1 FINTECH FOUNDATIONS: CONVERGENCE, BLOCKCHAIN, BIG DATA, AND AI

David L. Shrier and Alex Pentland

1.1 INTRODUCTION: THE CONVERGENCE REVOLUTION

Technology supercycles are accelerating. The digital transformation in financial services has attracted more than $500 billion in investment since 2015.[1] Fueling a disruption of the financial services industry, the transformation is in turn propelled not by one technology trend but rather by several that are converging contemporaneously and often in concert with each other.

The pace of technology-driven change is undergoing an increase of an order of magnitude or greater. It took nearly fifty years for landline telephony to achieve 60 percent market adoption. Social media usage and tablet computing reached that same level of penetration within seven years.[2] As networked communications promotes faster and faster adoption of new technologies, incumbent organizations such as conventional banks and asset management firms find themselves under threat from technology-enabled competitors. During the COVID-19 pandemic, certain segments of financial services digitized at an even greater rate—for example, in 2020, use of digital payments in lieu of cash doubled that of spring 2019.[3]

In this chapter we discuss the impacts of the maturation of an array of technologies within a few years of each other, which is creating a seismic shift in a traditional industry.

1.2 DISRUPTIVE TECHNOLOGIES

A cluster of technologies have been achieving new levels of utility and market adoption in recent years, creating new capabilities that had not been possible with only one or two of the technologies used by themselves. Mobile communications, artificial intelligence, big data analytics, and distributed ledgers are among the most prominent that are driving structural change in the financial services industry, in turn creating dynamic and widespread shifts in how consumers and businesses are served globally.

Mobile communications Mobile communications technology, developed in the late 1800s by Nikola Tesla and Guglielmo Marconi,[4] has become an economic engine that has enabled the developing world to catch up with, and in some cases overtake, developed economies—particularly with respect to financial services. World Bank research shows—and the University of California, Berkeley, confirms—that for every 10 percent of increased digital penetration in a given economy, there is more than a 1 percent increase in GDP.[5] This digital connectivity has been accelerated by the widespread deployment of mobile communications systems because it is no longer necessary to lay down terrestrial cable through remote, inaccessible, or underdeveloped regions, "leapfrogging" a generation of technology. In Africa, for example, nearly two hundred thousand Kenyans were shown to escape extreme poverty in only eight years through use of mobile money provider M-Pesa.[6] Popular Chinese mobile platform WeChat has more than 1 billion users (including 79 percent of the Chinese market when including all of its apps), and more than 50 percent of these users engage in financial services on the platform.[7]

Artificial intelligence Artificial intelligence (AI) has been posited for millennia and developed on modern computers for decades. Only recently, however, has it achieved widespread

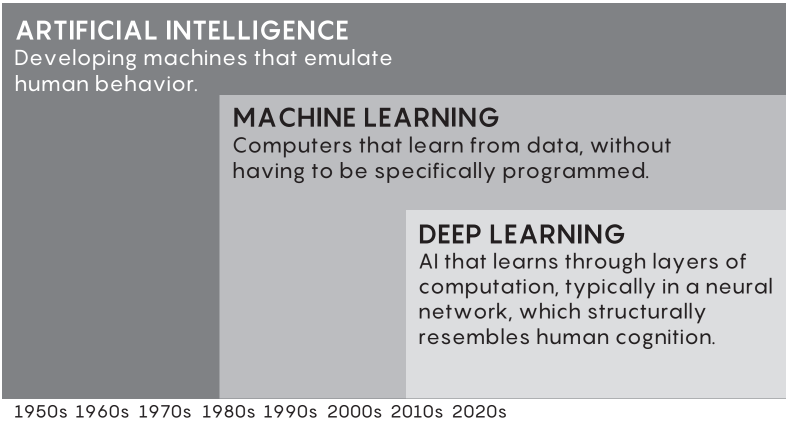

FIGURE 1.1
Different flavors of AI.

and deeply functional adoption. Figure 1.1, adapted from an *Oracle Big Data Blog* posting,[8] illustrates how different flavors of AI have emerged over the past forty years.

Note in the timeline that experiments with AI expert systems have been conducted on digital computers since the 1950s. One famous example is MIT scientist Joseph Weizenbaum's ELIZA digital therapist from the 1960s, which often fooled people into thinking they were interacting with a real human being.[9] More recent work around deep learning enables better speech-to-data translation, such as that found in Apple's Siri system.[10] A seminal event in the evolution of AI adoption was Google's 2015 release of TensorFlow, a software library for machine learning that made computing power and neural networks easily and widely available in an open source library.[11]

Big data analytics With the advent of better AI systems, we have seen the rise of big data analytics, which is now being used across multiple industries and has introduced profound changes in how we work, play, and live. Patterns can be extracted from all sorts of data flows, ranging from intraday trading activity to the physical movement of consumers in and out of retail stores (enabling better predictivity of system-wide sales). Big data analytics can even be used to detect whether company executives are lying or shading the truth on earnings calls. It can also be used to better manage cash liquidity systems, stabilize economies, or reconfigure a workforce.

