: Making Sense of Big Data

Julia M. Puaschunder

Columbia University, 116th and Broadway, New York, NY 10027, United States,
Julia.Puaschunder@columbia.edu, https://blogs.cuit.columbia.edu/jmp2265/
Aix-Marseille University, Aix-Marseille School of Economics, Julia.Puaschunder@emle.eu,
https://juliampuaschunder.com/

ABSTRACT: Economics is built on the idea of scarcity. The rational agent is meant to decide efficiently given scarce resources. Friedrich von Hayek challenged mainstream economic scarcity focus by proposing that economics is also about making sense of an abundance of knowledge that is dispersedly shared throughout society. Born out of the internet and digitalization, we live in the age of the advent of big data science. Computational advantages in light of enormous data storage and analytic powers have allowed to gain unprecedented information transfers out of our use of technology. Today's most profitable corporations in the world all derive value from big data insights. But extracting sense from a massive amount of data generated online on a constant basis has also become an enormous environmental burden, which is often not discussed or thematized. While marketing ensures to frame cloud storage as something light and intangible, the reality is that data hoarding has become an environmentally-burdensome practice that may not be sustainable given the pace digitalization is advancing, e.g., with 5G and Artificial Intelligence (AI) encroaching every aspect of human life. Big data is stored in facilities that resemble warehouses with enormous electricity consumption for the cooling of real-time data processors. In light of the rising trend of data storage as well as environmental conscientiousness demands at the same time, we may revisit Hayek's idea of the knowledge paradigm and connect it to scarcity. Economics is called for providing models that explain how to make sense of data efficiently. In particular, Friedrich von Hayek's knowledge paradigm could offer insights on how and what kind of information should be stored to become knowledgeable and what kind of information warrants for scarcity of being neglected to conserve and simply been forgotten over time to be in line with overall efficiency and sustainability demands to pass the earth onto future generations meaningfully and viably.

KEYWORDS: behavioral economics, carbon footprint, data storage, digitalization, discounting, environmental conscientiousness, environmentalism, internet, knowledge, law & economics, scarcity, Sustainable Development Goals (SDGs), sustainability

Introduction

Mainstream economics focusing on the scarcity problem analyzes the world as an equilibrium between forces exercised by rational individuals. The rational expectation framework, however, misses dynamics of social phenomena. A long time ago, Friedrich von Hayek (1945) argued that the problem society faces is that knowledge is not completely given to a single mind. Behavioral economics later proved that even if information is provided to human beings, they are not fully capable of working out the implications of knowledge (Kahneman 2011). Hayek (1945) pointed out that the problem economics tries to solve is to utilize knowledge that is not fully given to one single mind and is not observable in its totality. Hayek (1945) therefore advocates for not focusing on efficiency in the allocation of scarce resources in economics. But rather maximize our ability to secure the best use of resources known to any of the members of society.

This article makes the case that the time has come to revisit the knowledge and scarcity debate in light of the contemporary trend of big data analytics. Data has become the gold of our times with the most valuable companies in the world being big data processing entities. The storage and processing, however, creates an enormous environmental impact, which is not thematized. Cloud storage marketing lets data storage and processing appear as a lightweight floating task, when the reality, however, lies in deforestation for setting up data

storage centers that constantly require enormous electricity for information processing, data storage and cooling capacities.

Discussing the dichotomy of the wish for scarcity and knowledge management in terms of big data reaping entities having a significant environmental impact could be inspired by the insights of the knowledge paradigm having challenged neoclassical scarcity focus. First, the knowledge paradigm insights could aid in what information is useful and work towards a taxonomy of the value of knowledge. Second, knowledge theory could help make sense of the future value of information in order to weed out what kind of knowledge should be conserved. Thirdly, understanding scarcity as a benefit in light of cognitive capacity and environmental constraints could help make the case to forget some useless knowledge. Abandoning data hoarding for detected parts of uselessness information reaping to a degree that is harmful in terms of its environmental footprint, could be the scarcity innovation our times are calling for.

