CHAPTER 7: ARTIFICIAL INTELLIGENCE

7.1. Sector overview

Today's economy is a data economy as organizations use data and analytics to drive productivity and innovation. But this is transitioning into the algorithmic economy, in which many more organizations invest in artificial intelligence (AI) to automate processes, develop new products and services, improve quality, and increase efficiency (New, 2018). AI represents a cross-cutting, horizontal issue that is relevant to all firms and sectors engaged in trade. Using data, AI has the potential to impact virtually every sector of the economy given its ability to make and test assumptions (without human intervention), allowing it to learn autonomously. AI's impact on economic productivity holds the potential to be much broader, as various aspects of it can be understood as being "general purpose technologies" (such as microprocessors) that have historically been influential drivers of long-term technological progress as they affect most functions in an economy (Cockburn et al, 2017). By extension, AI-based applications can benefit both trade in goods and services, for example by optimizing route planning and enabling autonomous driving, reducing logistics costs through cargo and shipment tracking, and using smart robots to optimize storage and inventory (WTO, 2018).

AI's emerging role in international trade is based on the transformative impact of the Internet and other digital technologies. The rapid growth in the volume and diversity of data produced by digital platforms, wireless sensors, billions of mobile phones, and other sources, when combined with low-cost, widely accessible, and increasingly sophisticated cloud-based data storage services provides a platform for firms (of all sizes) from around the world to develop and deliver or use AI-based services. AI will have its biggest impacts on more routinized information-based functions, which tend to be services, (e.g., making loans, processing accounts, or analyzing medical tests) (ITIF, 2018). It is in relation to services that there are both significant opportunities—the WTO (2018) estimates that the share of services trade could grow from 21 to 25 percent of total global trade by 2030—but also peril, in that services trade liberalization (in terms of new and meaningful services market access and addressing the non-tariff issues that affect services) has long taken a back seat to traditional trade goals of reducing tariffs on goods.

Likewise, the McKinsey Global Institute (2017) estimates that the potential for data analytics in digital trade is significant, in part, as many firms are (still) capturing only a fraction of the potential value of data. There is a large and significant gap in the degree of digitalization between certain sectors and within sectors, where a few leading firms often lead a large group of laggards in terms of developing and using advanced digital capabilities (McKinsey Global Institute, 2017). This gap is a major factor shaping competition in an economy as leading firms are more profitable and successful. Legacy firms face the challenge of adapting to new digital technologies, like AI, which opens up the opportunity for firms that have mastered these technologies to provide their services to help them close the gap. Likewise, it also presents opportunities for those firms "born digital" that have AI/ML at the heart of their business model to either disrupt the incumbent players or to provide their services to help them catch up.

As this chapter outlines, firms using AI as part of their business model (or even as their entire business model) depend upon the ability to collect, use, transfer, and share a large volume and diversity of data to train and deploy their services, and these firms need to maximize the value of their investments in AI expertise and systems by deploying them as widely as possible across sectors and borders. Absent regulatory, market, trade, and other artificial barriers, these firms should be able to leverage modern ICTs to do this remotely as a form of services trade. However, this is not the case in many scenarios, which raises trade policy concerns around the rules and regulations that affect data, intellectual property, and market access (for key service sectors and for the remote delivery of services).

7.2. Profile of firms interviewed

Mindbridge Ai

Mindbridge Ai is a small, but rapidly growing, data analytics firm based in Ottawa, Canada[119]. Established in 2015, Mindbridge Ai has around 70 staff (its staff doubled in size in 2018). More than 230 customers in seven economies use Mindbridge Ai's AI auditor tool. Mindbridge Ai's main target is the external auditor services sector. In its short history, Mindbridge Ai has developed an impressive track record. In 2018, the Canadian Advanced Technology Alliance gave Mindbridge Ai its outstanding product achievement award, Accounting Today said its AI auditor was the top new product of 2018, and Mindbridge Ai's CEO Eli Fathi was named AI Leader of the Year at the Canadian FinTech & AI Awards. Mindbridge Ai also won the Central Banking's FinTech and RegTech Global Award for Best Machine Learning Solution for Regulatory Compliance.

Mindbridge Ai uses a hybrid of techniques—from decision-based rules and statistical methods, to ML and AI—to perform real-time data analytics, pattern recognition, and anomaly detection in order to help various organizations investigate or audit past activity, detect active inadmissible behavior (e.g., fraud), and prevent potential transgressions. Mindbridge Ai has two core products/services: its cloud-based AI Auditor platform and AI Advisory, which provides custom data analytics services for clients.

Mindbridge Ai's main product, AI Auditor, is used by leading certified practicing accounting firms and governments worldwide to detect anomalies in financial data. Through the automated ingestion and analysis of financial datasets, AI Auditor detects anomalies. AI Auditor's results are presented through an intuitive interface that augments the capability of auditors and investigative professionals by allowing them to focus on anomalous transactions. This significantly reduces the risks associated with manually analyzing samples of the transactions, while also delivering deep insights on the financial datasets.

