

A Work Project, presented as part of the requirements for the Award of a Master Degree in International Management from the NOVA – School of Business and Economics.

Artificial Intelligence in the Value Chain of Book Publishing

Luca Kraus | 29924

A Project carried out on the Master in International Management Program, under the supervision of:

João Castro

04.01.2019

Abstract

The book publishing industry is threatened by a shift in customer preferences and an increasing number of publishing options. This paper seeks to find an answer to the question of how artificial intelligence can support book publishing to protect itself against these current and future threats. The focus lies on the identification and evaluation of concrete applications of artificial intelligence. Porter's value chain analysis was applied to the book publishing industry and basic concepts of artificial intelligence were analyzed and introduced. Based on these explanations 9 promising applications could be identified: content analysis, automatic language generation, automated proofreading, automated book design creation, personalized advertising, propensity modeling, automated metadata creation, demand forecasting, and conversational commerce. This work contributes to an overall understanding of the opportunities artificial intelligence offers the book publishing industry and identifies concrete applications book publishers should focus on.

Key words: Book Publishing, Artificial Intelligence

1. Introduction

The book publishing process starts with the author. A writer's thoughts and ideas trigger the manuscript process resulting in a finished book the end consumer enjoys. While the writer seeks to earn an income, the writer also aims "to inform, educate, entertain, or to spark a debate" (Greco et.al, 2014, p. 4). To achieve these goals, an author needs someone willing and capable of reviewing, publishing, and distributing the final book to the reader. It is at this point, the book publishing industry comes into play. Book publishers have the knowledge, the resources, and the distribution channels necessary to ensure a successful publication and strong sell-through (Greco et. al, 2014). However, the book publishing industry has been challenged by a shift in consumer's preferences, digitalization, and upcoming new ways of publishing, which make traditional publishing partially or completely obsolete. First and foremost, self-publishing, characterized by rapid growth and threatening traditional publishers in their core competencies (UKSG, 2018). Furthermore, known and established distribution channels and thereby the customer definition itself are challenged. Traditional bookstores in the United States and Europe become less important. The German bookseller market has been facing a decline in sales through physical bookstores for years (Boersenverein, 2018) and large chains like Thalia can only grow in their online segment (Boersenblatt, 2017). Others like the large and well-known American book retailer Borders Group, Inc. were forced to close their bookstores entirely due to bankruptcy in 2011 (Noguchi, 2011). Consequently, establishing a direct connection to end consumers is now more important than ever. The book publishing industry must find solutions to those changes to ensure its survival.

A solution could lie within artificial intelligence. Artificial intelligence is omnipresent today and has been disrupting industries for years. A variety of applications is being used in different industries like retail, banking, and manufacturing and especially sales and marketing departments have been shaped significantly (Heath, 2018, SAS, 2018, Wodecki, 2018). So far, artificial intelligence plays a minimal role in book publisher's strategic development, as

first research shows (Heinold, Spiller & Partner, 2018). Despite the fact that it might enable them to face current and future threats and changes. Consequently, an analysis of possible applications dealing with the potential of artificial intelligence for the book publishing industry seems to be overdue. Therefore, this work offers a first overview of promising developments and recommendations for book publishers.

2. Literature Review & Theoretical Framework

After a thorough analysis was conducted, a necessity to distinguish between research in the publishing industry overall and in the book publishing industry specifically could be identified. Explicit research about artificial intelligence applications in the book publishing industry does not exist. Valuable studies that deal with the publishing industry in general are very limited, but can be found. The German publishing consulting Heinold, Spiller & Partner (2018) introduced artificial intelligence applications for the publishing industry mainly focused on solutions offered by third-party providers. Some of the presented solutions are valuable for the book publishing industry or could be in the future. They are included in the following work (see section 5). However, the introduced applications are only transferable to a limited degree and as such the main purpose of this work is to join the two worlds of artificial intelligence and book publishing. Therefore, the main research question of this paper is the following: *How can artificial intelligence support the book publishing value chain?* The overall objective is to introduce concrete applications of artificial intelligence that can be exploited in the respective steps of the book publishing value chain. Furthermore, explicit recommendations for book publishers will be given with regard to most promising applications and which should be focused.

