Artificial Intelligence and the Book Industry

White Paper

Tom Lebrun
René Audet

About the Authors:

Tom Lebrun is a legal expert specializing in copyright and digital law. As a doctoral candidate at Université Laval, he is currently studying the subject of text generation by artificial intelligence. He regularly publishes work on the topics of copyright, digital culture and the relationship between law and AI. His research is funded by the Fonds de recherche du Québec – Société et culture.

René Audet is a full professor at the Department of Literature, Theatre and Film at Université Laval. Specializing in contemporary literature and digital culture, he is interested in modern narrative forms, books and digital literature, editorial innovation and issues surrounding the dissemination of scholarly work in a digital context. He is director of both the Québec hub of the Littérature québécoise mobile project and digital publisher Codicille..

Cover Image: © Tom Gauld, "Novel–Writing Algorithm," 2019.

Contents

Introduction

At the moment, artificial intelligence (AI) is a source of fantasy and, often unfounded, anxiety. It seems to find its way into all spheres of society: consumerism, medicine, industry, daily life… And yet, so far, few efforts have been dedicated to its application by different actors in the book publishing ecosystem. This deficit is even more critical since major players of digital (Amazon being the largest) have invested hugely in AI for a number of years now. The objective of this white paper is to illustrate opportunities that technology associated with AI present for the book industry and its stakeholders. Some steps or courses of action could be undertaken immediately in order to take advantage of this tool and respond to growing competition.

After a brief explanation of the meaning of the term artificial intelligence, the book industry's supply chain will be examined in logical sequence, from a book's conception to its sale, by way of the writing, publishing, promotion and distribution stages. Midway between a positive study (of what already exists) and a forecast (of what may occur), we have attempted to provide as many examples as possible of current uses of existing technology. Thus, the objectives of this paper are:

- To provide a snapshot of AI practices already in place in the book industry;
- To provide key understandings of the requirements and limitations of this technology for each sector;
- To suggest steps for identifying business ventures/opportunities founded on AI.
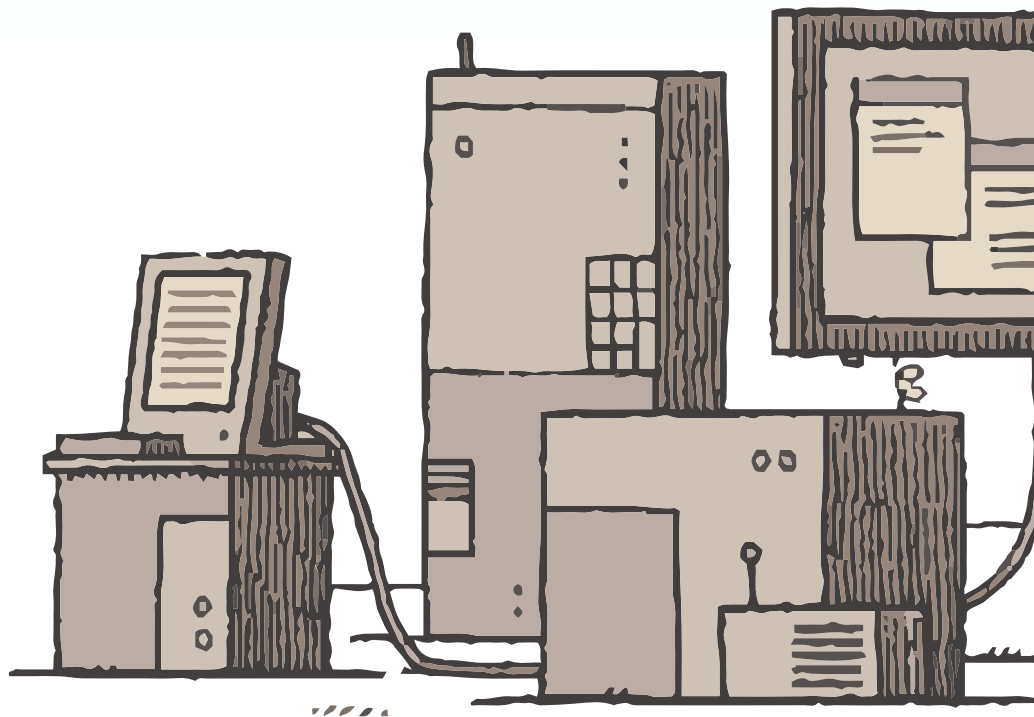
What Is Artificial Intelligence?

Artificial intelligence (AI) is a **generic, all-encompassing** term with many definitions. Experts tend to agree on two main ones: the first presents AI as a system that is able to imitate an activity normally reserved for human beings, while the other defines AI as a system's capacity to react appropriately to its environment (in other words, it is a system able to provide a satisfactory answer to a given problem).

If neither of these definitions are totally satisfying, it is because AI refers to an extremely diverse array of technologies, from the spell check feature in a 1990s word processing program to the self-driving cars we may be riding in one day. For some years now, the discourse around AI has tended to involve **machine learning**, another generic term that refers to many different technologies, commonly referred to as algorithms. An **algorithm** is a bit like a recipe: a sequence of instructions to be followed in order to obtain a result. The particularity of contemporary AI is that some algorithms can now learn to perform new tasks on their own. For them to do this, they must be *fed* enormous amounts of **data**.

Although definitions of AI vary, the objective is always the same: use data to **optimize** certain processes and/or **automate** others. In this sense, AI can be used to discover otherwise invisible patterns that may help improve business intelligence. AI could contribute to the book industry in a great number of ways, notably by means of specialized applications or software: to complete translations, to transform text into audio or to help edit manuscripts. But in any case, AI always requires prior collection of data. Lots of data.

Data Collection, a Prerequisite for Using AI

In general, the greater the quantity of available data, the more likely it is that AI will produce satisfactory results. In this regard, the book industry is in a much more favourable position than other cultural sectors, to the extent that in certain regions of the world, such as Quebec or France, the quantity of available data in some sectors of the book industry might lead one to believe that there is no need to collect more. This impression is misleading: it is in **maximising** the volume of data collected that we can **reveal** business prospects that have thus far been unexploited. But why try to apply AI solutions in the book industry?

AI and the Book Industry: Natural Allies

All economic and cultural fields are different and the book publishing ecosystem has a **unique status**. Often the inheritor of specific cultural policies, the book industry distinguishes itself in its custom of feedback exchange between various actors (the bookseller informing the promoter, the latter informing the publisher, etc.) and even in the solidarity between those who work in this cultural niche that holds significant cultural capital. This custom of **collaboration** has often been developed thanks to specific regulatory structures in the context of growing economic pressure provoked by the dominance of some online retailers, particularly Amazon. This pressure is all the more crushing given that the book industry is very segmented when it comes to different types of publishing (scholarly, literary…) and different economic models (self-published, state-funded, funded by a private business), making **planning** a response to changing business models somewhat difficult.

Nonetheless, the book publishing ecosystem persists and withstands the pressure, undoubtedly bolstered by a shared idea of culture, where competition is offset by community spirit. The daily lived realities are the same: insufficient time to see projects through and the constant search for human, technological and financial resources to ensure operations and to explore structural transformation. Planning for this cultural niche's immediate future must be done and specific actions must be undertaken in order to establish new methods and models. This white paper will outline a possible course of action: the idea of a concerted effort by book industry actors in the use of AI (and even an eventual pooling of collected data).

Definitions

Artificial Intelligence

AI represents a variety of technologies that historically have tried to simulate human intelligence using computers. Today, AI usually refers to machine learning, a field of study founded in statistics with the objective of allowing computer systems to learn independently (without being explicitly programmed for a given task) or under supervision using provided data. To do this, machine learning relies on different kinds of algorithms.