This paper is structured as follows: First the knowledge paradigm will be presented in relation to the neoclassical scarcity problem. Then the advent of data storage and processing facilities in the digital age will be captured. The knowledge paradigm and scarcity demands will then be applied to the rising data storage critique with special attention to the environmental impact of the erection of contemporary big data centers. The discussion will offer a future-oriented look into research avenues.

The knowledge paradigm in relation to scarcity

Building on Carl Menger (1871), Friedrich von Hayek (1945) challenged the classical economics notion of scarce resource allocation being the main concern of society. Modern economics, Hayek (1945) pointed out a long time ago, is all about allocating scarce resources rather than making sure that the right knowledge is used. In Hayek's (1945) view, individuals should rather be educated to decide for themselves which knowledge is useful for them.

Hayek (1945) also argues that the real economic problem society faces is that knowledge is not completely given to a single mind. Behavioral economics also proves Hayek's (1945) further point that even if we had all knowledge being represented to us, human beings will never be capable of working out the implications of knowledge fully by themselves (Kahneman 2011). As Hayek (1945) outlines, the problem economics should solve is to utilize knowledge that is not fully given to one single mind and is not observable in its totality by society. Hayek (1945) therefore advocates for not focusing on efficiency in the allocation of scarce resources. But rather maximize our ability to secure the best use of resources known to any of the members of society.

The classic knowledge problem is believed to be solved by making sense of data. Knowledge equips to perform better. Decentralized information exchange exists. In order to make sense of information spread out over the population, we cooperate, have training on the job and build scientific knowledge. People also do not need to know everything as there is an implied valuation of customers' goods in prices. Prices convey information and help society to coordinate in market economies.

Building on Friedrich von Hayek's (1945) knowledge paradigm that postulates that dispersed, incomplete and contradictory information is held by all of humankind, only a few pieces of information survive throughout history. Knowledge as dispersed, incomplete and oftentimes contradictory messages spilled out by a broad variety of diverse and separate people is similarly opaque to measure in terms of future value (Hayek, 1945). The writer of a book knows little about the success in the reception of the book by the audience, such as the painter or composer are in doubt about the future value if their arts created. Entrepreneurs are believed to be those who utilize knowledge that is not given completely to determine market dynamics better than others. Entrepreneurs are those who have a better sense about the reception of goods by the market in the pursuit of creating things that are valued by the

market with an elevated price tolerance higher than the sum of a good's production and distribution costs.

Knowledge is good for economic planning (Hayek, 1945). Knowledge puts one in a better position to decide. The more knowledge one has, the better solutions one can find. Knowledge is also the underlying feature of the discovery process of property rights, when one comes up with genius ideas. Virtus of markets are discovery processes that rely on property rights. Empirical Legal Studies address questions by making sense of data.

The knowledge paradigm is different from the prevalent neoclassical economics scarcity consideration. Neoclassical economics is built on foundations of scarce resource allocation capabilities determining better outcomes for the rational market actor. Neoclassical scarcity calculus is insofar convenient as there is a clearly measurable allocation maximization problem to be solved with mathematical models. Scarcity has become the center stage of production and consumption models in neoclassical economics. Mainstream economics focusing on the scarcity problem analyzes the world as an equilibrium between forces exercised by rational individuals.

The neoclassical rational expectation model, however, misses dynamics of social phenomena. Scarcity does not acknowledge information asymmetries in markets, which are in reality complex social phenomena. The knowledge paradigm maintains some kind of rationality assumptions but does not systematically rely on the use of mathematical formalization. Not primarily focused on a state of scarcity in the economy, the knowledge paradigm is rather concerned with the abundance of dispersed and tacit knowledge in society. The knowledge paradigm advocates for trying to solve the coordination in the use of knowledge. Making sense of information – that is not given in its entirety and not palpable by a single bounded rationality constraint mind – has become the center stage of the knowledge predicament.

In the perception, evaluation and decision model, perception is related to knowledge and evaluation to scarcity, which both lead to decision quality outcomes. The problem of perception and knowledge is that they are at least partly subjective and unexplainable to logic as behavioral economics outlines vividly (Kahneman, 2011). There is also the problem of subjective probability and utility in the knowledge problem, which is not addressed by the equilibrium theory proponents. The theory of choice and rational decision making assumes that decisions are based on perception plus the subjective evaluation of the individual. Behavioral economics addresses bounded rationality with expected utility coupled with subjective probability estimates (Simon 1986).