In a first step, an ideal-model of the book publishing value chain is introduced following the value chain analysis by Michael Porter. Secondly, basic definitions and concepts of artificial intelligence are introduced for the reader to develop an initial understanding of the topic. Thereby, the explanations and descriptions of “Artificial intelligence – a modern ap-

proach” by Russel and Norvig serve as the basis. Lastly, the above-mentioned concrete applications and recommendations are discussed.

3. The Book Publishing Value Chain

It is necessary for the main objective of this work project to identify aspects of book publishing suitable for an integration of artificial intelligence. In order to analyze these particular parts of book publishing, the usage of a framework is crucial to follow a structured and comprehensible approach. Michael Porter’s value chain analysis as a concept of strategic business management meets those requirements and can reveal a perspective of opportunities and challenges (Walters and Rainbird, 2007). The value chain offers the opportunity to single out the “strategically relevant activities in order to understand the behavior of costs and the existing and potential sources of differentiation” (Porter, 1985, p. 33). Following this statement, one can say that the value chain by Porter helps to identify all activities within an organization and reveals their interaction (Porter, 1985). This is highly valuable for the purpose of this work, since a comprehensive understanding of book publishing’s activities and processes is indispensable to establish a linkage to artificial intelligence.

The following can be considered as an ideal-type model of a firm’s value chain in book publishing.

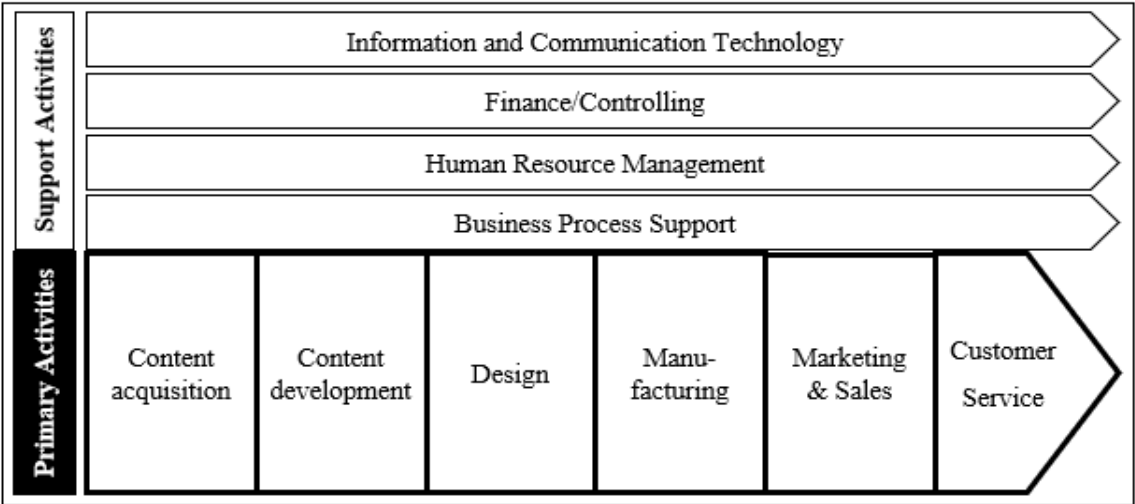


Figure 1: Generic book publishing value chain after Porter (1985).

Porter distinguishes between primary activities and support activities. Primary activities are the core of a firm's value creation process. Support activities include all activities which are necessary and crucial for the core process to succeed but are not part of the value creation process in particular (Porter, 2008).

The support activities in the book publishing industry include information and communication technology, finance and controlling, human resource management, and business process support (Mostert, 2006, Clement and Blömeke, 2009). Since this work is focused on potential applications which are particularly interesting for the book publishing industry the support activities are not being discussed any further.