Algorithm

An algorithm is a sequence of operations leading to the solution of a problem or completion of a specific task. Not all AI uses the same algorithms: the algorithms useful for operating self-driving cars are not the same as those used in medicine for cancer detection (though they have many similarities). Most algorithms useful to the book industry fall into a category called Natural Language Processing (NLP). However, other types of algorithms may be used, notably for image recognition (for analyzing book covers, for example).

Data

Data is the base unit of digital information. For instance, a photograph taken with a digital camera is a set of data, as is an e-book or statistics showing the number of books borrowed digitally from a library or the inventory in a book warehouse. The quantity of data circulating in the world, immense and increasing daily, had led us to refer to this time as an era of big data. In the book industry, many types of data could be useful for AI, such as statistical data related to reading, data on sales in bookstores or even data on the content of the books themselves (whether they contain graphic scenes, whether there are strong female characters, whether the language is formal or informal, whether the story takes place in Mumbai, etc.).

Metadata

Metadata is simply data that describes other data. For example, in online communication, the data is the content of the conversation (the content of the e-mail), while the metadata is the information that describes the context of the conversation (the names and addresses of the participants, the size of the message, the date and time of the exchange, etc.). Metadata is useful for organizing information so that data can be used for a particular purpose.

Predicting Reader Expectations

Whether you are a publisher or an author, understanding your audience is essential when making a book. Literary history is riddled with famous examples of publishers rejecting works that were destined for greatness, often due to presumptions about the tastes of their identified readership. For instance, Gide advised Gallimard not to publish Proust… before expressing regret some years later. The most well-intentioned authors and publishers must always wrestle with the **unknowable** when it comes to the books they decide to write and publish: will the text find its audience?

As mentioned in the introduction, contemporary AI essentially relies on analysis of data provided. At the risk of holding them up as examples, companies like Netflix already apply AI technology to data they collect (listening time, preferred content, etc.) in order to **identify** what audiences want to see, which they do successfully. The series *House of Cards* was conceived when Netflix identified an audience that was hungry for both Kevin Spacey and political thrillers. The same system is a possibility in the book industry. For example, it is possible to extract data (purchases, wish lists) from sites such as Goodreads in order to create a topography of what the public wants to read.

In short, AI can help identify reading trends with much greater precision than customary business analysis tools, which are typically simply sales statistics. In this sense, AI is a wonderful tool for discovering **business prospects**. Niche markets that an author or publisher would not have identified as profitable could rise to the surface and find their audience. For example, a romance novel about a pop star with hints of fantasy, or a feminist novel featuring Marguerite Duras in the #metoo era.

Facts and Figures

According to an article in the *New Yorker*, the algorithm presented in the book *The Bestseller Code* could analyze texts and then predict with 80% accuracy whether they had appeared on the *New York Times* bestseller list.

The Needs

Of course, none of this can be done without preliminary collection of **data** on literary works, readers and their reading habits. Some is already available, particularly data from libraries, but crucial collaboration between actors (bookstores, publishers, authors and readers) must be considered in order to take full advantage of current AI technology. This data must be quite **diverse** – on published books, the number of books sold and the number of books read (these are not the same thing), themes, settings where plots take place, time periods, and even more specific data, such as the language register, plot line, the number of characters and their personalities, etc. All information is useful; it is when we cast as wide a net as possible that AI helps us uncover fields for exploration and investment. The condition for success is **cohesion** across the book industry production chain. We have been referring to data from the book industry, but in the era of **big data**, there are other sources, notably information made available by government institutions (data on ideal post-secondary education according to students in France, for example, available at data.gouv.fr). The book industry can make use of this by cross-referencing with its own data, possibly providing it with strategic information, for example cultural trends among young adult readers.

The Limitations

The book industry is a unique ecosystem. We must always keep in mind that avant-garde work, artistic productions deviating from the norm, will inevitably be a **blind spot** for AI. One of the risks is a sort of standardization of cultural production, since AI can only extrapolate from and react according to the data it has been given. As seen in the introduction: **AI is a tool that attempts to use data from the past to predict the future based on probability**. Therefore, AI could not recommend publishing a text by Proust, since the famous author's writing style differs too much from the style of the average text.

Insightful use of AI would mean working in concert with such systems but still thinking critically: AI must remain an **aid** and not an inexpensive means of replacing employees in the book industry ecosystem. In this sense, AI is not so much pure *artificial intelligence* as it is an element of **augmented intelligence**.

Prospects

- AI is used to find **correlations** and can therefore be useful for identifying important associations: for example, that the majority of readers of a certain crime writer are also passionate about historical biographies, which could inspire the writing of a crime novel with a historical figure as its hero or heroine.

- An analysis of readers' wants using AI could detect budding interest in a theme (ecological fantasy, for example) far enough in advance to allow authors or editors to create works on that subject.

- Collecting data on the audience can also serve to optimize communication and marketing using AI (by assessing the potential effect of a work's title on sales in certain target markets, for example).

In-Depth Knowledge: Octavio Kulesz

"The potential of artificial intelligence (AI) in art and culture is, without a doubt, extraordinary. There is already a huge variety of applications offering automated systems made for co-writing, writing articles, proofreading, translating, generating summaries, appraising manuscripts and recommendating texts.

As it stands, advances in AI systems in the field of writing and publishing seem inexorable, implying both opportunities and risks. Use of these powerful tools could lead to higher productivity among authors and publishers, as well as generalization of more efficient marketing methods for reaching readers. However, the amazing capacity of huge digital platforms to connect all the links in the chain could mean that they become inevitable for any creative act in the future. The steady improvement of these AI tools in the hands of a small number of businesses could therefore have a negative impact on employment, replacing proofreaders and translators, but also journalists, publishers and even authors. In the years to come, to guarantee a literary and publishing landscape in which pluralism and diversity have increased, it is essential to protect the cultural ecosystems whose disappearance could lead to irreversible loss, and also to promote the emergence of new applications and local AI startups so that market concentration can be avoided."

Octavio Kulesz is director of Editorial Teseo, a publishing house based in Buenos Aires. He is a researcher in the fields of social media and digital culture in emerging economies and a consultant for UNESCO's Convention on the Protection and Promotion of the Diversity of Cultural Expressions.

Overview

One of the greatest opportunities that AI presents for the book industry is the possibility of analyzing and mapping out literary trends **that do not yet exist**. We can already imagine the creation of tools to target literary themes in unprecedented ways, not only for the general public, but also for niche markets, by searching for more specialized texts. AI could help identify the advantages of writing and publishing unique works, which might otherwise intuitively be cast aside by professionals in the field. To accomplish this, it is imperative that data collection be enriched and not limited to metadata (the genre, author name, date of publication, etc.). Major players in the digital world (like Netflix and Amazon) will no doubt – if they do not already – seize upon the opportunities AI affords to inform artistic creation in advance. It is therefore crucial that the book publishing ecosystem invest in AI as soon as possible.

Recommendations

- Stakeholders in the book industry already possess significant batches of data that represent an important, exploitable asset. It is essential that data collection be greatly **increased and diversified**. The greatest diversity is required, whether it is on reading trends or works currently in circulation (themes, settings, time periods where stories take place, etc.). Diversification of data will lead to a more precise topology of what readers crave, which will ultimately be useful to the whole book industry production chain.

- Possibilities for collection, maintenance and storage must be **considered before being implemented in the book industry supply chain** so as not to be captured by big players (who would not redistribute data or acquired skills, let alone results of analyses). Institutions, libraries and associations have their role to play in this area.

- It is of utmost importance that proposed solutions involving AI be evaluated with a critical eye. AI is a conservative tool by nature, since its findings are always based on data from the past: its use will determine its future impact on the book industry.