Distributed ledgers In our 2016 book *Frontiers of Financial Technology*, we laid out the evolution of network communications, culminating in distributed ledger. Distributed ledger technology, or DLT (popularly known as blockchain), is itself a product of several technologies and methods that had been developed decades earlier, such as Byzantine consensus and distributed computing. As of this writing, nearly $300 billion in value has been created in a nominal market capitalization of the digital tokens that are enabled by DLT.[12]

When we bring together DLT, AI, big data analytics, and better mobile communications networks, we create a perfect storm of efficiency improvements and new capabilities—as well as jobs disruption that PricewaterhouseCoopers estimates could impact millions of financial services workers. Chapter 2 will delve further into blockchain and DLT.

1.3 THE CONVERGENCE REVOLUTION

The convergence revolution opens up the possibility of finally addressing the financial services inclusion gap, whereby 3.5 billion people around the world are underserved or unserved by the banking system. One of the interesting artifacts of this transformation is that it allows new entrants to offer the services. Whereas previously an incumbent bank may have worried about nonbank finance companies, today the threat could equally come from a telecom provider or a digital platform company such as Google, Amazon, Alibaba, or Tencent.

It's not only technology that's driving this change. Over decades of oligopoly, banks grew accustomed to an incumbent mentality and customer service suffered, leaving them vulnerable to the "neobanks."[13] Futurist Benjamin Palmer said, "The only time I talk to a bank is when I've done something wrong or they have."[14]

The new wave of challengers brings a completely different mind-set about customers. We can see this illustrated in the Net Promoter Score (NPS). The NPS is a normalized measure of customer satisfaction developed by Bain & Company and Satmetrix.[15] It measures the likelihood of a customer recommending your product or service to someone else. The scale runs from –100 (everyone hates you) to +100 (everyone loves you). Citi reportedly has an NPS of –41 (yes, that's negative 41).[16] The average for all financial services is about +18. Apple, which recently has been making inroads into consumer credit cards and payments in collaboration with Goldman Sachs, has an NPS of +47. Given these figures, conventional banks should be worried about *platform envelopment*, whereby digital platform companies, already integrated into people's lives and exhibiting warm feelings of trust and dependency, keep extending their services until they encroach on other offerings (such as, in this instance, those offered by incumbent banking institutions).[17]

But let's not count out fintech companies. The low-cost money transfer app Transferwise is reputed to have an NPS of +76.[18] Fee-bearing activities such as money transfer and credit cards are the heart of a bank's profitability—much more profitable than investment banking or other areas, consequently attracting the lion's share of venture capital investment.[19]

We thus see the convergence revolution creating competitors on both left and right flanks of the conventional banking industry, all while they are laying off tens of thousands of employees through the adoption of technologies such as AI.[20] Layoffs might be attributed to global market activity, to Brexit, to poor trading performance, or to some other factor, but they are made possible by digital convergence.

MIT Sloan School of Management professor Erik Brynjolfsson predicts a workforce realignment of 50 percent or more, on par with the Industrial Revolution of the 1800s, owing to digital disruptive technologies such as AI. That said, whereas AI may destroy a large number of jobs, a large number of jobs in new industries will also be created.[21] With the advent of robots comes the advent of robot repairmen. The job title "Interactive UX Designer" didn't exist thirty years ago, and thirty years from now there will undoubtedly be jobs like "Quantum Communications Specialist" and other titles we lack the foresight to envision today.

Convergence is creating financial access and the ability to bring the benefits of robust financial services to nearly half the world's population to enjoy. It is also causing large-scale disruption in the conventional financial services sector. In this book, we will explore an array of topics that explain how, why, and in which direction the industry might evolve.

One of the key technologies that is converging is distributed ledger, colloquially known as blockchain. In the next section we will undertake a high-level examination of this technology.

1.4 BLOCKCHAIN BASICS

Something exciting happened with a sausage.

The story of blockchain, which we discuss in this chapter in terms of both what it is and why it is interesting from a business and societal standpoint, has taken decades to play out. But first we will describe how blockchain is the computer systems equivalent of sausage, a humble bit of food eaten by those who can't afford steak (or Beyond Meat, if vegetarian).

Sausage holds a special place in cuisine. Somewhat grayish, lozenge shaped, made up of little bits of this and that, including what some may believe is fiberglass wall insulation, it becomes palatable with enough ketchup or horseradish or mustard or possibly all three, a dryish conflation of left-overs that begs for seasoning. Not unlike haggis—a Scottish national treasure made out of the parts of the sheep that no one wants to eat, which can be survived only when doused in Scotch whiskey (ideally while reciting the Robert Burns poem "Address to a Haggis")—sausage benefits from a food companion in order to be consumable.