Primary activities and part of the core process of publishing a book are the following (Greco et alii, 2014; Clement and Blömeke, 2009):

Content acquisition: Content in book publishing is not created by the publishers themselves. The content creator is the author. By acquiring a manuscript from the latter, the publisher gains access to the content. This is one of the most crucial moments in the whole process. An editor reviews the manuscript and evaluates the literary skills of the author and the relevance of the content – the first part of the editorial process. Thereby, a significant challenge is the whole book publishing process itself, or more precisely the length of it. The editor must assess and understand what consumers might be interested in 12 to 18 months ahead (Greco et alii, 2014). Considering the fast-paced time of the 21st century, where consumer's preferences shift quickly, book publishing is challenged by lagging behind current trends.

Content development: The chosen manuscript passes through the second part of the editorial process. The focus is on the further development and enhancement of the content. This includes three major stages: 1. Substantive edit and revision. An editor takes a very close look at the manuscript and the content itself. Small or major changes can be made, some parts might have to be rewritten. 2. Copyedit. In the second step accuracy is focused. This includes

a grammar, spelling, punctuation, and consistency audit. 3. Proofreading. The final step of the editorial process, in which the emphasis is on surface misspelling and grammar mistakes. The editor perfects the manuscript (Laube, 2012; The Writing Center, University of North Carolina at Chapel Hill, 2018).

Design: This step deals with the visual management of the book. The style is conceptualized including layout of book and cover. This is done in alignment with the story, the target group, and the image of the publishing house (Mostert, 2006; Greco et alii, 2014).

Manufacturing: The fourth part of the value chain in book publishing is the actual production of the book itself. Manufacturing comprises prepress, printing, paper binding, or – in case of digital books – the preparation of the e-book. In general, the printing is outsourced as most publisher do not have their own printing plants (Greco et alii, 2014).

Marketing & Sales: Marketing in book publishing includes well-known activities like price setting (mainly restricted in Europe by the book price fixing), advertising, campaign crafting, and promotion affords. It is already present in the content development and design stage. A book's publication is ultimately determined by the marketability of the content it offers (Petersen, 2018). Thereby, marketing to gain the end consumer's attention becomes more and more important. However, publishers exploit the opportunities of sufficient end consumer marketing to a limited degree and it is manly concentrated on a few top titles. Furthermore, as mentioned in the introduction bookstores are less important than before and they used to be the main way of introducing new books to end consumers.

Sales is the process of selling the final product through various channels of distribution to the consumer. The distribution process is “one of the most serious and costly aspects of the publishing business” (Eaglen, 1982) and changed dramatically since the 1970's. Again, the shift in consumer's preferences and the decline of sales through traditional booksellers underlines the necessity to adapt to the new environment.

In brief, book publishers must rethink their marketing and sales strategies and processes to find new and direct ways of reaching out to the end consumer (Forsyth, 1997; Banou, 2016).

Customer Service: The last stage of the book publishing value chain aims to ensure customer satisfaction. This includes handling inquiries, orders, and complaints of any type throughout the whole sales cycle (Greco et. al, 2014).

4. Artificial Intelligence

The following chapter aims to introduce readers to the world of artificial intelligence. Definitions, basic concepts, and methods are discussed briefly to develop an initial understanding of the topic.

4.1 Definitions of Artificial Intelligence

The rise of artificial intelligence results from human's curiosity in understanding how we interact with our environment. How we think, perceive and understand reality. This drive for a deep understanding of cognitive processes is crucial for the field of artificial intelligence. It results in attempts to recreate those processes, to create something that can imitate a human being (Wodecki, 2018). However, human beings are complex and do not always act rationally. Several aspects must be considered. Therefore, a definition of artificial intelligence itself is challenging. It highly depends on what exactly the researcher or creator of artificial intelligence wants to achieve with it and what his or her own understanding of the field is. As a result, there are many different definitions. Most of them view artificial intelligence as a sub-field of computer sciences (Forbes, 2018). For example, the English Oxford Living Dictionary (2018) defines artificial intelligence as "the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages". However, this definition appears to be insufficient to deal with the complexity of the topic. Consequently, one must break down the concept and find a way to order the different aspects and definitions.