Broadening the Horizons of Creation

Can AI write? The question is not as absurd as it might seem at first glance and it's becoming less and less strange, as it's already used daily in the world of journalism and communications. Astoundingly, many existing tools can already write. Your word processor's spell check and the autocorrect on your smartphone fall under the category of AI, as well as the word prediction technology seen in some messaging services. However, these tools are merely aids for writing. Other, much more advanced (energy–and data–hungry) tools literally **write texts**… Sometimes they even compose fiction and poetry, with varying degrees of success depending on the conditions and our expectations.

In 2019, the company OpenAI developed an AI program called GPT-2, which generates fictional or journalistic texts for quick and superficial reading with satisfactory results. The GPT-2 system caused **something of a stir** in the field of AI, since it was generally accepted that mastery of natural language (and therefore of writing) was still a far-off objective. Though the texts generated by the algorithms in question are not perfect (a close reading of examples provided by OpenAI brings to light some incongruities, among other things), this spectacular development **inspires both admiration and curiosity**. Several free, user-friendly tools can give you an idea about their functioning and require little in the way of computer skills:

- OpenAI's Scratch Pad
- Write with Transformer (WWT)
- AI Dungeon

Very recently, OpenAI released a new and improved version of its system called GPT-3. The tests currently being carried out (their programming interface has been made available to the public, though there is a waitlist), show that generated texts seem to have greater coherence. It is difficult to predict future developments in text generation, but although it is unlikely that such texts will be hailed as literary masterpieces anytime soon, they could be profitably integrated into the **mass market** where recurring narrative models are the norm: Harlequin novels, to cite a well-known example, had amply demonstrated the viability of this niche long before the arrival of computer technology.

Facts and Figures

Currently, the most high-performing text generation system requires 45TB of text to function. To put this into perspective, this is the equivalent of roughly 15 000 000 times the entirety of Marcel Proust's *In Search of Lost Time.*

Apart from such uses, it is more likely that AI will be used as a **writing aid**. Some authors already use it this way to produce literature: in Japan, a manga generated by AI entitled Paigon is the product of an aesthetic snapshot of the work of Osamu Tezuka, father of the famous Astro Boy. In California, the software program Granthika allows its creator, award-winning author Vikram Chandra, to map out narrative elements of a text during its creation in order to keep track of complex plotlines and character development. Meanwhile, Montréal poet David "Jhave" Johnston uses AI to generate rough drafts that he then reworks as a poet. To cite another example, novelist Robin Sloan uses AI to generate dialogue for one of his characters (which happens to be an AI program!).

Writing a text represents only the most visible and spectacular examples on the **spectrum of possibilities that AI affords** the book industry. AI can also be used to write custom summaries of novels according to potential readers' tastes and sensibilities or to provide direction for book design or even illustration. *Machines Upon Every Flower* (Gregory Chatonsky and Karmel Allison, Anteism Press, 2018) for example, is a collection of poems generated by AI using one word as a starting point, which were then interpreted by a second AI program to produce images. Another example of application of AI, *1 the Road* by Ross Goodwin, is described as a "novel" written by a car during a journey from Brooklyn to New Orleans. An unabridged version (not proofread or edited) was published by JBE Books in 2018... Again, the possibilities are endless.

Prospects

- AI could lead to development of personified creative aids that imitate an author's "voice" in writing or even drawing, for example.

- AI could allow for discovery or genesis of new creative methods, particularly by cross-referencing data (for example, an AI-driven application could automatically produce a travel diary by commenting on family photos).

- New prospects may emerge from the cross-referencing of data on books, open databases, and data from the Internet of Things of the future. Reading a cookbook could lead to suggested recipes being displayed on a connected refrigerator screen and orders being automatically placed at a supermarket.

The Needs

As with the creation of a topology of readers' tastes, writing by AI requires data, and lots of it. Impressive quantities of text must be analyzed by AI systems, leading to **high costs** in terms of processing power, preparation and maintenance by qualified personnel, but also for the validation and calibration of data. However, the advantage of these systems is that they are **reproducible**: if they are shared, the cost of development could be considerably diminished.

Facts and Figures

Writing generated by AI is also used in journalism. Approximately one third of news published by Bloomberg is machine-generated.

The Limitations

In order to prevent any infractions and in the absence of established laws and legal precedents when it comes to using AI on data protected by copyright, it is recommended that publishers and authors use **standard contracts** that explicitly outline how works can be mined and analyzed for the purpose of developing and improving AI-driven solutions. It is recommended that the data referred to in such contracts be either from the public domain or governed by Creative Commons licenses that allow for reuse and modification.

For Example: Fable

Bringing your favourite literary characters to life digitally and allowing you to interact with them: this is Fable's experiment. By increasing the number of possible communication media (virtual reality, but eventually smartphones, computers or tablets as well), this company seeks to turn the often impersonal character of the virtual assistants we currently use (Siri, Alexa, Google Assistant, etc.) on its head. This strategy could lead to increased **emotional engagement** with characters that have good qualities as well as faults. So, you could have a conversation with Ahab from *Moby Dick* in your living room, ask Hermione Granger her opinion on American politics or even (more chillingly) have Stephen King's monsters respond to your voicemail on Halloween.

In-Depth Knowledge:
Allison Parish

"In 2020, the use of artificial intelligence in literary composition is entirely mainstream. Technologies such as spell check, word completion, and predictive keyboards (like iOS QuickType), each of which qualify as "artificial intelligence," are as ubiquitous in word processing software as they are in text message apps and e-mail clients. Nearly all published text in the world is composed using tools like these, so it's not unreasonable to claim that nearly all published text in the world is "co-authored" by artificial intelligence. For this reason, the very idea of "computer-generated text" is ultimately incoherent, and represents a failure to trace a text back to the context of its composition, and the human labor that created it. Still, the question of authorship constantly looms over uses of artificial intelligence in writing. In particular, large neural network language models like OpenAI GPT-3 (or its predecessor, GPT-2)—which appear to autonomously produce long stretches of coherent text—spur deep philosophical questions about the nature of authorship, while quotidian technologies like spell check and predictive keyboards do not. In fact, there is little difference between the two from the standpoint of authorship, and text produced by GPT-3 can be understood as "machine generated" only in the trivial sense that an e-mail is "co-authored" by an e-mail client's spell check algorithm. In the case of both spell check and GPT-3, it's important to recognize that these systems are themselves created by people, result from human choices, and carry human biases. For example, GPT-3's engineers made choices about its architecture, and chose which texts to include in the model's training set, and chose how to make their tools available for others to use. I argue that texts produced by GPT-3 are not authored by "the machine" (as is frequently claimed), but instead by an amalgamation of the model's engineers, the writers whose works were appropriated in the training set, and the end user of the model who puts the generated text to a particular use. As a poet, my goal is to invent new forms of language. I'm interested in using artificial intelligence not to replace authorship, but to augment it, and make it strange. Poets have been using statistical and computational techniques to do just this for at least a century, from Tristan Tzara's *How to Make a Dadaist Poem* in 1920, up to Lillian-Yvonne Bertram's remarkable *Travesty Generator* published in 2019. I believe that present-day techniques in artificial intelligence and machine learning have great potential to contribute to this tradition."

Allison Parrish is a computer programmer, poet, educator and game designer whose teaching and practice address the unusual phenomena that blossom when language and computers meet. She is the author of many computer-generated works, including *Articulations* (Counterpath Press, 2018).