Mindbridge Ai's custom model for the Bank of England's Fintech Accelerator is an example of its tailored services. The Bank of England gave Mindbridge Ai approximately 100,000 data points of desensitized, historic regulatory credit union data (going back seven years) to develop an AI-based model for anomaly detection (e.g. reporting errors, compliance issues, and fraud). Mindbridge Ai combined AI and ML with more conventional data science techniques to produce a risk score for each data point, allowing anomalies to be easily identified. Mindbridge Ai's initial project with the Bank of England was successful and has been extended.

Pondera Lab

Pondera Lab is a three-year old data analytics firm based in Mexico City, Mexico[120]. Pondera Lab has 12 staff, with specialties in data-related law, econometrics, and data analytics and science. Its goal is to help private sector firms and government agencies use AI to better organize, analyze, and visualize data to help make better business decisions. As part of this, Pondera Lab provides a holistic suite of consulting and AI-based services in advising clients on how to incorporate new technology to collect and explore data, helping show clients how to plan and strategize using AI and data analytics (including capacity building of technical skills if needed), and providing either off-the-shelf or custom-built AI and ML models for clients. Pondera Lab serves clients in Argentina, Bolivia, the Dominican Republic, Mexico, Panama, Peru, and the United States.

[119] "About – Mindbridge Ai," Mindbridge Ai website, https://www.mindbridge.ai/about/.
[120] "Pondera Lab - about us," Pondera Lab website, http://ponderalab.com.mx/en/about-us-2/.

Data is at the center of Pondera Lab's business. However, Pondera Lab itself does not collect data; instead, it helps its clients develop and use technology to better collect, organize, and analyze their own data. In Pondera Lab's three years, it has found that its service is among the cutting edge and often ahead of where the actual market is in terms of firms and government agencies recognizing that AI and ML can be used to drive efficiency and innovative new services. Pondera Lab found that besides large technology firms, many other large Mexican firms and government departments are still at a relatively low level of awareness about the potential to use data and AI to help their businesses.

At the heart of Pondera Lab's business model and competitive position is proprietary AI and ML models. Pondera Lab uses these to provide either basic data-driven business intelligence models or advanced models that provide predictive abilities and learning processes to drive efficient business services (such as logistics and marketing). A generalized and indicative case for Pondera Lab is a client that already collects, or has the potential to collect (but lacks the technology), a significant amount of data from customers. However, the data is "messy" in that some parts of the firm may be collecting data, while others do not. Different parts of the same firm may use different platforms which do not connect to each other, which can lead to data that is not being collected, aggregated, and stored in a standardized manner. This often leads to data not being analyzed on a consistent basis or in a manner that provides actionable business intelligence.

Certn

Certn is a small (12 full-time staff) start up based in Victoria and Toronto, Canada, that has developed an AI (using proprietary AI and ML systems) and data analytics services (also based on proprietary AI and ML systems) that are focused on helping clients analyze prospective customers, employees, and renters (for example, individuals applying for a loan or bank account, prospective renters, and job applicants) [121]. Certn's AI and data analytics services are hosted on a cloud-based platform. Certn's services collect a wide range of data in order to create comprehensive profiles of prospective customers and applicants and provide (predictive) advice to its customers. At its founding in late 2016, Certn's focus was on helping customers evaluate people's credibility (especially those that do not currently have access to financial services) in order to help promote financial inclusion, while reducing risk for financial institutions, landlords, and employers. Certn's services allow its customers to effectively validate identity, and make better risk decisions while satisfying 'know-your-customer' (KYC) and 'anti-money laundering' (AML) requirements. At the moment, Certn works only in Canada, but it is expanding into the United States.

Certn's two main services are screening, through its main platforms ---"Basic eID" and "Softcheck." Certn provides the rapid screening of employees, contractors, taskers, and tenants by checking for criminal records from around the world, credit reports, and motor vehicle and driver records. It can conduct both basic and enhanced identity verification. For the former, this includes being able to use its "Basic eID" to instantly confirm a person's age and credit details, but extends to using a range of data sources to generate multi-choice questions and answers that only the true identity owner should know. For the latter, Certn allows customers to use any Internet-enabled devices (such as a smart phone) to take a photo of their physical ID and a selfie, which combined with Certn's enhanced e-ID, uses a proprietary mix of AI (including the subfield of computer vision, which focuses on training computers to interpret and understand visual objects) and ID experts to determine if an identity document is authentic and belongs to the user.

Certn's Softcheck identifies risk using real-time public information. It is designed to reduce the instances of high-risk hires, tenants, and customers by delivering automated, intelligent customer

[121] "Why Certn," Certn website, https://certn.co/.

screening, and to provide advice that helps businesses make decision about customers/applicants. Softcheck is updated daily and is curated using both natural language processing (NLP) and ML, with expert oversight from compliance personnel. Certn's API (a key function for software) allows it to provide a "plug-and-play" service to financial institutions, commercial property managers, financial technology companies, real estate technology companies, credit resellers, and others to assist with their risk management, identity verification, and compliance needs. Certn's API gives clients seamless access to their databases.