Blockchain is like a technological sausage. It is a kind of database with special features that help it solve certain problems, and in and of itself it doesn't represent new technology. The individual components that go into making blockchain—how it accepts new data, how it secures that data, and how it makes that data visible and available—were around for years before Satoshi Nakomoto's 2008 white paper and the origins of the first large-scale instance, the bitcoin blockchain. Like the scraps of meat that go into making a sausage, these elements were assembled into blockchain, which assumed a new kind of importance by virtue of both composition and timing. And to stretch the metaphor, properly seasoned it makes for an exquisite dish. The financial form of blockchain, cryptocurrencies, had a market capitalization at this writing of about $250 billion; notably during the COVID-19 crisis, as the general equities markets suffered, cryptocurrencies proved resilient. From January 1 to May 31, 2020, the Dow Jones Industrial Average was down 12 percent and the Financial Times Stock Exchange 100 Index was down 20 percent, but traded tokens had a return of 31 percent.[22]

Blockchain has certain features, like immutability, a distributed nature, and a consensus mechanism, that combine to create digital trust that the system is telling you the truth about your data. You don't need to rely on the assurances of governments or of banking institutions, neither of which was particularly trusted after the 2008–2009 global financial crisis. Instead, you can trust in the technology.

1.5 WHAT BLOCKCHAIN IS

Blockchain is simply a database that has interesting features, making it suitable for certain applications that require trust and transparency. These could include financial systems, supply chain management, or even the access and sharing of health data. Also termed "distributed ledger," emphasizing the record-keeping aspect of what distinguishes it from other types of databases, blockchain offers a high degree of assurance that your information hasn't been corrupted by hackers or other bad actors.

As described in more detail in *Basic Blockchain*,[23] parts of blockchain were born out of the distributed computing movement. The concept was that when your computer was at rest, the idle processor cycles would be used for more than running the once-ubiquitous screen saver. Instead, with projects like SETI@home, thousands upon thousands of computers had their downtime applied to the fascinating problem of searching for extraterrestrial intelligence by analyzing radio telescope signals. It was citizen science, and it was empowering the crowd to create new intelligence; but it was also a clever programming hack to take a difficult problem, for which an individual science lab lacked the resources to solve, and make it more manageable. We see many instances of this distributed computing model today aside from blockchain; it could be argued that the edge computing industry, expected to be worth $43 billion by 2027,[24] was born out of a humble hack to find little gray men.

So the first essential element of blockchain is its distributed nature. Thousands of servers all over the world are involved in calculating the math that goes into making blockchain cryptographically secure. For the sake of this discussion, we're going to ignore other variants of distributed ledgers like R3 Corda. (R3 Corda is technically not a pure blockchain but a distributed ledger technology. To understand the distinction, take a look at https://www.r3.com/blockchain-101/.)

Space seems to feature prominently in the history of blockchain. Another critical component of how blockchain works is how it decides to accept new information into the database. The concept of Byzantine consensus originated out of SRI International to solve the issue of unreliable computers in space.[25] As Douglas Adams wrote, not only is space "big," it also contains lots of radiation from sources like the sun, which doesn't interact well with computer systems. A fault-tolerant model is required, where even if one computer has its data corrupted by a stray bit of radiation, you can still land your spaceship safely.

As NASA was designing and building America's fleet of space shuttles, it needed to ensure that the computers in the shuttles would be reliable. An elegant solution was generated out of a thought experiment about how to manage trust among a group of rivalrous generals in the ancient Byzantine Empire (called, appropriately, the Byzantine Generals Problem). The solution was the creation of a particular kind of voting mechanism—even if you couldn't necessarily trust one particular general, you could trust the group of them voting together. In NASA's case, four identical computers were built into the space shuttle, and three of them had to agree in order to verify a particular assertion. In the case of the original blockchain, 51 percent of the servers in the distributed network needed to agree that something was true in order to accept it as a new piece of information incorporated into the database.

This construct, a distributed network that has a consensus mechanism to accept new data, means that blockchain is *cyber resilient*. Unlike a centralized system, where a hacker who penetrates the computer can modify the data inside of it with impunity, blockchain requires compromise of many different systems in order to inject new data. It is designed to withstand distributed attacks, which consist of large-scale mobilization of computers that seek to overwhelm a system from many points at once. The need for hack-resistant data systems has become exacerbated in recent years, with $6 trillion in cyber breach–related losses estimated for 2021.[26]

As an aside, no one knows who Satoshi Nakamoto, the progenitor of blockchain, actually is. Much like the question "Who wrote Shakespeare's plays?," it is of academic interest, but today it doesn't really matter. Shakespeare, too, stole bits of legend and lore and structure from others in crafting his masterpieces. We have our *Hamlet*, we have our bitcoin, so functionally it's less important which hand authored the creation.

Another critical component of blockchain's resilience is the use of Merkle trees. The basic concept is that as a blockchain grows, each new link in the chain is inexorably (and mathematically) tied to every link behind it. If, for example, your blockchain represents a financial ledger, you can't go change a historical entry (perhaps in an effort to commit fraud) without altering every other entry in the chain.[27]

1.6 WHAT BLOCKCHAIN IS GOOD FOR

The hot topic in the blockchain world is identifying natural use cases that scale. Areas such as digital payments appear to be a natural use case, and others are exploring applications ranging from aerospace to health care.