Overview

At this time, in a literary context, AI is used more as a writing tool to complete and correct texts than as an autonomous writing agent. It is not likely that this situation will change radically in coming years, even if the possibility should not be excluded. This is reassuring in the sense that AI continues to fall into the category of augmented intelligence, rather than true artificial consciousness. Nevertheless, applications using AI could already be applied to writing in **innovative** ways, unearthing novel writing dynamics or revealing new approaches for the future of books. This is already happening in journalism, where AI produces many articles that adequately summarize sports, financial or political news.

Recommendations

- Standard contracts authorizing use of works protected by copyright for the purposes of private AI research must be put in place. The added value of such contracts is that they would also be of great importance in establishing a coherent strategy for use of AI in the book industry.

- Development of creative or animation tools using AI requires a lot of data and processing power. It is therefore in the interest of actors in the book industry to **share the costs**, which could also be entirely or partially financed by public institutions in order to benefit smaller actors.

Support for Publishers

If there is an aspect of the book industry ecosystem that can truly be transformed by use of AI, it's publishing. Although it may not be totally revolutionary, AI can help improve and optimize many tasks regularly done by publishers. Again, the preferred expression would be **augmented intelligence**: the role of each individual at the heart of the publishing process must remain a vital part of tomorrow's publishing world. Even so, a panoply of AI tools already has the capacity to fulfill several functions.

As it is, AI could be used to **considerably optimize** a publishing house's work, freeing up time and money for other tasks. As well as **improving product listing** (to improve distribution of a publishing house's titles thereafter, of course), translation tools could produce rough drafts in previously inaccessible languages, reducing the time (and therefore money) needed for such projects. AI could also read manuscripts and allow for preliminary classification of proposed texts, for example categorizing them according to genre, making it easier to associate manuscripts with a publishing house's readership. AI could identify optimal timing for publishing a title, produce more accurate **sales predictions** and suggest printing an appropriate number of copies, allowing publishers to take risks when it comes to more challenging, less profitable titles. Other tasks could also be automated, such as attribution of royalties according to number of sales, detection of copyright infractions, or even text formatting according to a publishing house's style guide.

Facts and Figures

Publishers receive a huge number of manuscripts, of which on average only 2% are published. Among these, 1% become bestsellers. Using AI to carry out an initial screening, if only to categorize titles according to theme, would save publishers a considerable amount of time.

Prospects

- Some digital assistants already recommend titles using voice interaction. Simon & Schuster has endowed Amazon's Alexa with a "Stephen King Library" that recommends titles by the author based on user preferences. The smart speaker even asks questions to better calibrate its responses. The program's functions could be extended to recommend works by other authors.

- The rapid evolution of AI in fields such as sound engineering makes it conceivable that it will soon be possible to **automate transformation of manuscripts into audio books**. In China, a company called Sogou already automates reading by AI and reproduces authors' voices. In the West, companies like Lyrebird are doing something similar. It's easy to imagine that we will soon see an application that allows us to select our favourite reader, such as an actor or even a member of our family, to read a new book to us.

- The possibilities afforded by AI can improve the self-publication process. Already used by larger platforms to format manuscripts, applications would translate texts into new languages, giving self-publishing authors access to previously unattainable markets.

The Needs

Publishers are in a **privileged** position for data collection in the book industry, both on titles produced and on the audiences to which they are sold. The structure of this domain and most importantly, the presence of meaningful financial and human resources, make them natural candidates for developing AI solutions. As all AI systems require large quantities of data, the largest publishing houses would likely have an undeniable advantage over the smallest ones, given the scope of their production, and therefore data collection. However, we should not forget that a certain number of businesses already offer AI services at an **accessible** cost to smaller organizations. Whether it is done externally or internally, AI is a **reality** in the publishing world.

Facts and Figures

According to the Association of American Publishers, the audio book is the format currently experiencing the most growth (no less than +36.5% in the first two months of 2019). Developing text-to-speech tools would be an appropriate way to benefit from this groundswell.

The Limitations

By determining the potential market for a given work, AI participates indirectly in the publication of texts. An AI analysis tool might suggest that the manuscript is too long for its identified readership, that the number of pages per chapter be reduced or even that offensive content be revised. AI could just as well suggest not publishing works with significant literary value: as a tool based on statistics, it obviously cannot make an aesthetic judgment. An automated system would quite simply reject the majority of the great authors of the last few centuries, either because the sentences are too long, because the syntax is too complex or because the storyline has no resolution. Similarly, translation done by AI will never rival the creative work of the best translators in the field, even if it can produce rough drafts cheaply. In short, classification done by AI, its analysis of markets and its importance as a publishing aid should always be taken with a grain of salt: ideally **human decisions remain of utmost importance** prior to (for configuration or choice of algorithms) and following automated processing.

Facts and Figures

The number of self-published e-books represented 34% of the American e-book market in 2017. The self-publishing sector currently benefits little from use of AI.

For Example: QualiFiction and Booxby

QualiFiction and Booxby are companies that propose using AI to help publishers evaluate manuscripts. According to Elisabeth Mol of the online magazine *Lettres numériques,* the software program LiSA, developed by QualiFiction, analyzes a manuscript and then provides the publisher with a report on a large number of parameters, such as the level of suspense or the level of intellectual engagement required, the complexity of the sentence structure or even how innovative the text is compared to other texts in the publisher's catalogue. In addition to providing **sentiment analysis**, allowing it to gauge reader's feelings throughout the book as well as their frequency, the system adapts to the editorial line of the publishing house, ruling out youth fiction for example if it only publishes detective novels. LiSA even goes so far as to produce a curved line representing the structure and plotline of the manuscript and assigns nothing less than a "bestseller score" to the text… Booxby is a company offering essentially the same service: text analysis with the purpose of discerning the best marketing strategy for promoting it to the public.

Overview

Use of AI in the publishing world affords obvious opportunities for optimization of certain tasks that could save money and time, both **precious assets** in a cultural sector that is under pressure. Amazon and other players of this calibre are no doubt already capitalizing on the possibilities of AI. Automation of manuscript classification, translation assistance, detailed planning of releases or even possibilities for transforming titles into audio books and other such **promising developments** for the sector should not be left only to big players.

Recommendations

- Information acquired about catalogues thanks to their being processed by AI is precious data. In sharing them with other actors in the book industry ecosystem (for example, through sectoral agreements), this data could become the driving force for considerable **return on investment** for publishers. Pooling of data will lead to economic synergy that should be seriously considered.

- To promote data collection and eventual dissemination to other actors in the book industry, investment from the **public sector** would undoubtedly be critical. Inciting actors to share their data, whether by providing open source tools or by offering financial incentives (tax credits, innovation grants, etc.), would also be an approach worth exploring.

In-Depth Knowledge: Virginie Clayssen

"I can speak to uses I have had the opportunity to observe directly or have seen presented. One concerns discoverability. Technology based on AI can actually contribute to the classification of texts after the fact (in the case of massive digitization) by generating relevant metadata. I am thinking of the work carried out by Pythagoria for FeniXX, a company that manages digitized versions of books from the 20th century that are no longer available, which it does on behalf of publishers. The project produced concrete results: it greatly facilitated production of metadata for the books concerned through their digitization, as well as their classification, allowing these books to have greater discoverability. Another task that we would like to be able to entrust systems based on AI with concerns accessibility of digital books. One of the most onerous tasks when making a digital book accessible for visually impaired readers is integrating descriptions of images when they appear in a text. We can imagine gradually instilling in authors and publishers the habit of including such descriptions in new books at the beginning of the publishing process, but we also need to add such descriptions to all the titles that have already been published and digitized. Both require use of automation that is as complete as possible, which AI allows. Such projects have been undertaken by LIA (Libri Italiani Accessibili) in Italy. EDRLab in France would like to continue its research in this area. Although use of AI calls for constant vigilance (critical reflection on "digital labour," for example), it seems important that actors in the book industry pay close attention to these technological advances, as much to the potential disruptions as to the possible benefits they could entail."