Softcheck uses data from a range of services, including:

- Criminal Record Searches: Public criminal and court records from around the world including PACER (USA), 350+ courts, boards and tribunals in Canada and records from 240 other economies (Interpol, economy-specific government and state agencies, and police forces).
- Adverse Media Scan: Softcheck uses AI to check over 200,000 sources of adverse media/negative news, sorts relevant articles for risks, and identifies individuals before they appear on a sanctions list or court record.
- Fraud watchlist: Softcheck scans thousands of watchlists dedicated to reporting fraud, including governing regulatory bodies (financial and securities commissions) from around the world.
- Known Affiliation: Softcheck searches police, government, and public databases for known affiliations to gangs, terrorist organizations, and other negative groups.
- Sex Offender Registry Check: Softcheck searches registries for every state, province, and territory in hundreds of economies.
- Public Profile Scan: Softcheck searches social media platforms for public social media profiles, positive news articles, and high-risk behavior.

7.3. Role of Data in Firms' Business Models

Each in their own way, the firms interviewed for this chapter both rely on data in their respective business models and show how they are all focused on specialized services to help other firms use data (and their AI) to help themselves. They are similar in that they are all outside parties offering their AI-based services to clients, rather than developing or otherwise applying AI-based services within established firms. As Certn put it: "business is data."

It is helpful to explain some background on AI and its connections to trade. AI is a branch of computer science devoted to creating computer systems that perform tasks characteristic of human intelligence, such as learning and decision-making. AI overlaps with other areas of study, including robotics, natural language processing, and computer vision.[122]

AI offers many functions:

- Monitoring: AI can rapidly analyze large amounts of data and detect abnormalities and patterns.
- Discovering: AI can extract insights from large datasets, often referred to as data mining, and discover new solutions through simulations.
- Predicting: AI can forecast or model how trends are likely to develop, thereby enabling systems to predict, recommend, and personalize responses.

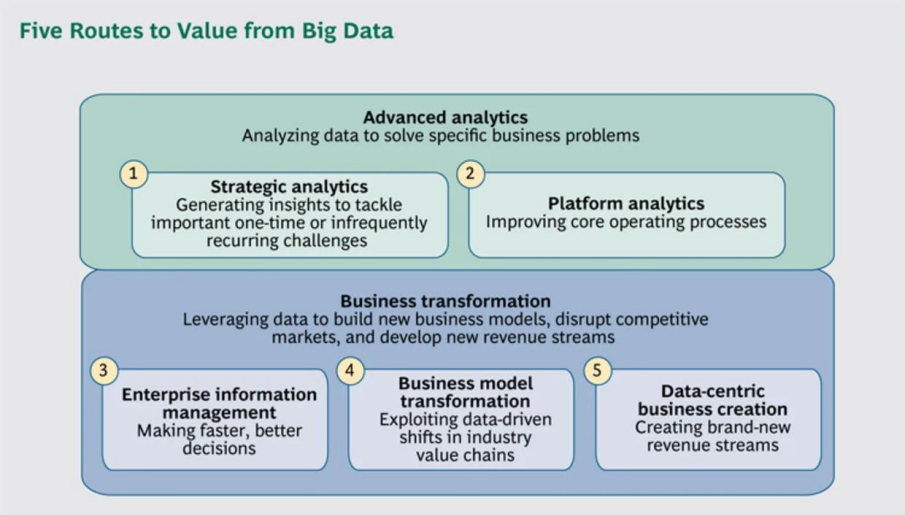
[122] The Information Technology and Innovation Foundation (ITIF). ITIF Technology Explainer: What Is Artificial Intelligence? (Washington, D.C, September 4, 2018), https://itif.org/publications/2018/09/04/itif-technology-explainer-what-artificial-intelligence.

- Interpreting: AI can make sense of patterns in unstructured data such as images, video, audio, and text.
- Interacting: AI can enable humans to more easily interact with computer systems, coordinate machine-to-machine interactions, and engage directly with objects.

Machine learning is an important subfield of AI. It focuses on building systems that can learn and improve from experience without being explicitly programmed with specific solutions. This compares to traditional data analysis, where software aggregates, organizes, and performs basic analysis of historical data for a human to interpret and use as the basis for predictions, insights, or programming feedback. Within ML, an important development is deep learning. Deep learning involves processing multiple layers of abstractions of data and using these abstractions to identify patterns—much like the way people learn through changes in the configuration of the neurons in their brains in response to various stimuli. Obviously, a benefit of ML systems is that they are able to analyze data at a speed, scale, and depth of detail that is beyond human analysis (Reavie, 2018).