Russel and Norvig (2010) developed a promising differentiation. They distinguish between two dimensions: (1) thoughts versus behavior, and (2) human level performance (including human's irrationality) versus rational (ideal) performance. Derived from these two dimensions, four approaches of artificial intelligence can be generated:

1. *Acting humanly (behavior + human performance)*: This approach is based on the Turing test, developed by Alan Turing in 1950. The test is a tool to assess if a machine is able to act and react like a human. A human interrogator must decide whether written answers come from a machine or from a person. If the interrogator cannot decide or fails to decide correctly, the machine passes the test and proof of the existence of artificial intelligence is given (Henry et. al, 2003). To pass the test, a machine needs the following capabilities: natural language processing, knowledge representation, automated reasoning, machine learning, computer vision, and robotics (Russell and Norvig, 2010).
2. *Thinking humanly (thoughts + human performance)*: This group of definitions follows a cognitive modeling approach. A prerequisite is an understanding of human's cognitive processes through introspection, psychological experiments and brain imaging. The goal is to develop a program that shows a human like input-output behavior. The output of the problem-solving process itself, which means if the problem is solved correctly, is not paramount to this approach. The reasoning and steps behind the process and the similarity to a human's solution finding process are decisive.
3. *Thinking rationally (thoughts + rational performance)*: Following the laws of thought¹, this approach focuses on the development of programs that can solve all solvable problems that are based on logical notations. However, this perspective of artificial intelligence makes it very difficult to include informal knowledge and does not take into con-

¹ Any of several principles of logic including the law of contradiction, the law of excluded middle, and the law of identity (Merriam-Webstar Dictionary, 2018)

sideration the difference between solving a problem principally and in practice. Therefore, the applications of this approach are limited.

4. *Acting rationally (behavior + rational performance)*: The fourth approach is based on the concept of rational agents. Rational agents are expected to “operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change and pursue goals”. They act in a way to “achieve the best outcome or, when there is uncertainty, the best expected outcome” (Russell and Norvig, 2010, p. 4). The acting rationally approach includes the laws of thought and components of the Turing test. However, rational agents have two advantages over the others. Firstly, they are based on a more general approach. The overall goal is to achieve rationality. For example, this includes but is not limited to the laws of thought. Secondly, rationality can be based on mathematics. It can be defined and broken down to develop agents. It is not subject to the burden of irrational human behavior.

When considering the research question, the concept of rational agents is the most promising approach to analyze and evaluate artificial intelligence solutions suitable for an implementation in the value chain of book publishing. The outcome a program provides is the critical success factor. This is most valuable for the research question, as it is unimportant if the artificial intelligent used in book publishing (1.) acts like a human, (2.) follows processes similar to human cognition or is (3.) solely amenable to the laws of thought. Consequently, this work project will concentrate on rational agents as the basis for possible applications.

4.2 Rational Agents

The concept of rational agents is complex, and a detailed explanation exceeds this work’s scope. Furthermore, a precise and comprehensive understanding of the concept is not necessary for the purpose of analyzing potential applications of artificial intelligence in the

publishing industry. Therefore, the following solely constitutes a description of the basic assumptions.

As previously mentioned, an agent is expected to receive information from its surrounding, interpret those signals, and act accordingly while adapting to changes and sustain action over a given period of time (Wodecki, 2018). To enable an agent to fulfill those requirements, an agent needs *sensors* to receive information from the environment and so-called *actuators* to act upon the environment. The agent can take specific actions based on the information it perceived through the sensors about the current environment and its knowledge gained through data explicitly given to it or accumulated in past interactions with environments (Wodecki, 2018). In other words, the spectrum of possible actions depends on current information and the agent's history of perceived data. Russell and Norvig (2010) call the latter the percept sequence of an agent. Furthermore, included in an agent's capability set is a measure of efficiency, allowing the agent to evaluate its degree of success at a given task. As the name indicates, the main objective of a rational agent is to act rationally based on the information it has. According to Russell and Norvig (2010) rationality is subject to four conditions: (1) The success defining *measures of performance* (2) The agent's *knowledge of the environment* it interacts with (3) The *spectrum of possible actions* (actions the agent can implement at a given moment) (4) The *percept sequence*

This leads to the following definition from Russell and Norvig (2010) of a rational agent:

“For each possible sequence of percepts, a rational agent should choose the action that most likely maximizes its effectiveness function, taking into account the acquired information about the state of the environment (percepts) and possessed knowledge about this environment” (Russell and Norvig, 2010, p. 37).