Virginie Clayssen is director of heritage and digitization at Editis and president of both the digital committee of the French Publishers Association and EDRLab (European Digital Reading Lab). An architect by training, she speaks regularly at conferences on digital publishing in France and around the world.

Optimized Promotion and Distribution

Although these two sectors are usually connected, the potential impact of AI is not the same for promotion (the commercial dimension, representation and advertising of books) as it is for distribution (the logistical dimension, storage and delivery of books).

When it comes to **promotion**, sometimes assumed directly by the publisher depending on the sector, potential use of AI is mostly related to automating recommendation systems. For example, the *New York Times* and the *Wall Street Journal* already use machine learning systems to adapt their newsletters according to individual readers' preferences and habits. Not surprisingly, a similar procedure could be considered for promoting books. For example, the **standing order form** from a promoter (perhaps a publishing house) that determines which books will be sent to retailers (bookstores, big-box stores, etc.) can be automated, at least partially, thanks to AI.

This makes it easy to imagine promoters concluding contracts with bookstores to gain access to certain subsets of their data, allowing them both to benefit. The promoter could then see reading trends in the bookstores it services and react accordingly in real time (manually or automatically) by sending titles corresponding to the needs and reading habits of those booksellers, thereby helping them to better manage their stock. If data on books is sufficiently refined and makes categorization according to extremely precise criteria possible (such as the setting, whether the hero or heroine has a dog, the absence of violence, etc.), then the promoter's tools will be all the more efficient and relevant for helping bookstores give recommendations. The promoter will of course maintain its role in the strategic management of book promotion but will eventually benefit from an undeniable advantage through exploitation of available data using AI.

Conversely, when it comes to **distribution**, AI's massive and disruptive incursion is inevitable. While systems that have already been digitized are often based on analysis of representative data sets (extraction of a "core sample" of information related to the system's functioning so that necessary modifications can be made), new algorithms allow us to adapt to the situation in real time, in other words to adapt to the flow of ingoing and outgoing information. For the publishing distribution chain, this means that it would be possible to automate book orders from the retailers to distributors and printers, and even to publishers, in real time. For example, in 2012 Amazon filed a patent for anticipatory shipping. To avoid breakage, which could hinder the smooth operation of the company's entire supply chain, the system anticipates that a part will need to be repaired at a given moment and has it repaired before it breaks. Similarly, the Chinese company JD.com uses drones not only to deliver to remote areas, but also to automate transfer of merchandise between warehouses. Some drones transport up to five tons of merchandise.

From receiving and stocking books to completing orders and delivering them, as well as invoicing, all aspects of distribution could benefit from automation using AI and robotics. This could be a boon for the distribution sector: an AI system intended for promotion and for distribution, linked to data upstream and downstream in the book industry supply chain, could certainly optimize operations in the sector, reducing the infamous 25 to 35% return rates, for example. But for this to happen, cooperation between actors when it comes to collection and availability of data is necessary.

Facts and Figures

According to the 2019 report by the France's Syndicat national de l'édition (SNE), 202 900 tonnes of books were transported from publishers to retailers and 50 370 tonnes of unsold books were sent back between 2015 and 2017. This number could be radically lowered if distribution and promotion were optimized using AI. Local printing (in small print runs) as well as printing on demand are solutions that could work in tandem with this movement to **minimize** the flow of goods, also advantageous from an ecological point of view.

Prospects

- Standard order forms could be managed by AI and adapted to retailers' needs in real time by decreasing or increasing the quantity of books (for example, by generating selective standing orders adapted to clients' wishes for each bookstore in a geographic area).

- Using data collected by retailers, promoters could more successfully determine the type of advertising to use.

- For promotion and marketing, AI makes it possible to find titles similar to the one that must be promoted and to suggest a more appropriate communication strategy, which would considerably reduce the usual return rates.

The Needs

Again, to be useful to the book promotion and distribution sectors, AI must have access to large quantities of data. It is therefore particularly in this two-pronged sector's interest to access retailers' data (visitor numbers, but also data on readers) and publishers' data (describing the content of its books in detail). Agreements must be considered in advance, requiring **collaboration** between interested partners, whether they be promoters/distributors, retailers, publishers, etc. It is not necessary that all actors in the field decide to participate, even if there is certainly strength in numbers: when it comes to introducing AI solutions that will eventually be adopted, this potential asymmetry will obviously be taken into account to establish sensible strategies. The existence of established norms, such as the ONIX format for transmitting data in the supply chain, will no doubt be helpful in this process.

The Limitations

In the publishing sector, good distribution does not mean uniform distribution across a territory. Although AI can help determine the best way to promote and distribute a book, it would be risky to give it complete autonomy. The authors we believe in or who are opening up new horizons in literature, not to mention new market prospects, could be promoted and distributed improperly by automated systems, simply because the data we use to train AI and on which it is founded does not provide the system with any point of reference. Human control, intervention when it comes to the parameters or content suggested by an AI system, is required at **each key stage** of promotion and distribution. For example, when it comes to promotion, an automatically generated standard order form is certainly desirable, but it must remain manually editable for sales representatives according to the needs expressed by retailers. Obviously, they should be able to adjust the generated standard order form (by means of a selective standing order form or a standing order adjustment), since it is possible

In-Depth Knowledge:
Véronique Guèvremont

"Introduction and use of AI in the book industry supply chain raise many questions about the repercussions this phenomenon might have on diversity of cultural expression and about the necessary adaptation of public policy in this area. On the one hand, use of AI presupposes access to a mass of data that not all actors are necessarily willing to share. For example, big data on reader tastes and habits is sought after by authors and publishers in order to generate and sell works likely to attract the attention of a large audience.

However, some data holders, such as public libraries, have legitimate concerns about the impact of sharing such data on the creation of new content and the evolution of cultural diversity. On the other hand, AI and recommendation algorithms can be wonderful tools for increasing discoverability of a variety of literary works in a digital environment. In order to benefit from this potential and develop appropriate algorithms, it is necessary to have access to data generated by readers. It is also important that this data reflects national, regional and local tendencies, so that discoverability does not rely only on data generated by large platforms, which would produce the inverse effect of standardizing reader tastes and habits due to improper use of recommendation systems. In this respect, state intervention in the publishing sector could be desirable in order to support, but also to set limits using appropriate legal instruments on the collection, sharing, analysis and use of data by different links in the book industry value chain."

Véronique Guèvremont is UNESCO chair on the Diversity of Cultural Expressions and a full professor at Université Laval's Faculty of Law and Graduate School of International Studies. Co-head of the "Arts, Media and Cultural Diversity" line of research at the International Observatory on the Societal Impacts of AI and Digital Technology (OIISIAN), she is mainly interested in protecting diversity of cultural expression in a digital environment and in the treatment of cultural goods and services in commercial agreements.

that a bookseller would like to obtain fewer, or even more copies of a title for personal reasons (such as managing stock or local enthusiasm for a book due to an author's visit, etc.).

For example: Tamis and StoryFit

Tamis and StoryFit are companies that use AI for text analysis with the primary objective of improving referencing, thereby enhancing title recommendations for readers. Machine learning and emotional recognition allow **text classification** using keywords (indicating for example whether a title has elements of suspense or is well-written) to ensure better promotion. With these applications, it's possible to give a second life to works that are becoming dated: by identifying specific keywords for the works, the publishing house can effectively **promote** them to the right audience at the right moment (if a related topic resurfaces in the media, for example). Some of these companies go so far as to **automate** attribution of BISAC codes (one of the codes used to classify books) and **recommend** the number of copies to be published and distributed according to the book's content and its potential readership.