These firms all show that AI can be leveraged in a variety of ways to: generate new business insights; improve core operating processes; enable faster, better decision making; take advantage of changing value chains; and create new data-centric businesses. As BCG (2013) points out, competing through lower cost, better products, and innovation are not new, but driving all these with data and AI is now central to a firm's success, especially as data processing and storage costs have decreased by a factor of more than 1,000 over the past decade. While more and more firms are realizing that there is value to be extracted from the massive amounts of data being generated every day, only a few are truly incorporating AI into their business. These firms are some of the latter. Their strategic and tactical application of AI shows how it can affect all aspects of a firm, as Figure 11 below shows.

Figure 11. Example on how Firms Can Extract Value from Data.



Source: BCG (2013)

Furthermore, the broad application of AI by these firms shows that business models and capabilities in virtually every sector of the economy are being reshaped by AI's use of data, with applications that will impact a wide range of sectors from education, travel and leisure, and finance to media, retail, and advertising. Incumbent firms may rely on certain standardized data to make decisions, whereas new firms (such as those profiled here) are using new data sets (such as orthogonal data), which leads to reshaped competition in and between sectors. A general example is that insurance companies are using telematics data (i.e., location/GPS data) to derive insights into customer behavior (such as driving), which is beyond the usual demographic data used for insurance underwriting (McKinsey & Company, 2017). Another example is in marketing, whereby firms use behavioral characteristics to engage in

micro-segmentation[123], in addition to other key demographic characteristics (e.g., breaking a person's geography down from economy to state to city to neighborhood or for behavior that could mean frequent purchases, seasonal-only purchases, or window shoppers).

Interviewed firms' business models differ in terms of where data comes from and how it is used. For Mindbridge Ai and Pondera Lab, they are not data collectors (i.e., not data controllers) themselves but rely on clients to provide access to their data to either help develop or to use as part of their AI and data analytics services. For example, Mindbridge Ai's data-driven business model is indicative in that clients use AI Auditor to process their data, so that they receive benefit from the insights that the platform is able to provide about their data. AI Auditor allows clients to learn at three levels: at the local level in "unsupervised" learning; at the "tenet level" in how clients respond to insights and use it to change business operations; and at the "service level" through curated learning through the insights provided by ML and AI.

Certn is also not a data collector, but a data aggregator in that its services are based around identifying, accessing, and analyzing data from established third-party sources. Certn's value-add derives from accessing a broad range of databases and sources to identify the right person and to correctly attribute information about this person to them when undergoing assessment by a customer.

For all firms, a major challenge is not only collecting data, but in how they organize and analyze it. Their experiences are indicative of the fact that for every firm, in every sector, "big data" means something different in terms of how the data comes from various sources and how it appears in multiple formats. In many cases, data arrives unstructured, which requires firms to develop algorithms to analyze and organize it into useful information. This process lies at the heart of their data-driven, value-added services.

7.4. How Policies and Regulations Impact Firms' Business Models

Firm interviews revealed two main types of rules and regulations that affected their use of AI for digital trade: (i) the rules on data, and (ii) source code protection.

Laws and regulations on data collection, transfer, storage, sharing, and use affected how all of these firms were able to use AI to engage in digital trade. The impact of these laws and regulations, such as for privacy and regulatory oversight, can be either direct or indirect depending on whether the firm is a data controller (i.e., collector and manager of data) as opposed to simply providing data analytic services for clients to use with their own data. However, a clear point that came out of each interview is that AI-based firms need to be supremely vigilant in reviewing the legal and regulatory environment in each market in which they operate to assess compliance-related risks.

Data protection and privacy rules are central to how AI-based firms operate. How economies set the rules about collecting, sharing, and using personal data can have a major impact on AI. In some cases, firms outlined how it reduces the availability of data that AI can use.

Beyond mandatory compliance activity (due to local laws and regulations), policymakers also need to recognize that there is significant pressure from clients about how firms manage and protect data. Many of the firms interviewed mentioned that much of their compliance activity, and parts of their delivery

[123] Micro-segmentation involves layering hundreds or even thousands of data points to identify granular clusters of individuals. See: https://blogs.oracle.com/oracledatacloud/targeting-in-the-age-of-micro-segmentation

of business services, was done to satisfy their clients' perception of compliance risks, even if it was not legally mandated.

All firms mentioned the importance of protecting their AI/ML, whether this is through source code protections, the use of contractual arrangements, and/or the use of technical and administrative controls to manage access and use. This meant avoiding certain markets due to the risk of hosting their AI-based services on local cloud services, due to the unacceptable risk this would pose to their system as they could not trust local cloud providers. Some economies have formal or informal rules/practices that make source code disclosure a requirement for market entry.