One can derive the following conclusions from the definition: (1) A rational agent needs a clearly defined measurement of effectiveness to guarantee a valuable evaluation of the agent's success. (2) The consequence of the agent's behavior must be considered when evaluating the

rationality. An agent acts rationally if the consequences – the changes in the environment caused by the agent’s behavior – maximize the measure of effectiveness. (3) The evaluation system to define whether or not an agent acts rationally is not mainly based on the correctness of reasoning, which stands behind the decision an agent makes. Logical interference is not a prerequisite of rational actions (Wodecki, 2018).

This definition is universal and forms the basis for most applications of artificial intelligence (Wodecki, 2018). Therefore, it is crucial to have an understanding of it in order to proceed with basic methods of artificial intelligence and finally with concrete applications for book publishing.

4.3 Methods of artificial intelligence

In the following chapter a short overview of two basic methods of artificial intelligence is given: (1) Learning, and (2) Natural Language Processing (Russell and Norvig, 2010). The capability of rational agents *to learn* is crucial for every application of artificial intelligence that deals with the interpretation of data and acting on the given information derived from the data (Harrington, 2012). The second presented method, Natural Language Processing, covers the topic of how rational agents understand human languages. Both methods seem to be particularly important for dealing with suitable applications in book publishing.

Learning: As mentioned above, a rational agent must have the ability to interact with its environment and to adjust its behavior based on the data it gathered at an earlier stage or from the current environment it operates in. This adjustment of behavior implies the ability to learn. Therefore, the ability to learn is a crucial part of any artificial intelligence system in order to ensure autonomy and development of the system. If a rational agent was not capable of learning, the developer would have to foresee any possible situation and provide a specific solution. The system would not be intelligent (Alpaydin, 2010; Wodecki, 2018).

Raschka (2015) determines three fundamental types of learning: supervised learning, reinforcement learning, and unsupervised learning.

Supervised learning: This type of learning can also be called learning from examples and rules (Wodecki, 2018). An agent receives a number of input-output rules, information about how to act in a certain situation. In a second step, the agent learns a function that connects the input to the output (Russel and Norvig, 2010). Consequently, supervised learning aims to make predictions about future or unknown data based on labeled training data. For example, the task for an agent is to recognize cats on images. In order to be capable of doing so, the agent must learn what cats look like. In this case the agent would receive images of cats as input and the output would be provided by a teacher, who says, “That’s a cat”. At a certain point, the agent should be able to identify images of cats itself, based on the previous training data. The system solves unanswered problems on its own (Russel and Norvig, 2010; Wodecki, 2018).

Reinforcement learning: In contrary to supervised learning, reinforcement learning is mainly useful in complex scenarios with multiple input parameters. Training data and rules are not reliable enough for an agent to work sufficiently. The agent must perform several actions based on the given inputs (Wodecki, 2018). In such a scenario, the agent learns through interactions with the environment itself and the consequences of its behavior. The decisive factor is a so-called reward signal, which can normally be derived from the current status of the environment. Depending on how well the agent performed a reward or punishment is given. Reinforcement learning follows a trial-and-error approach, an agent learns from a series of reinforcements (Raschka, 2015).

Unsupervised learning: The third type of learning aims to group inputs based on their similarity. The data is unknown, and it is the agent’s primarily task to find a hidden structure. Neither does the agent learn through sample data, nor through explicit feedback. The agent identi-

fies patterns in the input and clusters them into meaningful groups. Clustering helps to identify connections and relationships one could not derive by simply analyzing the given but disordered data (Harrington, 2012).