Contrary to the scarcity conundrum in economics, which is a closed-ended world; the knowledge paradigm is an open world assumption of knowledge breeding innovation. In an Edgeworth box description, knowledge allows to better allocate contents in the box. "The economics of knowledge and the knowledge of economics" captures this importance of knowledge (Boulding 1966). Knowledge is dispersed, partly tacit and constantly-changing in progress. The scarcity problem, which is connected to evaluation of the perception of knowledge, is valuable for the description of one individual; while the knowledge paradigm allows to make inferences of the complexity of society.

The knowledge paradigm has been prominent in describing market process that use and create tacit and dispersed knowledge. In the knowledge world, market behavior is captured being unpredictable and an extremely complex social phenomenon. According to Israel Kirzner (1973, 1985), entrepreneurship is a discovery process in the real world, where knowledge is dispersed. The entrepreneur has better knowledge and can use it in a more efficient way. Spontaneous order rules the cooperation in markets featuring decentralized decisions leading to entrepreneurial discovery and activities. Discovery is a learning process afar from the mainstream equilibrium idea. Markets are understood as discovery processes of entrepreneurs that are gaining a profit (Knight, 1921/2013). Dynamic views of competition

focus on satisfaction of others' needs in a discovery process that is targeted at searching for better ways to satisfy others' needs. This stands contrary to the neoclassical idea of competition as a state of affairs where many producers offer the same good and are powerless on prices in competition.

In a spontaneous order that results from human action, prices and markets are formed by the highly complex interplay of human actions. In Hayek, order and the knowledge retrieved in the spontaneous order allow for forming correct expectations about future market occurrences. The rules of the spontaneous order are followed consciously but also unconsciously. Sometimes people wish to not abide by the natural order and therefore should change based on a moral authority.

Overall, the knowledge paradigm appears to be better to cope with a world of truly divergent perceptions. The knowledge paradigm explains how market innovations are built on discovery and learning in comparison to the scarcity perception of the world that limits economic output to figure out how to cope with scarce resource allocation task. The knowledge paradigm also addresses the subjective elements of any rational choice. While the closed world model is about *what* one does not know, the open world model is acknowledging *that* one does not know.

Data centers in the age of digitalization

The big data revolution evolved in the digital age by server-based media storage. Computational power and storage ability have progressed ever since the advent of the internet. Today, data is stored in unprecedented dimensions thanks to modern hardware and technology advancements. Solid-state drives (SSDs), object storage and software-defined storage are some of the inventions in the new millennium that help cope with a rising digitalization and data transfer trend. The Internet of Things with smart devices being connected to every task throughout the day, but also the sharing economy explosion, have driven enormous trends in data storage and computational analysis. The COVID-19 pandemic has exacerbated digital realities and with it the amount of transferred data grew at an unprecedented rate.

Big data centers record our digital footprints all along our digital consumption, online behavior and interaction. These digital archives embrace the estimated 5.3 billion globally connected individuals –accounting for 65.7 percent of the world population – who use the internet and interaction digitally (Durmus 2024).

Big data insights promise better life quality, corporate service improvement but also efficiency gains and preparedness for trends. The ability to make sense of complex data relations and find patterns in constantly-floating data streams is therefore extraordinarily gratified in today's economy. The intelligence and bottom-up on the ground information sourcing capacities are unprecedented features of the digital age.

The global IT infrastructure, however, is these days challenged by environmental concerns arising from big data storage facilities. Big datacenters have 24/7 running servers that need space and to be cooled constantly. The temperature regulation for generators and backup batteries imply enormous energy consumption resulting in data centers having a growing impact on the environment. Solutions to lower energy consumption have been to store data in cooler parts of the world or near wind farms. Rebound effects that lead to higher technology consumption due to an illusion of unlimited data storage capacities, however, demand for further and more innovative strategies to cope with the surmounting data capacities that are stored.