Data-related laws and regulations that support the role and flow of data

Data privacy laws

For all firms, data privacy laws and regulations are central to their use of data in the economies in which they operate. Data privacy laws define who is legally authorized to collect, store, and use one's personal information, and they are intended to protect individuals from three types of injuries: harm to one's autonomy (such as involuntary disclosure of sensitive information); discrimination (such as denied access to housing, credit, or employment); and economic harm (such as in the case of identity theft or fraud) (McQuinn, 2018). The starting point for firms in their decision to provide services in other economies (i.e., engaging in digital trade) is conducting a legal review of local laws and to compare these against laws in their home economy and major "benchmark" economies, such as the United States and the European Union. This analysis considers how local laws would impact how the firm typically uses data at home and whether it can (generally) deploy existing data analytic services with no changes.

Pondera Lab uses legal services specializing in data-related issues, especially those related to privacy and the legal framework for how the firm will access and use their clients' data as part of their services (the latter will be addressed below). Pondera Lab manages this for all its clients, whether based in Mexico or elsewhere in North or Latin America. On privacy, Pondera Lab develops individual legal contracts to account for any privacy-related legal requirements of a client's home economy, but this is often built on Mexico's privacy framework, which Pondera Lab considers to be robust. This tailored approach works for most clients given their home economies have compatible and comparable privacy frameworks. If clients have specific privacy concerns, they are able to specify these as part of the contract they sign with Pondera Lab.

Mindbridge Ai builds to a "high water mark" in terms of meeting the strict privacy framework in which it operates, and complements this with externally audited compliance measures, which are then demonstrated to clients in product use and customer service. Given it is the strictest, Mindbridge Ai considers the GDPR to be the "high water mark" of data-protection regulations. The focus on GDPR is also driven by the fact that non-compliance penalties are huge. Beyond Europe and the GDPR, Mindbridge Ai has to manage the challenge of differential privacy protections, such as different state-level requirements in the United States.

Certn's use of data is mainly affected by Canada's federal privacy law as well as some provincial privacy laws (especially in British Colombia, Alberta, and Quebec). Furthermore, Certn keeps updated (and where necessary, makes adjustments) on the latest interpretations of the law, as issued by the Canadian Office of the Information Privacy Commissioner. Certn finds Canada's privacy framework to be generally supportive of data-driven innovation. Certn's internal data privacy and protection procedures are regularly audited by Canadian government authorities. In the case of an audit, Certn's cloud storage provider (Amazon AWS) makes it easy by providing all the relevant documents and certifications about how it stores and protects Certn's data.

From its launch, Certn has worked with Canada's privacy regulators and other government agencies and leading privacy compliance professionals from across North America to ensure its data-

management processes are both legal and ethical. Certn has worked with the U.S. Federal Trade Commission (Fair Credit Reporting Act), the U.S. Department of Housing and Urban Development (Fair Housing Act), and the Canadian Office of the Information Privacy Commissioner. Certn never looks at race, religion, sexual preference, family status, or any other characteristic protected by any economy's human rights legislation. Certn does not analyze photos, nor does it review content that is not public. For example, Certn never looks at social media profiles that are set to "private."

Data storage, data protection, and data accessibility

All firms stressed the importance of understanding data governance issues as part of their daily engagement with customers in their home economies and overseas. Highlighting (again) the importance of cloud services, all firms emphasized that their AI-based services are designed for the cloud, in part, as it provides ready and dependable access, good cybersecurity protections, and scalability to meet client demand. However, as with Pondera Lab, the firms need to tailor these cloud services for each client given local legal requirements, which affect cost, complexity, and accessibility of their services.

Many firms go beyond the strict legal requirements in how they manage and protect their data. For example, Mindbridge Ai stated that their broader approach to data protection emphasizes a focus on both administrative and technical compliance with data protection requirements. Administrative compliance relates to contractual controls about data access, use, and protection a firm sets with its client. Firms use technical controls for data access and protection, such as two-factor authentication and monitoring and logging of data access and use. Furthermore, many firms mentioned their use of third-party certifications to prove that their firm's IT systems, or that of their provider (in the case of cloud storage services), are designed to keep its clients' sensitive data secure.

Pondera Lab relies on globally distributed cloud services. Regarding data access and use, Pondera Lab needs to develop a legal framework with each client to govern how it will access and use their data and how it can deploy or develop AI-based platforms as part of its services. A key question for clients is how to manage cloud service arrangements so that Pondera Lab can access its clients' data. For example, can Pondera Lab use its preferred cloud provider to host its AI/ML platforms, which then needs a legal and technical framework to access data in the client's cloud service provider to develop or provide AI-based data services. Otherwise, Pondera Lab needs to develop a legal framework to manage its access to and use of a client's own cloud service provider in order to develop customary AI/ML platforms and/or upload off-the-shelf data analytics platforms.