What limits scale? Blockchain is a complicated database. It typically uses numerous computer cycles in order to cryptographically secure its data payload. While many things could be put onto a blockchain, this does not mean they should be. In many instances, other forms of databases, like a relational database (such as Oracle) or a large-scale unstructured database (like Hadoop), are better suited to managing the data, although some newer iterations of blockchain technology are specifically designed to address historic limitations.

New technology isn't an innovation unless it is applied to a problem at scale. Many hundreds of different use cases for blockchain have been posited. However, not all of them represent a large enough market opportunity or impact to justify the investment in converting technology and process infrastructure to blockchain, and not all of them require the trusted consensus that blockchain delivers. Over the past five years there have been numerous proofs of concept to experiment with blockchain applications. The question now being posed is, aside from the bitcoin blockchain, what are useful applications at scale?

Many believe that core financial systems will be rebuilt using blockchain. Major trading markets such as Nasdaq have active research and prototyping efforts underway in blockchain, and the stock exchange of Suriname has announced its intention to be the first in the world to trade listed equities in blockchain and to serve the broader Latin American and Caribbean region (as opposed to the well-established array of "blockchain exchanges" that allow the trading of cryptocurrencies, Suriname is actually going to trade the stocks of its major companies).[28] International payments and money transfer, traditionally the domain of companies like Western Union or SWIFT, are getting upended by companies like Bitpesa and Ripple.

Notably, government is getting into the act. In the wake of the "Libra backlash," where an overwhelmingly negative reaction arose to Facebook's efforts to create a digital currency outside the reach of government, and the moves by China to corner the digital currency market with the creation of the digital RMB coin, other governments are looking more seriously at national and multinational central bank digital currencies.[29]

Other domains outside of financial services show promise: for example, supply chain management, where a network of interlinked market participants need to trust in the veracity of the data and have perverse incentives. Two and a half percent of aircraft replacement parts are counterfeit, risking flight safety, according to former MIT Future Commerce student Eleanor Mitch, who has created 14bis to solve this problem with distributed ledger.[30] Health care and medical research could be enhanced by management of “granular patient consent” and better secure data management, concepts being developed by companies such as BurstIQ.

The period from 2021 to 2025 should see blockchain applications achieving scale and adoption in mainstream corporate and government environments.

1.7 WHAT BLOCKCHAIN COULD ENABLE

If we step beyond the immediate potential applications of blockchain, it becomes interesting to explore what might be possible with the evolution of blockchain.

Taxes could become automated, relieving the time and expense burdens on people and companies, reducing cost and friction to enable governments to function.

Intellectual property rights for music or video domains could be automatically managed, with payments flowing seamlessly alongside consumption, perhaps enabling better protection of the rights of individual artists.

The distributed, immutable, and transparent characteristics of blockchain could make it possible to better manage elections, with every vote counted (and counted only once). Citizens could have greater confidence in the security and reliability of the mechanisms of governance. We could even make the plebiscite more feasible at scale, allowing direct democracy in a way not seen since the days of ancient Athens (only this time, women and poor people can vote).

Blockchain is still an immature technology. Its final evolution and its greatest use cases are yet to emerge but should become more apparent within the next few years.

At its heart, blockchain is a database. But there is an even bigger revolution around financial services than blockchain, and it is happening in the world of big data and AI. The next section delves into this area.

1.8 BIG DATA AND AI FUNDAMENTALS

What is "big data"? What distinguishes it from other data?

When we speak of big data, we typically think of it as having certain characteristics that distinguish it from other kinds of data that don't need as sophisticated a set of methods and tools in order to process and understand it. The characteristics of big data are commonly called the "three V's" of volume, velocity, and variety:[31]

- **Volume:** Big data is, by definition, "big." The exact definition of "big" is a moving target, as processing power and storage continue to increase in capacity. A few years ago, "big" was terabytes of data. Then, petabytes. Now, we think about exabytes of data, or 10^{18} bytes. One exabyte is equivalent to about 320 billion copies of the King James Bible.[32]
- **Velocity:** Big data is data that is changing rapidly. The locations of every inhabitant of New York City or Shanghai are sets of data that change from minute to minute, as well as being high volume. Individual purchase transactions on credit cards for all of Brazil's day to day are a rapidly moving data set.
- **Variety:** High diversity is another dimension of big data. Instead of simply having repeating text or numbers of uniform size that fit neatly into rows and columns, you might have a mixture of different kinds of unstructured data, such as video or free-form text.