Natural Language Processing: The ability to understand and interpret received information is crucial for an agent to perform successfully. The research question of this work is how artificial intelligence can support the book publishing industry in its value chain. Information in book publishing is nothing else than written language in form of books and stories. Consequently, an agent must be capable of understanding and interpreting human language to some extent. Natural Language Processing (NLP) “explores how computers can be used to understand and manipulate natural language text or speech to do useful things [...] so that appropriate tools and techniques can be developed [...] to perform desired tasks”. (Chowdhury, 2005, p. 51). This definition seems to be very basic at a first glance. However, it provides a very detailed understanding of what NLP tries to achieve. Thus, NLP is a way to develop agents that can be used in the complex world of human languages. Any agent that can utilize text for a given task belongs to the field of NLP in a certain way.

Liddy (2001) identifies the following applications as the most common ones, confirmed by Russell’s and Norvig’s (2010) explanations: (1) Information retrieval: a user seeks information about a specific topic. The application finds relevant documents that may contain the desired information. Best-known applications which work based on information retrieval are search engines like Google. (2) Information extraction: certain key elements of information are being acquired from different texts. Goal is to identify those elements and find relationships among them. Information extraction includes machine reading, where the system is able not only to extract information from a text, but also to learn by reading the text. (3) Question-Answering: a technique that is similar to information retrieval. However, question-answering systems show explicit text passages where the answer might be found.

(4) Text classification: the system assigns a given text into a predefined category. Well-known examples are language or genre identification. (5) Summarization: a larger text is being summarized in a shorter high-quality version of the original document. (6) Dialogue Systems: systems that are capable of directly interacting with end-consumers. For example, answer inquiries or orders.

5. Applications of Artificial intelligence in book publishing

The previous sections introduced an ideal-model value chain of book publishing and basic models and concepts of artificial intelligence. The following section aims to find opportunities for integration and concentrates on potential applications of artificial intelligence in the book publishing value chain. The main purpose is to answer the research question: how artificial intelligence can be exploited by book publishers. The identified applications will be described and evaluated based on the importance for book publishing and the development status.

(a) Content acquisition:

Content analysis: As previously mentioned, the screening process of manuscripts is a difficult task in the work of an editor. The difficulties lie within the time commitment required and inefficiencies due to the high rejection rate. Artificial intelligence tools can support editors in manuscript screening by analyzing a manuscript with regard to different key factors and evaluating the potential of a manuscript to become a bestseller. The technique behind this is called text mining, which is mainly based in this case on the supervised learning, unsupervised learning, and information extraction mentioned in section 4.3. The main objective of text mining is to discover patterns in texts (Aggarwal and Zhai, 2012). The system receives training data from the developer, bestsellers in this case. Thereby, the agent learns what a bestseller is and can predict if a manuscript has the potential to become a bestseller as well. Furthermore, the system can extract information from the screened manuscript. For example,

genre and sentiment, which can be highly valuable information for publishers in order to evaluate if a manuscript matches their requirements and program. A system based on this method currently exists. The German start-up *QualiFiction* developed a software to analyze texts and predict the bestseller likelihood. After feeding the software with over 15.000 best-sellers, it is now capable to predict the success of a book in form of a score with an accuracy of 80%. The more data the software receives in form of bestsellers, the more accurate it becomes. Moreover, the software gives a detailed analysis of the story development, how it might not match a bestseller's requirements, and the publisher's unique key factors, all within a few minutes (QualiFiction, 2018).

Evaluation: Content analysis tools based on artificial intelligence could be highly valuable for publishers. As a first screening tool, these prediction systems could support editors in the screening process and help to identify manuscripts worth looking at more intensively. An implementation would be possible in the short term.