Overview

Although use of AI in the book distribution sector has been a reality for a number of years, especially for its biggest actors like Amazon, the promotion sector still has a long way to go when it comes to implementing this technology. Beyond establishing previously unknown correlations between pieces of information, AI can analyze data in real time and therefore help automate some decision-making. Again, **cooperation** between

different actors in the book industry supply chain when it comes to access and availability of data is certainly an avenue to be pursued.

Recommendations

- **Agreements about collection and use** of data and metadata could be made between promoters/distributors and other actors in the book industry, both upstream (publishers) and downstream (retailers, mainly bookstores), in order to create shared software tools and returns on investment benefitting many sectors in the supply chain.

- Exchange protocols between distributors, promoters and clients (such as the NEEDA standard in Quebec) must be **reevaluated** in light of new types of data collection for purposes related to AI.

- For distribution specifically, financial and human investment in modernization of storage spaces will be costly, since it will mostly entail use of robotic equipment. This must be taken into account, notably by **public authorities**, to establish mid- and long-term strategies with small and mid-sized organizations so that they are not left behind in the transition.

Facts and Figures

Using AI algorithms to make individualized book recommendations for readers, German publisher Ebner Verlag recorded conversion rates (in other words, sales after advertising) up to 20% higher than previous rates.

Improving Access to Books

Book sales can take place online or through mass distribution, but here we take a closer look at transactions in bookstores and the public equivalent, borrowing in libraries. There is a point on which online sales differ from sales in bookstores and borrowing in libraries: the setting and the feeling of trust. Other than the obvious value of proximity (and stock immediately available on shelves), the places where book transactions occur have an advantage when it comes to readers: a **personal connection**, founded on trust and competence, is established with booksellers and librarians. In this context, AI at first does not seem like a useful instrument and may even be viewed negatively. This is a reductive view, since in helping with research and recommendations, AI can open up many business opportunities in a sector where sustainable economic solutions are in high demand.

Not surprisingly, the event of purchase or borrowing is the most useful data point for contemporary AI, notably algorithmic recommendation systems. According to figures from the Federation of European Publishers (FEP), **610 000 new titles** were published in Europe in 2017 alone. No bookstore or library can keep such a dizzying number of works in stock, much less employ personnel with memory good enough to recommend them all. AI presents one (or many) solutions to these problems. Though the majority of such applications and software are still being developed, we already have an idea about the potential uses of these technologies. For example:

Facts and Figures

According to a Harvard Business Review study, companies in all fields using AI to improve their sales could increase their number of potential clients by more than 50% and reduce costs by 40% to 60%.

A reader enters a library or their favourite bookstore. They're going on a vacation to Barcelona soon and want to bring some reading, maybe a detective novel. Of course, the way things are now, any good bookseller or librarian will first rely on their memory and experience to give a recommendation. In the future, they could also use AI-driven software. Thanks to a detailed profile of the reader's preferences, the bookseller/librarian of the future will use this software to make the very specific recommendation of a detective novel set in Barcelona (and maybe even excerpts of self-published text that bear resemblance to the reader's trip). Taking it a step further, we can imagine defining even more specific criteria, for example to find a novel that is also appropriate for young readers (so that the client's children can read it afterwards).

To develop such applications or software, trust between the reader and the bookseller or librarian is primordial. The level of trust will determine whether or not the reader will share personal preferences and benefit from new recommendation tools in the future.

Prospects

- If the work of referencing books is carried out properly, we can imagine an AI-driven system that allows booksellers and librarians to navigate existing catalogues in very specific ways according to the needs and wishes of potential readers (a book with lots of suspense, a book that takes place in Japan, etc.).

- The trusting relationship established with the bookseller will allow creation of databases **unique to each bookstore**, while libraries could no doubt pool their resources more easily. Reader profiles could be refined over the years, not only according to books read or purchased, but also according to insightful suggestions by booksellers and librarians. This refining of recommendations (according to reader tastes, by professionals and the computer system) will be all the more relevant and complete, since it is always evolving.

- It would of course be possible to send **automated suggestions** to readers of books that correspond to their profile. Founded on the internally established database, these automated recommendations will be extremely relevant.

- Other types of cross-referencing of data can also be imagined: a bookseller's catalogue could include not only published works, but also self-published works that could be ordered thanks to on demand printing. The sale of self-published works recommended by AI could develop a complementary market, helping to **promote local authors**. There is also the possibility of cross-referencing with data from other cultural sectors, such as cinema or theatre, that will no doubt lead to novel recommendation opportunities.

- A bookstore or library's catalogue could be evaluated in real time in order to generate a list of orders, to plan orders or even to indicate which titles do not sell or get borrowed and suggest they be removed from shelves or (in countries where fixed book price law does not apply) propose reducing the price.

The Needs

Not surprisingly, the software that will assist professionals with book recommendations also requires data. In addition to detailed data on books that are for sale or in the institution's catalogue, information on each reader from the bookstore or library must be compiled and cross-tabulated. By isolating personal information (which allows individuals to be identified and is subject to rules, notably ones that have to do with legislation on the protection of personal data), the proposed analyses would help improve reader profiles and get insight into their reading preferences: the size of book they like to read, their preferred genre, their desire to discover new things versus their desire to stay in familiar literary territory, their desire to read more books with a feminist dimension, etc. This data would simply be collected with the help of an integrated questionnaire for use by professionals, who could manage it with readers' approval.

Facts and Figures

According to Bowker, the official ISBN agency for the United States, more than 1 680 000 ISBN codes were issued in the year 2018, an increase of 40% compared to 2017. This extremely dynamic sector is not fully catalogued or exploited by booksellers (especially because of legal and regulatory frameworks), but AI technology could facilitate this.

The Limitations

The purpose of AI-driven recommendation tools is not to replace booksellers or librarians, but to assist them by allowing them to navigate larger catalogues while refining and organizing knowledge about their clientele for example. Recommendations and reader "profiles" are only **templates** representing limited, partial versions of individual tastes and indexed titles. If this were not the case, recommendations of books sold on the largest commercial websites would hit the mark every time, which they do not. AI recommendation systems are not perfect, which is why the professions of bookseller and librarian will always be relevant. Nonetheless, AI will be an invaluable tool for these professionals because

of the **trust** their clientele has in them. This trust should be preserved and valued, notably by respecting the rules in place regarding client files and personal data. This "augmented" approach is undoubtedly the **secret weapon** of booksellers and librarians against algorithmic competition from the giants of online sales.

For Example: HarperCollins' BookGenie

In 2017, HarperCollins added a recommendation tool called BookGenie to Facebook. Readers seeking recommendations for their next books could click on the "Message" button and begin communicating with a chatbot. Available only on the publishing house's English page, this tool proposed to outdo recommendation systems that were based only on previously-read books by interacting with readers and suggesting titles that took into account their mood and reading preferences. Obviously the chatbot would only recommend HarperCollins titles and not ones from competing publishing houses.

In-Depth Knowledge: Clément Laberge

"Unlike other types of cultural content, books are still difficult for machines to decipher. Their content often remains inaccessible to algorithms that make connections between all types of information found on the web and therefore cannot help to improve their discoverability. In this context, I think we must see AI as above all a powerful means of offering books the place they deserve in this large web of human knowledge and imagination. AI must help equip booksellers and librarians so that they more effectively help readers navigate all written material, which no longer always takes the traditional form of a book and is no longer necessarily distributed by the usual commercial channels. Libraries have a fundamental strategic role to play in AI development in the book industry, mainly because they have access to impressive quantities of bibliographical data and statistics on borrowed items. This data should be shared as much as possible with all local actors in the book industry. If not, AI will remain the privilege of web giants and only be used for commercial purposes. For this reason, libraries should be the cornerstone of any AI strategy for the book industry."