Data provides the feedstock for a particular kind of AI known as *machine learning*. When we talk about AI, there are several different flavors. *Artificial general intelligence* (AGI) is a machine that can think like a human and learn any new idea or skill in an unsupervised manner. We are many years away, at least, from widely applicable AGI—some experts think many decades.[33] Other kinds of AI include *expert systems*, where human knowledge is boiled down to a set of discrete rules (*if A, then B*), and *machine learning*, where probabilistic models iteratively improve themselves. Systems like Google search and voice-enabled software that enables you to call your bank and simply ask it a question (like “What’s my account balance?”) are built on machine learning algorithms. Machine learning is being applied in numerous areas within institutional and retail financial services, on everything from market forecasting to fraud prevention.

Part of why Google works so well for searching is that its AI has been trained on more than 1.2 trillion searches per year.[34] This massive amount of unstructured data has provided a critical advantage for building a really smart AI, at least for the particular application of internet search (all of those inputs have helped Google with other applications, of course). Financial institutions as well have a significant advantage in building new machine learning applications because they are able to leverage large-scale data sets that are under their control or influence.

1.9 BIG DATA AND FINANCIAL SERVICES

Many exabytes of data (and soon, many zettabytes, 10^{21} bytes) are being generated from people using credit cards, debit cards, and mobile payment platforms like Apple Pay, Amazon Pay, Venmo, and Alipay. These data streams are tagged with rich information, including not only the location of the purchase transaction but also the category of goods and services and the time of day, and, once analyzed in aggregate, can indicate trends of behavior.

The concept of using computers in the banking system has been around for a long time, but it was really the advent of ATMs in the 1960s and 1970s that led to consumer behavior data from electronic devices being used in banking business strategy. For example, after deploying its ATMs, Citi discovered that people withdrew money after work and on weekends, which was antithetical to the conventional banking hours of 9:00 a.m. to 3:00 p.m. Data from ATM network usage led Citi to add branch hours and increase its business through the application of analytics. This, in turn, enabled Citi to engage with its customers around other kinds of financial products, and the business enjoyed rapid growth thanks to extended hours, a practice that rapidly spread across the conventional banking industry.

Mobile communications networks became more prevalent in the 1980s and 1990s. With the advent of high-speed data in the next twenty years, we will see an explosion of information about people and their movements around towns and cities, data that informs financial services. Africa is expected to have 600 million smartphones by 2025,[35] versus 1.2 billion mobile phone lines overall; even the least developed areas on the planet are now getting some form of connectivity. With that connectivity comes the ability to understand human behavior and improve the delivery of financial services. We explore “leapfrog” possibilities from technologies such as mobile financial services in more depth in chapter 4, and around Africa specifically in chapter 9.

Institutional financial services have also enjoyed meaningful innovation through the application of AI and big data. High-frequency trading would not be possible without sophisticated AI systems. Quantitatively driven investment firms like Two Sigma, Renaissance Technologies, and Bridgewater have generated outsize returns for their investors by employing armies of mathematicians and computer programmers to exploit big data analytics for alpha generation.

1.10 DATA QUALITY

When we’re trying to build quantitative models that interpret, analyze, or forecast market price movements or an individual consumer's credit behaviors, we are relying on big data streams to power those models. When models are initially created, the baseline assumption is that the data can be relied on, but the reality is often much different.

One big data set we were working with, which involved several terabytes of data, had anonymous location information for mobile phone users. Fair enough as far as it went. However, we found some outliers such as data that suggested an individual was inside the city one minute, was half a mile out to sea fifteen minutes later, then was suddenly back in the city. Such movement is not likely in a real-world setting, at least not until we have commercial teleportation technology. Savvier data science teams will build data assessment engines to automate the evaluation of data quality as it is being ingested and interpolation and extrapolation systems to bridge gaps that may appear in data to facilitate model development.

As AI systems are built and trained using big data, it is incumbent upon technology and business professionals to put into place processes to audit and improve data quality so that the systems in turn can produce better answers.

1.11 SOCIAL PHYSICS AND UNLOCKING POTENTIAL

With the advent of widespread big data / analytics and improved AI systems, we have seen the emergence of new computational social sciences that offer many benefits to financial services. *Social physics* is a dominant example of this new approach to understanding, predicting, and even changing human behavior at scale. The key idea is to use universal regularities in human behavior, such as the "long tails" in distributions of behavioral variables, to explain the data, rather than trying to model the data using general-purpose models like linear regression or neural networks.

For example, conventional lending relies on linear regression models offered through credit bureaus, requiring three years of data and suffering from the "credit trap" that in order to get credit you must have credit, thus excluding 3.5 billion people around the world from meaningful participation in financial services. With social physics, we can in theory improve credit modeling and predictions of credit behaviors by 30–50 percent versus credit bureaus, and we may be able to do this with just one month of data rather than thirty-six (once the model is trained).[36] This holds exciting potential for addressing financial inclusion in the developing world, where most people have thin credit files or no credit file at all.