Mindbridge Ai uses cloud services with data centers in Canada, the United States, and the European Union as these are the three core markets in which the firm operates. From this base of operations, for every client that operates in another jurisdiction, Mindbridge Ai conducts a review of a client's regulatory environment (at the state/provincial and federal level) to check whether its services comply with local requirements. If Mindbridge Ai does not find any regulatory issues, it will offer to provide its solutions from cloud services from data centers based in Canada. Mindbridge Ai stipulates for these clients that its services will be governed by Canadian law, and if the client does not accept that, then the service is not provided. Thus far, Mindbridge Ai has not had to decline services based on this jurisdictional clarification. Beyond ensuring its services are compliant with (mandatory) local laws, Mindbridge Ai uses voluntary, externally accredited and audited certification measures to demonstrate its commitment to best practices in terms of data protection and security.

Intellectual Property Protections - Mixed

There is no single template for firms involved in AI/ML in how they seek legal protections for their inventions and the underlying data they use.

Given that the intellectual property laying at the heart of their business model is intangible (in the form of source code), it is susceptible to exposure and theft. There are several potential scenarios that pose a risk to source code. There is the threat posed by hackers gaining unauthorized access to the software

hosted on a foreign cloud service provider. At the other end of the spectrum are mandatory source code disclosures, as considered by a number of economies. Mindbridge Ai and Pondera Lab mentioned the importance of implementing specific measures to protect their AI-based services, involving legal, technical, and administrative arrangements.

Mindbridge Ai pursues strict internal control over the intellectual property that lies at the heart of its AI- and ML-driven data analytic services. However, it has not run into IP-related issues that negatively affect its operations. Yet, Mindbridge Ai is aware of the potential risks to its IP in certain markets that it does not currently operate in. The source code at the heart of Mindbridge Ai is largely protected in how its software is developed and deployed. Mindbridge Ai only does service development on data centers based in Canada, and therefore, is protected by Canadian law. Mindbridge Ai also employs internal controls to carefully manage product development so that source code is not inadvertently exposed and therefore copied or reverse engineered. Furthermore, Mindbridge Ai files relevant IP filings to ensure its products are protected from hostile filings. As with most cloud-based platforms, the compiled form in which its product/code is deployed to cloud services in Canada, the United States, and the European Union for clients to use and feed data into means that it is largely protected from reverse engineering.

Pondera Lab recognizes that the custom models it develops and uses are the result of its intellectual capital, so it needs to ensure that it uses IP protections to maximize the value from them. To protect its algorithms, Pondera Lab registers relevant IP in the United States. While U.S. law explicitly mandates copyright protection for software[124], actual protection of software has been significantly limited due to case law (i.e., the U.S. Supreme Court's Alice Corp. v. CLS Bank case). Copyright protects against literal infringement of the text of the program. In this regard, source code can be protected under copyright as literary work.[125] Pondera Lab registers its IP in the United States and sees the provisions on source code protection in new trade agreements (The Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) and the United States-Mexico-Canada Agreement (USMCA)) as holding the potential to change this situation in some economies.

Data-related laws and regulations that limit the role of data

Privacy

When considering providing services in a particular economy, firms outlined how they each assess whether the new market has privacy rules which require changes to their systems. Each firm then weighs up the market opportunity against the impact (technical and financial cost) of having to modify its core services to account for these local laws. The latter may be higher than the former, especially in the case the contract is likely to only be a one-off.

AI-based services are a highly competitive sector and so having to modify fairly standardized cloud-based analytics platforms for each and every economy may not be viable for every firm given the 'back end' cost in terms of ICT services, data engineering, and other activities that are related to providing the firm's services. This highlights the importance of economies of scale to AI-based models and the incremental impact that differential data-related laws and regulations have on a firm that provides AI-based services, in that major requirements (such as local data storage and source code disclosure) obviously entail significant costs, but that smaller requirements in aggregate can be just as significant a barrier to entry.

[124] https://www.law.cornell.edu/uscode/text/17/101
[125] https://www.law.cornell.edu/uscode/text/17/101

Mindbridge Ai's strategy is to essentially develop its AI-based services so that they are compliant with the strictest privacy framework in the market in which they want to target—i.e., to build to a "high water mark." However, Mindbridge Ai does run into privacy legislation that can be problematic to comply with, so it sometimes avoids working with clients in certain sectors. For example, America's HIPAA requirements have constrained its ability to provide its solutions in America's health services sector. In some cases, Mindbridge Ai has provided its Auditor AI service to clients that are covered by HIPAA, but in order to do this, the client has requested that Mindbridge Ai not use its U.S.-based cloud services, but instead deploy its AI-based platform onto the client's on-premise data center. These types of requests are not strictly required by law, but are based on how the client chooses to be compliant with HIPAA. These custom deployments add initial costs to deployment (as opposed to using cloud services), but the real costs and complexity come after deployment, when Mindbridge Ai has to provide software updates and customer support services, as it does not have remote access to the client's platform, for example, to check the access logs and other details when something goes wrong. This adds considerable technical difficulties, especially if Mindbridge Ai were to accept such requirements from a growing number of clients, as it would effectively remove a major benefit of using a centralized, cloud-based service.