Clément Laberge has worked for several businesses in Quebec and France's publishing sectors. He has played an important role in many large digital distribution projects. Now an independent consultant, he is particularly interested in challenges related to distribution and discoverability of all types of digital cultural content.

Overview

When it comes to sales, AI can help booksellers and librarians find their way in larger and larger catalogues and even open up new markets, such as self-published works, to them. In this case too, it might be better to call it augmented intelligence rather than artificial intelligence. The necessary acquisition of data might seem overwhelming to some, especially when we know that the biggest players have a head start, with Amazon being in the lead. However, we must remember that libraries have already been collecting relevant data on readers for many years, data that could translate into a considerable competitive advantage. Also, bookstores have an undeniable advantage over strictly online businesses, since they can interact with readers, obtaining feedback that is more specific than simply a given number of stars next to titles (for example: "I like the beginning, but not the end," "I don't find the story believable," "This reminds me of another book," etc.). This data, if collected, constitutes an asset the competition does not have access to and therefore represents great value for the rest of the book publishing ecosystem, including publishers, promoters and even authors.

Recommendations

- Whatever they may be, tools for the purpose of improving recommendations and sales using AI can only be relevant if they are the result of **cooperation** with publishers, who are in the best position to produce analyses of data on their titles. On the other hand, bookstores and libraries are in a favourable position when it comes to collecting data on readers. Theoretically, these data sets represent even greater wealth if they are shared between different stakeholders, whether they are independent actors, public institutions or organizations involved in circulation of data on books.

- Public actors and libraries must examine the data that the latter possess. The resulting competitive advantage of this resource should be made available to the book industry, whether by means of special sectoral agreements or by pooling assets.

Literary Prospects, by David "Jhave" Johnston

"Currently, the act of writing literature remains a craft of the mind, heart, and ear. Yet it seems safe to assume that in the near literary future, much (writing technology) will change and much (mind/body/society, love/death/language) will remain the same.

It is now normal to imagine a near-future auto-complete function that is user-configurable. Deep-learning models updating in real time trained on the writer's words (zero or few-shot data). Writers selecting genres, styles, idioms, content-region, and language-complexity from an evolving dynamic topological classification map, and then a custom auto-complete model (the archetypal assistant) generates coherent text. Alternate pathways emerge as branching tentacles.

Beginning writers, and children, learn how to read/write within real-time feedback (learning-surveillance) pipelines. Adult writers become editorial curators, culling an Ai avalanche. The danger here is as it has always been: conformity, predictability, stale replication, content reinforcement loops. The advantage is always-on creativity, the infinite Ai muse. Hybrid man-machine authors. Mass collective Ai-human authorial fields: evolving open semi-auto writing spaces.

Yet, as I often say, contemporary artificial intelligence (mostly neural networks) is feeble at subtle context, narrative flow, psychological acuity, continuity. A new superior subtle nuanced Ai derived from neurological architecture, incorporating embodied experience data, will be needed. Once active, future Ai might eclipse all but the strongest authors in both versatility and speed. Meat-brains then become mere data, and autonomous Ai algorithms may write an elegy (or sports, distractions, leisure, gossip, weather, news) to the ravaged earth, its subjugated creatures, and its diminished collapsing ecosystems."
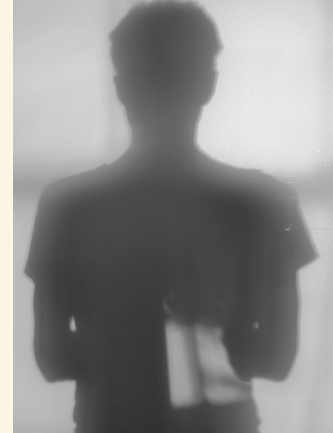
David "Jhave" Johnston is a Canadian poet, videographer, and motion graphics artist working chiefly in digital and computational media. Jhave completed his PhD at Concordia University in 2011 and taught between 2014 and 2017 at the School of Creative Media, City University of Hong Kong, after which he returned to Montréal. He recently published *ReRites* (Anteism Press, 2019).

Conclusion

Artificial intelligence in the book industry is a reality. Players such as Amazon have invested in it heavily for years, as they quickly saw the value of their own data, and will certainly continue to place their chips on it in the future. In the face of this competition, each of the stakeholders in the book industry must think about how to cope with this fundamental movement that will soon profoundly change many jobs in the field. To adapt, we must keep in mind that AI technologies will be even more relevant and effective if they have been well thought out before being applied to the intended ecosystem.

The key to introducing AI to different links in the chain is undoubtedly **exploitation of data that is already available and that the competition does not possess**, even if it means creating asymmetrical agreements between interested players to access it. This data exists and all actors in the book industry, particularly libraries, have access to this precious asset; in this sense they are endowed with considerable strategic resources.

Using this data as leverage, much like initiatives for introducing technology or eventually pooling resources, will of course be met with many obstacles, whether it be the inertia of convention, the cost (in time or in money) of implementation or the state of the law and regulations. These frameworks should not be minimized. Though they differ in every country, they could still potentially serve as springboards for efficient, relevant AI solutions for the book industry if the public actors concerned mobilize around the issue. Over time, creating a network of the collected data will facilitate transmission of information, beneficial for the entire ecosystem.

To ensure good planning and alignment with existing resources, any AI strategy must take into consideration the three following components:

I. ACCESSIBILITY OF DATA	
Evaluating data that is already available	In terms of AI, possession of data that competitors do not have is the biggest advantage. The book publishing ecosystem has a strategic resource that is inaccessible to online players: statistics focused on distributors, data from publishers or even those collected by libraries and bookstores. It is important to examine and become familiar with it.
Establishing relevant partnerships	Each player in the book industry supply chain presides over data that they alone have and has a need for data they do not possess. In this context, making agreements with the most relevant players to one's business model in order to diversify data resources could be a lucrative strategy.
Increasing the quality and the volume of data collected	As operation of AI applications is related to the volume and quality of available data, it is important to plan out an optimal collection scenario by determining which data sets are unharvested and defining information that will be most useful to collect, as well as establishing a strategy to accomplish it.
Finding safe, reliable solutions for data storage	Protection of information is of utmost importance for everyone, but there are different ways to ensure this. A large publishing house would probably not wish to store its data on an external server, while a small bookstore may prefer to share online storage to be more cost-effective.

Having standardized and easily exportable data	The success of AI solutions depends upon the compatibility between different kinds of data it draws on. Establishing protocol for standardization of data in agreement with all players concerned is a necessary prerequisite. Public authorities will have their role to play when it comes to this fundamental question.

2. DATA PREPARATION

Identifying problems that could be solved by AI	An audit of the sector (macro) and of different businesses (micro) must be carried out in order to evaluate which decisions can be automated, which processes can be optimized, which strategic information could reveal important correlations, etc. This step will lead to establishing objectives that will be integrated into the operation of these technologies.
Categorizing and labelling data according to objectives	Collected data is useless without the crucial step of data preparation. Categorization and labelling must be done with a specific objective in mind (which could be intersectoral, between companies or one that benefits only a specific actor).

3. EXPLOITATION OF DATA

Defining priority development objectives	Taking into account the availability of data and the most obvious needs and establishing a plan of action where specific developments are prioritized will allow for formulation of a coherent strategy, notably when it comes to data collection and subsequent standardization.
Involving relevant authorities	In order to direct common projects or to help finance development of shared prototypes or software, institutions and government actors must be mobilized. Pooling of data calls for cooperation, notably for software development (also made possible with the creation of a favourable regulatory framework).
Identifying relevant AI solutions	Since AI is a constellation of different technologies, computer expertise is required at this stage to evaluate implementation of AI solutions, which could be quite vast and require specific software development or draw on preexisting solutions.