New methods are emerging to ensure data security while still making data useful. The conventional method of data analytics is to bring all of the data together in a single repository, a *data lake*, and then perform analytics on it. While perhaps convenient for the data analyst, and while data insights can be more profound when data is combined and shared, copying and centralizing data is also highly cyber insecure. The OPAL Project (www.OpalProject.org) says that instead of bringing the data to the code, you bring the code to the data. OPAL is built around the premise that data is left atomized, it is placed in various secure repositories, and carefully vetted algorithms are sent to each data repository. These algorithms can extract insights. So, for example, you could have telecom data, banking data, and health data all sitting in different repositories. *Without* having to copy the data over (which may be illegal in many domiciles, and in all cases creates security risks), a query can instead be sent to each of the databases and the software would merge these insights into coherent information.

The ability to not only predict but actually influence the large-scale behavior of people through the application of big data / analytics creates a set of ethical and moral challenges. The introduction of these kinds of sensing platforms brings with it meaningful ethical and moral risks. The most extreme example is the social credit score in China, where behavioral data from mobile phones and other communications networks is used to determine people's eligibility for loans, housing, travel, and other aspects of participation in society. China has elected to trade individual liberty for social stability. To a lesser degree, these kinds of trade-offs occur across a number of Western democracies as well. For example, the United Kingdom, and particularly London, has opted for a higher degree of surveillance in exchange for greater security, installing a network of closed-circuit TV cameras on streets and in vehicles.[37] Germany, on the other hand, has gone for greater personal liberty and attenuated the ability for the security services to quickly track down bad actors. Each of these societies has made a set of value judgments about data ethics.

1.12 ETHICS OF BIG DATA AND AI

As with many disruptive innovations, there is a gap between what we *could* do with a particular technology and what we *should* do. Splitting the atom led to the ability to annihilate millions of people in an instant, as well as the ability to power an entire country with clean energy and to cure cancer. Commercially viable steam engines enabled military organizations to rapidly deploy troops to conquer territory and to support those troops with appropriate logistics, and they enabled the creation of transcontinental shipping networks to invigorate trade and unify culture. Big data and AI systems likewise have fearsome possibilities that have already been used to the detriment of society, and they have tremendous utopian potential.

The governor in all these decisions is the human ethical framework. When AI programmers are creating a system, what kind of system do they choose to create and what limitations do they place on it? When business or government or individual users apply that AI technology, what do they choose to apply it to?

Professors Luciano Floridi and Joshua Cowls have proposed a rigorous ethical framework for designing and deploying AI.[38]

In it, they map five key concepts derived from meta-analysis of the thinking around this subject across forty-two countries:

1. benevolence (AI should do something good for us)
2. nonmalevolence (AI shouldn't do harm in the process of doing something good for us; having a robust data privacy framework is an example of nonmalevolence)
3. autonomy (humans should be able to make key decisions and regulate how autonomous the AI is)
4. justice and fairness
5. explicability (AI shouldn't be black box; the AI should be able to tell us how it makes decisions)

This work, and other research like it, helps create parameters around which AI systems creators can inculcate *ethics by design* rather than attempt to retroactively assess whether a given AI is falling within certain ethical parameters.

1.13 REGULATION

Concomitant with the discussion of data and AI ethics, new regulations begin to emerge around personal data, such as the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act in the United States. The European Union is also contemplating specific regulations around AI, and its ethical applications, in addition to the protections enshrined in the GDPR. Regulators are engaging in active consultative processes with the private sector, academia, and public interest advocates in an effort to shape meaningful regulations around the applications of big data and AI to financial services.

The most enlightened companies have proactive regulatory engagement efforts. Rather than waiting for an enforcement letter from a regulator as their first contact with the government, these organizations have teams deployed to share plans with regulators, understanding and addressing concerns proactively and helping to build capacity among government agencies on the implications of new technologies to foster a more informed regulator.

1.14 UNSOLVED PROBLEMS

While big data analytics has performed a number of miracles in the area of financial services, a number of thorny issues remain:

- Large amounts of data are not necessarily good-quality data. For example, credit bureaus are notoriously bad at managing data, securing the data, and providing meaningful insights from it. Correcting errors in an individual personal credit file can be laborious, highlighting poor governance of the data.
- Big data techniques can improve on bureau credit scores by 50 percent or more, but incumbent lobbying efforts and regulatory capture mean that the existing players are difficult to dislodge.
- The introduction of big data techniques for credit raises new questions. For example, Goldman and Apple endured a public relations mess with the launch of Apple Card, as some accused its data evaluation techniques of deliberately discriminating against women. The opacity, the lack of explainability of new algorithmic credit models, is another challenge for the industry.
- It can cost a bank anywhere from $13 to $130 to perform a "Know Your Customer" check on a new client. This legally mandated background check is an effort by the bank to avoid doing business with terrorists or other proscribed individuals. This accelerating cost trend has meant that a number of financial institutions have retrenched from inclusion efforts and have exited certain markets, which is known as *derisking*. The World Bank has identified derisking as harmful to economic progress in developing countries.[39]

- Identity lies at the heart of another unsolved data analytics challenge, that of *ultimate beneficial owner* (UBO). When a business account is opened, the financial organization needs to ascertain not only who has direct signatures on the account but also the upstream individuals who may have financial interest in the account. In a global platform, particularly in the high-growth emerging markets areas, this is a difficult task even for the most well-resourced financial institutions.