Currently the energy consumption of big data centers is estimated to range at around 1% of the entire world energy consumption and 2-4% of the global carbon emission, which could rise up to 5.5% by 2025 (Hongyu Zhu et al. 2023; Serverwala Cloud Data Centers 2024). Technology improvements to boost efficiency of data storage facilities as well as carbon footprint reduction but also renewable energy transitions are contemporary proposals to help

cope with a growing problem (Microsoft 2024). Transportation improvements and supply chain management alongside service design optimization are meant to reduce the impact of big data storage on the environment (Microsoft 2024).

Google emphasizes environmental stewardship in energy efficiency and renewable energy solutions but also circular economy principles (Google 2024). Green building standards are monitored but also advancements in cooling technologies and mindful water restoration projects contribute to the environmental sustainability pledge of Google. Amazon also focuses on renewable energy and innovative cooling techniques alongside improving power efficiency with AWS-designed storage chips (Amazon 2024). Meta, the company behind Facebook and Instagram, also uses renewable energy sources and stresses water stewardship. Energy efficiency innovations but also reducing e-waste via circular economy practices is aimed at reducing Meta's environmental impact (Meta 2024).

While it is obvious that the big data gatekeeper institutions are aware of the massive environmental impact and rising sustainability concern of the data storage facilities, the solutions all appear to be housed in technology advancements and energy source choices. Intellectual reflection on the need for selectivity and demand for scarcity to not conserve all information given, are – at least – not obviously thematized. This paper therefore advocates for attention to Hayek's (1945) knowledge paradigm to weed out what parameters should be used to dissect the wealth of information given and what priorities should be set to distinguish useful information to those pieces that should rather be forgotten.

Using the knowledge paradigm and scarcity calculus to focus on meaningful data insights

Big data management facilities have three main tasks: (1) Track incoming information, (2) Process this information and (3) Store the information. In all three components of data management, Hayek's (1945) knowledge paradigm in relation to scarcity offers invaluable cues how to handle the constant information inflow.

When it comes to tracking incoming information, a recentering of focus could occur by acknowledging that information is dispersed and in such a way spread out that human capacity is impossible to connect all dots. Computational power may aid in making sense of the massive influx of information but the making sense of data may be guided by human decision making fallibilities. First, data processing must become aware that human minds tend to favor the now over the past and future. Real-time tracking and analysis of data may thus be conducted at data storage facilities, which are easily available and form some gatekeeper function.

The data should then be processed for storage and weight given on the importance of information. Irrelevant information may not be stored in front-center locations but in the periphery or at data centers at less monitored locations where the storage may become more energy efficient. In the processing of information, the knowledge paradigm dispersed function of information should be considered. Information should be stored in neural network like webs of interconnections, which can highlight trends but also the different qualities of data. Like different shades, information should be ranked based on relevance. Relevancy and discounting elements should be integrated in the formation of information networks. Like mental neural networks, for example, the least frequently used data nods should be abandoned.

Lastly, in the storage of information, decentralized techniques should be pursued, in which decentralized resources are used for storage, while high-level analytics should be performed in data centers with analytic capabilities. Like in the beginning of the internet, browsers at academic institutions could be used to store information but also built-in storage devices could track decentralized and only relevant information be accessed. The

accessibility could be pegged to tax revenue in order to create the fiscal space to offset some of the losses implied by big data storage carbon footprint externalities.

Another decentralized information pooling attempt lies in the MIT project 'Solid,' which is a world wide web decentralizing project. Solid stands for Social Linked Data. Led by Professor Tim Berners-Lee, who is one of the inventors of the world wide web, the project offers linked-data applications that are completely decentralized and under individual users' control. This decentralized peer-to-peer networking web is aimed at granting transparency and privileged accounts on data processing. More focused on bestowing individuals their own personal data rights as private good, the decentralization aspect could help outsource some of the data storage and data processing tasks to decentral hubs. Technically, individual users are enabled to store their personal data in so-called "pods" (standing for personal online data stores) that can be hosted wherever the user desires it. Applications that are monitored by Solid can request data from users that have control over the permission to use their data. Having to request information could grant more transparency over data transfer. It would also allow individuals to have more control what kind of data they want to share. And would create a new branch of data brokerage, which would help gain a price for individual data sharing. However, not to mention potential data network information losses as the information pockets are oftentimes disconnected from each other in various pods. Novel inequalities could arise from the value of data and data sharing capabilities being distributed unequally throughout society. It is assumed that education and salary levels are positively correlated with valuable data production online. And there may be a societal divide between those who have better strategies and access to data brokerages online.