Certn's general view of data privacy at the domestic level in Canada and the United States is that it is generally supportive of data-driven innovation. At the broadest level, data privacy across both economies is similar in that they mention similar things, but in slightly different ways, which is where complications arise. The big difference is that managing these differences across the 10 provinces of Canada is much easier than across the 50 states of the United States. For example, two Canadian provinces require local data storage for personal data (outlined below). In the United States, different state laws and regulations affect fairly standard employee data in very different, often complicated, ways. The challenge in navigating these differences is that it is quite common for customers to be considering people/applicants who have lived in multiple states, therefore meaning the (same or similar) data for a single person can be simultaneously governed by multiple laws/regulations. Furthermore, U.S. states manage access to data in very different ways, which also complicates integrating the same data sources into a single platform. For example, criminal records in California require a manual application process for Certn to gain access, while in Colorado, the same data is available online and is accessible to automated services (such as Certn's cloud-based platform).

Certn does not provide services in the European Union as yet. However, Certn has made the up-front investment to build procedures into its AI-based services with the goal of eventually being GDPR-compliant. This has required a significant investment of time, money, and effort, including seeking outside legal and privacy consultants and expertise[126]. GDPR sets a high bar that is (at least initially) difficult for a small start-up like Certn to meet. During its early, formative stage, accounting for GDPR compliance was challenging and costly as it required Certn, as a start-up, to have the compliance regime of a big firm. This extended to physical security requirements at Certn's office and doing extensive background checks on its own staff. Furthermore, it required Certn to have full-time staff dedicated to

[126] For example, the GDPR calls for the mandatory appointment of a data protection officer (DPO) for any organization that processes or stores large amounts of personal data, whether for employees, individuals outside the organization, or both. In terms of potential the potential cost and impact of having a DPO, one study from the University of Milan Biocca, Ca' Foscari University Venice, and the Denver-based Analysis Group estimated that if the data protection officer provisions of the European Union regulation are implemented as written, it would cost each effected European small- and medium-sized enterprise as much as €7,200.00 in additional compliance costs per year. See: Lauritis R. Christensen, Andrea Colciago, Federico Etro and Greg Rafert, *The Impact of the Data Protection Regulation in the EU* (Denver: Analysis Group, 2013).

data protection and security at an early stage, which is a major cost. In essence, GDPR increases the up-front cost of entry for a start-up AI-based firm like Certn. However, in making an assessment of the market risk and opportunity, Certn judges that making this investment will eventually pay off in being able to target bigger clients by being able to show that the company complies with the GDPR. Certn's long-term strategy in aiming for GDPR compliance is that meeting such a "high water mark" standard will make it easier in expanding to other economies with a similar, but less burdensome, privacy framework, such as Australia and New Zealand.

Data storage, data protection, and data accessibility

Requirements to store data locally pose a major risk to AI-based firms as they cut them off from the data that lies at the heart of their business model. A growing number of economies are enacting data localization, for a variety of reasons, such as privacy, cybersecurity, digital protectionism, and guaranteed government access to data. AI benefits from the quantity of data (e.g., merging of data sets from different economies etc.), but also the diversity of data, in that AI predictions will be better if the algorithms have access to a greater range of data. Economies which enact data localization policies limit the ability of foreign AI firms to achieve economies of scale, while protecting the ability of local AI firms to exploit local economies of scale, but at the cost of lower-quality AI predictions and services.

A major issue for Certn is that two Canadian provinces have implemented laws mandating that personal data held by public bodies such as schools, hospitals, and public agencies must be stored only in Canada. The British Columbia Freedom of Information Protection of Privacy Act and the Nova Scotia Personal Information International Disclosure Act apply to personal information in the custody or control of public bodies. This requires Certn to store personal data locally in these two provinces, which it does through its primary cloud provider (Amazon's AWS), which happens to operate data centers in these two provinces. Therefore, Certn is fortunate in that this local data storage requirement does not disrupt its IT systems in a significant way. However, as it expands, it does raise the issue of cost and complexity from using multiple data centers (even if via a central provider) and not being able to aggregate all its data.

Mindbridge Ai uses cloud services based in Canada, the United States, and the European Union as these are its three core markets. Mindbridge Ai stipulates that for clients outside these economies its services will be governed by Canadian law, and if the client does not accept that, then the service is not provided.

Pondera Lab uses a global cloud storage service. In limited circumstances, Pondera Lab's clients (mainly some Mexican government agencies and financial firms) require that their data only be stored in Mexico. Local data storage requirements disrupt Pondera Lab's use of its preferred (in terms of cost, accessibility, and security) cloud service provider.

Furthermore, in limited circumstances, clients will specify that Pondera Lab will only be able to access their data from their premises, which is a major operational barrier to its services, given it relies on remotely accessible cloud services. The framework the client wants has a significant effect on the cost and complexity of the service Pondera Lab provides. Many of Pondera Lab's clients, at least initially, are concerned about sharing data with the firm. This is understandable, as these firms need to ensure their data (whether personal, operational, or transactional) is protected and secured (e.g., not disclosed to competitors). Clients need a clear understanding about how Pondera Lab will use and protect their data.