The array of challenges presented in this chapter represent only a small part of the areas of potential innovation and entrepreneurial action that can help address the needs of the financial services industry and of consumers and businesses more broadly.

Now that we have established the foundations of financial technology, we can begin to explore use cases, applications, policy, and societal implications.

NOTES

1. I. Pollari and A. Ruddenklau, *Pulse of Fintech H2 2019* (KPMG, February 2020), https://assets.kpmg/content/dam/kpmg/xx/pdf/2020/02/pulse-of-fintech-h2-2019.pdf.

2. H. Ritchie and M. Roser, "Technology Adoption," Our World in Data, 2017, https://ourworldindata.org/technology-adoption.

3. D. Thomas and N. Megaw, "Coronavirus Accelerates Shift Away from Cash," *Financial Times*, May 28, 2020.

4. T. Long, "Aug. 23, 1899: First Ship-to-Shore Signal to a U.S. Station," *Wired*, August 23, 2011, https://www.wired.com/2011/08/0823first-us-ship-to-shore-radio-signal/.

5. C. Qiang, C. Rossotto, and K. Kimura, "Economic Impacts of Broadband," in *Information and Communications for Development* (Washington, DC: World Bank, 2009), https://www.semanticscholar.org/paper

/Economic-Impacts-of-Broadband-Qiang-Rossotto/40e5bf124f1005 6e4b02d7149bf1a0bd1a221b37; C. Scott, "Does Broadband Internet Access Actually Spur Economic Growth?," December 7, 2012, https://colin-scott.github.io/personal_website/classes/ictd.pdf.

6. T. Suri and W. Jack, "The Long-Run Poverty and Gender Impacts of Mobile Money," *Science* 354, no. 6317 (2016): 1288–1292.

7. M. Iqbal, "WeChat Revenue and Usage Statistics (2019)," Business of Apps, updated March 8, 2021, https://www.businessofapps.com/data/wechat-statistics/.

8. P. Jeffcock, "What's the Difference between AI, Machine Learning, and Deep Learning?," *Oracle Big Data Blog*, July 11, 2018, https://blogs.oracle.com/bigdata/difference-ai-machine-learning-deep-learning.

9. J. Weizenbaum, "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine," *Communications of the ACM* 9, no. 1 (1966): 36–45.

10. Siri Team, "Deep Learning for Siri's Voice: On-Device Deep Mixture Density Networks for Hybrid Unit Selection Synthesis," *Machine Learning Research*, August 2017, https://machinelearning.apple.com/research/siri-voices.

11. C. Metz, "Google Just Open Sourced TensorFlow, Its Artificial Intelligence Engine," *Wired*, November 9, 2015, https://www.wired.com/2015/11/google-open-sources-its-artificial-intelligence-engine/.

12. "Top 100 Cryptocurrencies by Market Capitalization," CoinMarketCap, accessed February 2020, https://coinmarketcap.com/.

13. I. Vidili, "How Much Time Is Left for Banks? The Lost Battle of Incumbent Barclays vs Agile 'Neobanks,'" *Medium*, June 4, 2019, https://medium.com/@thesmartercrew/how-much-time-is-left-for-banks-the-lost-battle-of-incumbent-barclays-vs-agile-neobanks-430627a2602f.

14. Personal conversation with the author, August 2016.

15. "The History of the Net Promoter Score," Bain & Company, accessed February 2020, https://www.netpromotersystem.com/about/history-of-net-promoter/.

16. "Citibank Net Promoter Score 2021 Benchmarks," Customer Guru, accessed May 2021, https://customer.guru/net-promoter-score/citibank.

17. T. R. Eisenmann, G. Parker, and M. Van Alstyne, "Platform Envelopment," *Strategic Management Journal* 32, no. 12 (December 2011): 1270–1285.

18. M. Gasparello, "What's behind Transferwise Massive +100% YoY Growth?," Strategico.io, 2019, https://strategico.io/transferwise-digital-strategy/.

19. R. Ghose, S. Dave, A. Shirvaikar, K. Horowitz, Y. Tian, J. Levin, and S. Ho, "Digital Disruption: How FinTech Is Forcing Banking to a Tipping Point," *Citi GPS: Global Perspectives & Solutions,* March 2016.

20. A. Cohen, "HSBC Layoffs: 3 Reasons Why the Banking Giant Is Imploded," FastCompany.com, February 18, 2020, https://www.fastcompany.com/90465194/what-went-wrong-at-hsbc; S. Butcher, "Big Job Cuts Coming at UBS after Nasty Quarter," eFinancialCareers.com, October 22, 2019, https://www.efinancialcareers.com/news/2019/10/ubs-q3-results-2019.