Bibliography

Sofian Audry, «*for the sleepers in that quiet earth.*: Experiencing the Behavior of a Deep Learning Neural Network Agent through a Generative Artbook», *Proceedings of the 24th International Symposium on Electronic Art*, 2018, p. 270-277.

Alexandra Bensamoun & Joëlle Farchy, rapport de mission *Intelligence artificielle et culture*, pour le compte du Conseil Supérieur de la propriété littéraire et artistique, 2020, 106 p., https://www.culture.gouv.fr/Sites-thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-enjeux-juridiques-et-economiques-de-l-intelligence-artificielle-dans-les-secteurs-de-la-creation-culturelle

Margaret Boden, *AI: Its nature and future*, Oxford University Press, 2016, 208 p.

Annemarie Bridy, «Coding creativity: copyright and the artificially intelligent author», dans *Stan. Tech. Law Review.*, 2012, 28 p.

Dominique Cardon, Jean-Philippe Cointet and Antoine Mazières, «La revanche des neurones – L'intention des machines inductives et la controverse de l'intelligence artificielle», dans *Réseaux*, vol. 5, n° 211, 2018, p. 173-220.

Matt Carlson, «The robotic reporter. Automated journalism and the redefinition of labor, compositional forms, and journalistic authority», *Digital Journalism*, vol. 3, n° 3, 2015, p. 416-431.

Nathalie Casemajor, Guy Bellavance, Guillaume Sirois *et al.*, «Pratiques culturelles numériques et plateformes participatives : opportunités, défis et enjeux», rapport de recherche, FRQSC, 2019, 192 p., https://tinyurl.com/y87krndu

Grégory Chatonsky & Karmel Allison, *Machines Upon Every Flower*, Anteism Press, 2018, 64 p., https://www.anteism.com/shop/machines-upon-every-flower

Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, Basic Books, 2015, 352 p.

Ross Goodwin, *1 the Road*, [Français, Anglais], Jean Boîte Éditions, 2018, 143 p.

Gould Finch & Frankfurter Buchmesse, «The Future Impact of Artificial Intelligence on the Publishing Industry», *Livre blanc,* 2019, 24 p.

Tristan Greene, «How AI will change storytelling and become the next great art form», *The Next Web,* 2019, https://thenextweb.com/artificial-intelligence/2019/01/25/how-ai-will-change-storytelling-and-become-the-next-great-art-form/

Andrea Guzman, «What Is Human-Machine Communication, Anyway?», *Human–Machine Communication: Rethinking Communication, Technology, and Ourselves*, Peter Lang, 2018, p. 1-28.

Ian Heller, Michael Wu and Austin Garrison, «The Future of AI in Distribution. Artificial Intelligence, Machine Learning and the Implications for Distributors», MDM/Pros, 2019, 18 p., https://tinyurl.com/y73emhje

Véronique Heurtematte, «Intelligence artificielle, une nouvelle donne pour les bibliothèques universitaires», *Livres Hebdo,* 2019, https://www.livreshebdo.fr/article/intelligence-artificielle-une-nouvelle-donne-pour-les-bibliotheques-universitaires

Hervé Hugueny, «Enquête sur l'intelligence artificielle dans l'édition», *Livres Hebdo,* 2018, https://www.livreshebdo.fr/article/enquete-sur-lintelligence-artificielle-dans-ledition

Hervé Hugueny, «Plongée dans l'intelligence artificielle appliquée au livre», *Livres Hebdo*, 2017, https://www.livreshebdo.fr/article/plongee-dans-lintelligence-artificielle-appliquee-au-livre

David «Jhave» Johnston, *Raw Output / Responses (ReRites),* Anteism Press, 2019, 92 p., https://www.anteism.com/shop/rerites-david-jhave-johnston

Howard Lovy, «Artificial Intelligence and the Indie Author: AskALLi Advanced Self-Publishing Salon with Orna Ross and Joanna Penn», *Alliance of Independent Authors,* 2019, https://selfpublishingadvice.org/artificial-intelligence/

Francesco Marconi, Alex Siegman et Machine Journalist, «The Future of Augmented Journalism: A Guide for Newsrooms in the Age of Smart Machines», Associated Press, 2017, https://insights.ap.org/uploads/images/the-future-of-augmented-journalism_ap-report.pdf

Nils J. Nilsson, *The Quest for Artificial Intelligence, A History of Ideas and Achievements,* Cambridge University Press, 2010, 580 p.

Novipro & IBM, *Intelligence Artificielle, Comment débuter un projet d'IA dans votre entreprise,* Livre blanc, 16 p., https://tinyurl.com/y8dqq7zu

Allison Parrish, *Articulations*, Counterpath Press, 2018, 120 p.

Holly Lynn Payne, «Why Book Publishing Seeks Artificial Intelligence», *HuffPost*, 2017, https://www.huffpost.com/entry/why-book-publishing-seeks-artificial-intelligence_b_59482841e4b0961faacbe5d6

Gregorio Pellegrino, «Improving automatic image description in EPUB using Artificial Intelligence», 2019, https://www.youtube.com/watch?v=XZpgGNoBQoo

Adam Rowe, « Interactive Storytelling App Novel Effect Just Raised A $3M Series A », *Forbes,* 2018, https://www.forbes.com/sites/adamrowe1/2018/05/22/storytelling-app-novel-effect-3m-series-a/#1a5c0f1a120d

SAP, « Trends in the Wholesale Distribution Industry. Riding the Waves of Disruption in Wholesale Distribution. Adapting Business-to-Business Paradigms to the Experience Economy », Livre blanc, 2019, 12 p.

Satoshi Sato, « A challenge to the third Hoshi Shinichi award », *Proceedings of the INLG 2016 Workshop on Computational Creativity in Natural Language Generation*, 2016, p. 31-35.

Paul Sawers, « Huawei's StorySign app can translate kids' books into sign language », *Venturebeat,* 2018, https://venturebeat.com/2018/12/03/huaweis-storysign-app-can-translate-kids-books-into-sign-language/

Robin Sloan, « How to Write a Novel with Machine Learning? », Ars Live #25, Interview with Ars Technica, 2018, https://video.arstechnica.com/watch/ars-live-25-how-to-write-a-novel-with-machine-learning

SNE [Syndicat National de l'Édition], *Les Chiffres de l'édition, rapport statistique du SNE 2018-2019,* 24 p., https://tinyurl.com/y9thfejp

Pierre Talbot, *Livre blanc IA,* CCI Essonne, 2019, 20 p.

Vincy Thomas, « HarperCollins mise sur l'intelligence artificielle », *Livres Hebdo*, 2017, https://www.livreshebdo.fr/article/harpercollins-mise-sur-lintelligence-artificielle

Deb Vanasse, « Publishing by the Numbers », *IBPA Independant*, 2017, https://articles.ibpa-online.org/article/publishing-by-the-numbers/

Boris Vian, « Un robot-poète ne nous fait pas peur », dans *Œuvres romanesques complètes,* Tome II, Paris, Gallimard (Bibliothèque de la Pléiade), 2010, p. 1021-1025.

Sophie Viaris de Lesegno, « Les contrats de diffusion-distribution », LEGICOM, vol. 24, n° 1, 2001, p. 29-36, https://www.cairn.info/revue-legicom-2001-1-page-29.htm

"Publishing in the Era of Big Data", Kobo, [White paper], 2014, https://www.yumpu.com/en/document/read/42663323/publishing-in-the-era-of-big-data-kobo-whitepaper-fall-2014