7.5. Conclusion

As this chapter shows, there are a range of data-related laws and regulations that can affect how firms using or offering AI provide their services. AI-based firms stand to be major players in digital trade as technological change reshapes economies. It holds enormous potential to drive productivity and

innovation in service sectors, which comprise a growing share of many economies. Being based on low-cost, scalable global platforms means that AI-based services can basically be delivered from, and-to, anywhere. These interviews are indicative of the fact that just because a technology or service may be globally deliverable and accessible, does not mean that it ends up being that way due to government laws or client perceptions of regulatory risk. As this chapter shows, there exists a range of data-related laws and regulations that can affect how AI-based firms approach digital trade in providing their services and products across borders. At its heart, this chapter shows how certain laws affect the critical concepts of economies of scale and scope that are critical to the use of AI in digital trade.

A key theme in the interviews is that in cases where there are local data requirements, firms need to weigh up whether the compliance risk and cost is less than the market risk (in terms of bringing its service into line with a jurisdiction's laws and regulations in order to serve clients in that market). As one of the firms stated: there are two demands for data protection and security requirements, to be secure and to feel secure. Firms do the former as a matter of fact, but also have to do a range of things for the latter that may not be strictly necessary, as there is a market risk in clients not willing to take on additional perceived risk. But this highlights the broader impact that data-related rules and regulations can have on how firms manage data in that even if economies do not call for explicit local data storage requirements and call for the free flow of data, barriers to data flows may be the de facto result due to compliance risks. The indirect impact leads some firms to specify requirements above-and-beyond what is legally required, which affects how firms use data.

Designing data privacy and protection frameworks involves a complex process that must address a wide range of legal and regulatory issues. Economies of all sizes and levels of development are grappling with this challenge, which is understandable given the impact digital technologies have had on our societies and economies. The challenge for policymakers is to fully understand digital technologies and balance various competing goals, such as consumer privacy, productivity, and innovation.

The interviews highlighted that getting this balance right is a major challenge. Despite the significant benefits to companies, consumers, and economies that arise from the ability of organizations to use, share, and analyze data, including through AI-based technologies, a growing number of economies have enacted or are considering policies which may act as a barrier to AI-based digital trade.

Another key issue is the impact of data protection laws. A number of firms mentioned that they often have to work with policymakers, especially outside their core markets, who may misunderstand data privacy and data protection. Some policymakers justify restrictive data privacy laws and data localization requirements on the premise that they want their citizens' data to be protected by the laws of the economy. But as interviews with these firms demonstrate, the location of personal data storage is separate to holding the firm responsible for its management of personal data originating from an economy. If a firm has a legal nexus in an economy, the laws and regulations of the economy apply. That was most definitely the case for each firm interviewed. Similarly, for data protection, many policymakers associate the geography of data storage with cybersecurity. But the confidentiality of data generally does not depend on the geographical location where information is stored. Data security depends very much on various factors including the technical, physical, and administrative controls put in place by the service provider. For the firms interviewed, they relied on global, best-in-class providers in part as they recognize their ability to provide a high level of cybersecurity protections.

Firms also mentioned the cumulative impact of data-related laws and regulations. Due to the APEC members covered (those of North and South America), the firms interviewed detailed issues (mainly) about North America, Latin America, and the European Union. However, they also referenced the impact of policies they had encountered outside these regions, especially when considering expansion to new economies. This highlights the importance of adequate IP protection, especially source code protection. This is important, as most economies in North America and the European Union have a predictable and stable business environment, including a strong rule of law. When dealing with sensitive and potentially significant technologies, firms are understandably reluctant to enter into non-core

markets where they are not assured of the same degree of control and predictability. What this means is that firms may decide not to enter certain markets or take on one-off clients from economies outside these regions due to the regulatory uncertainty and risk and due to the corresponding cost and complexity involved in tailoring (often global) IT systems for local conditions. This highlights the broad spectrum of considerations that AI-based firms face in deciding whether to engage in digital trade—explicit vs. implicit requirements, regulatory compliance vs. market opportunity, client vs. government-driven requirements.

At the strategic level, given the critical role of data and emerging competition in AI, it is also worth considering the longer-term implications of divergent data-related policy frameworks in the key markets. Different laws and regulations can advantage AI firms in some economies, given the impact they have on economies of scale and local externalities, while disadvantaging foreign firms. The central role of data to AI means future trade policy will likely focus on these points of friction and/or interoperability. Where it is the former, current literature already shows that policies which limit services trade, for example by restricting market entry and foreign investment in services markets, or by impeding online cross-border supply, constrain the development of the digital economy (Roy, 2017). But a growing number of economies have started the process of enacting rules that protect data flows and address other AI-related trade issues, such as source code protection. Which side prevails in setting the global standard on data and digital trade will play a part in determining the impact of AI and other data-driven technologies in driving economic productivity and innovation.