21. "Jobs and the Fourth Industrial Revolution," World Economic Forum, 2016, https://www.weforum.org/about/jobs-and-the-fourth-industrial-revolution.

22. Bloomberg.com; Coinmarketcap.com.

23. D. L. Shrier, *Basic Blockchain: What It Is and How It Will Transform the Way We Work and Live* (London: Little, Brown, 2020).

24. Grand View Research, *Edge Computing Market Size, Share & Trends Analysis Report by Component (Hardware, Software, Services, Edge-Managed Platforms), by Industry Vertical (Healthcare, Agriculture), by Region, and Segment Forecasts, 2020–2027*, 2020, https://www.grandviewresearch.com/industry-analysis/edge-computing-market.

25. L. Lamport, S. Shostak, and M. Pease, "The Byzantine Generals Problem," *ACM Transactions on Programming Languages and Systems* 4, no. 3 (July 1982): 382–401.

26. CISOMAG Staff, "Cybercrime Will Cost the World US$6 Trillion by the End of the Year: Study," CISOMAG, March 23, 2020,

https://cisomag.eccouncil.org/cybercrime-will-cost-the-world-us6-trillion-by-the-end-of-the-year-study/#:~:text=In%20fact%2C%20the%20damages%20caused,US%243%20trillion%20in%202015.

27. S. Ray, "Merkle Trees," Hackernoon.com, December 14, 2017, https://hackernoon.com/merkle-trees-181cb4bc30b4.

28. N. Ramautarsing, "OuroX Signs LOI with the Suriname Stock Exchange and Takes First Steps to Develop an Innovative Securities Exchange for Latin America and the Caribbean," OuroX, May 2, 2019, https://medium.com/ourox/ourox-signs-loi-with-the-suriname-stock-exchange-and-takes-first-steps-to-develop-an-innovative-46115a5d3015.

29. D. Shrier, "The Future of Money Isn't Libra or Chinacoin, It's Federated," Forkast.News, September 24, 2019, https://forkast.news/opinion-the-future-of-money-isnt-libra-or-chinacoin-its-federated/.

30. M. Wood, "Award-Winning Startup Delivers Blockchain Traceability for US Airforce," Ledger Insights, June 11, 2019, https://www.ledgerinsights.com/blockchain-us-airforce-traceability-aircraft-parts-14bis/.

31. D. Gewitz, "Volume, Velocity, and Variety: Understanding the Three V's of Big Data," ZDNet.com, March 21, 2018, https://www.zdnet.com/article/volume-velocity-and-variety-understanding-the-three-vs-of-big-data/.

32. Starry Staff, "How Big Is the Internet?," Starry, July 29, 2019, https://starry.com/blog/inside-the-internet/how-big-is-the-internet.

33. N. Joshi, "How Far Are We from Achieving Artificial General Intelligence?," *Forbes*, June 10, 2019, https://www.forbes.com/sites/cognitiveworld/2019/06/10/how-far-are-we-from-achieving-artificial-general-intelligence/?sh=7e205096dc4d.

34. InternetLiveStats.com, accessed March 31, 2020.

35. "2019: The Mobile Economy Sub-Saharan Africa," GSMA, accessed May 28, 2020, https://www.gsma.com/mobileeconomy/sub-saharan-africa/.

36. D. Shrier and V. Raeburn, "Giving Credit, Taking Credit: The Emergence of Predictive Credit Analytics" (unpublished white paper, November 2018).

37. P. Bischoff, "Data Privacy Laws & Government Surveillance by Country: Which Countries Best Protect Their Citizens?," Comparitech, October 15, 2019, https://www.comparitech.com/blog/vpn-privacy/surveillance-states/.

38. L. Floridi and J. Cowls, "A Unified Framework of Five Principles for AI in Society," *Harvard Data Science Review*, July 1, 2019, https://hdsr.mitpress.mit.edu/pub/l0jsh9d1/release/7.

39. World Bank, "De-risking in the Financial Sector," October 7, 2016, https://www.worldbank.org/en/topic/financialsector/brief/de-risking-in-the-financial-sector.

This is a section of doi:10.7551/mitpress/13673.001.0001

Global Fintech

Financial Innovation in the Connected World

Edited by: David L. Shrier, Alex Pentland

Citation:

Global Fintech: Financial Innovation in the Connected World
Edited by: David L. Shrier, Alex Pentland
DOI: 10.7551/mitpress/13673.001.0001
ISBN (electronic): 9780262369534
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif by Westchester Publishing Services, Danbury, CT.

Library of Congress Cataloging-in-Publication Data

Names: Shrier, David L., editor. | Pentland, Alex, 1952– editor.
Title: Global fintech : financial innovation in the connected world / edited by David L. Shrier and Alex Pentland.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021030642 | ISBN 9780262543668 (paperback)
Subjects: LCSH: Financial services industry—Technological innovations. | Finance—Data processing.
Classification: LCC HG173 .G636 2022 | DDC 332.1—dc23
LC record available at https://lccn.loc.gov/2021030642