Discussion

In the ocean of information and wealth of knowledge to be derived from in the digital age, the economic conundrum of our times is more centered around making sense of information. Environmental sustainability concerns add the demand for being selective in what information should be conserved for posterity.

When considering the historic debate whether economics should be centered around scarcity or knowledge, Friedrich von Hayek's initial critique of the scarcity focus in neoclassical economics appears to have won. Contemporary data processing and storage capabilities dwarf scarcity concerns but elevate environmental footprint concerns that demand for selectivity. Selectivity appears to be the demand of our time to determine our legacy for posterity. Only if we manage to determine what information is valuable to stored for future generations, we will be able to leave the earth in a better condition than found.

Future research avenues may include a nomenclature of importance of data. A public mandate to lay open data analytic insights would help in determining what information gain should be enhanced in which ones should be monitored or even be curbed. New network technologies that function as decentralization processors should be innovated. How to discount the future value of information could be an additional feature to determine what information should be conserved. Going back to historic traces of information that led to breakthroughs with a time-lag will serve as additional anecdotal piece of information to determine what to conserve and what to abandon in our wealth of information digital age.

This paper served as first call to refocus back on Hayek's point to make sense of collectively shared knowledge for innovation. Better ways to determine what information is relevant to derive meaningful insights from are called for. Rather than teaching scarcity of the conundrum in economics to be solved, how to cope with abundance of information appears to be the more sustainable solution and key to success of the current generation. Only if we find ways to determine what knowledge is useful and set discounting parameters right on the future wealth of knowledge, we will progress and advance humanity within its scarce environment.

References

Amazon. 2024. *The cloud*. Retrieved at https://sustainability.aboutamazon.com/products-services/the-cloud.

Data Report. 2024. *Datacenter sustainability*. Retrieved at https://datareportal.com/global-digital-overview.

Durmus, Isil. 2024. "Data Centers and Its Environmental Impacts. GreenerByte." Retrieved at https://www.linkedin.com/pulse/data-centers-its-environmental-impacts-i%C5%9F%C4%B1l-durmu%C5%9F-q5xvf/.

Google. 2024. *Datacenters*. Retrieved at https://www.google.com/about/datacenters/.

Hayek, Friedrich von. 1945. "The use of knowledge in society." *The American Economic Review 35*(4):519-530.

Hongyu Zhu, Zhang, Dongdong, Goh, Hui Hwang, Wang, Shuyao, Ahmad, Tanveer, Mao, Daijiafan, Liu, Tianhao, Zhao, Haisen, Wu, Thomas. 2023. "Future data center energy-conservation and emission-reduction technologies in the context of smart and low-carbon city construction." *Sustainable Cities and Society*, *89*, 104322.

Kahneman, Daniel. 2011. *Thinking, Fast and slow*. New York, NY: Farrar, Straus & Giroux.

Kirzner, Israel. 1973. *Competition and entrepreneurship*. Chicago, IL: University of Chicago Press.

Kirzner, Israel. 1985. *Discovery and the capitalist process*. Chicago, IL: University of Chicago Press.

Knight, Frank. 1921/2013. *Risk, uncertainty and profit*. Wilmington, DE: Vernon.

Menger von Wolfensgrün, Carl. 1871/1950. *Grundsätze der Volkswirtschaftslehre* [*Principles of economics*]. Glencoe, IL: The Free Press.

Meta. 2024. Meta *Datacenters: Infrastructure for today and tomorrow*. Retrieved at https://datacenters.atmeta.com/.

Serverwala Cloud Data Centers. 2024. Are Cloud Computing Data Centers truly sustainable? Retrieved at https://www.linkedin.com/pulse/cloud-computing-data-centers-truly-sustainable-serverwala-xfojc/.

Simon, Herbert A. 1986. "The Failure of Armchair Economics." *Challenge*, Taylor & Francis Journals *29*(5): 18-25.