Automatic language generation: Artificial intelligence or more precisely the underlying algorithms can become authors themselves. Sport reports, product description and financial reports can already be generated by artificial intelligence. Retresco, a company based in Berlin, is one of the leading companies of artificial intelligence driven content automation (Glock, 2018; Retresco, 2018). Furthermore, the publisher ICON Group International has specialized on automatic generated books and has published over 1 million of them so far. Even though, these books are mainly market studies and reference books, the publications also include poetry (Weber, 2018). When it comes to the creation of novels and stories, the application of artificial intelligence is limited, though. The complexity of human languages especially in the form of novels is still something artificial intelligence struggles to reproduce. However, first attempts have already been made and artificial intelligence software is capable of writing shorter stories that come close to human like writing (Nigel, 2018).

Evaluation: Automatic language generation is something which is likely to influence and change book publishing in one way or another and should be followed and focused rather sooner than later. However, as of today concrete applications integrable in the book publishing value chain do not exist yet.

(b) Content development:

Automated proofreading and copy-editing tools: Proofreading and copy-editing are the essential but also very time-consuming parts of the editing process. Artificial intelligence could make the process more profitable by screening and correcting an accepted manuscript and support editors in the first step of a further development. Basic proofreading tools like Microsoft Word are well-known for helping writers to improve the quality of their writing on a basic level. However, artificial intelligence can do more than just checking spelling or punctuation mistakes (Kravariti, 2017). Software like Grammarly, PerfectIt, or Google's Grammar Checking Tool can correct grammar and incorrectly used subordinate clauses, suggest more appropriate words and analyze how professional a text sounds. However, these tools are mainly used for professional emails and text improving so far, as creative writing is also in this form of application an obstacle artificial intelligence has not been able to overcome yet (Fogden, 2018).

Evaluation: Proofreading and copy-editing by artificial intelligence tools provide interesting opportunities for book publishers to consider. Even though, existing editing software is not able to fully replace an editor in the editorial process – and it might never be – basic corrections are already possible today. Editing tools will most likely be developed further in the future and efficiency will be increased during the process (Fogden, 2018). To summarize, the benefits of proofreading tools in book publishing at this point are limited but it is quite certain that proofreading tools will play a role in the future.

(c) Design:

Automated book design and book cover creation: Design seems to be a very tricky task for artificial intelligence as designing is highly associated with creativity. As previously mentioned, human creativity sets limits to artificial intelligence. In fact, no software exists so far which can create the design of a whole book including the book cover. However, artificial intelligence does play a role in design (Vetrov, 2017). For example, website creation. A software build by a company called *The Grid* can retouch and crop photos, run A and B tests, and finally build a website completely on its own. Taking a close look at the applications for publishers, one should mention the advances in creating web page articles. A software developed by *Flipboard Pages* can find the best suited design for a given article depending on the article's content and the learned templates (Vetrov, 2017). So far, artificial intelligence is mainly being used in online design and not in the designing of books. Looking at the creation of book covers artificial intelligence is far away from being able to create a cover itself. First attempts have been made creating agents that can identify a book's genre based on its cover. The method behind this is again supervised learning and the ability to analyze a book's cover and assign it to a given category might be the first step in the direction of autonomously created book covers by artificial intelligence (Emerging Technology from the arXiv, 2017).

Evaluation: Valuable support of artificial intelligence in the design process of book publishing is still up in the air. However, publishers should be aware of the possibilities artificial intelligence offers, even though it might not be feasible at the current stage.

(a) Manufacturing:

Since a significant amount of the manufacturing step, especially the printing itself, is normally outsourced by book publishers, no concrete applications could be identified in this area that only involve the book publishers themselves. However, the issue of hoarding stocks does play

a role in section (e) in form of demand forecasting, which impacts the manufacturing step as well.

(b) Marketing & Sales:

Real-time personalized advertising: In today's world personalization is the key word in business and especially online-marketing. Personalized advertising aims to retarget customers who visited a website and have or haven't bought anything. Based on accumulated data collected via cookies an artificial intelligence software can find the visitors of a website and show them specific ads in real-time based on their interests through different channels (e.g. Social Media). Ideally, they return to the website by clicking on the ad, convert, and purchase something (Smith, 2018). This is broadly used by firms like Google or Facebook and for publishers, this solution offers a way to reach out to the end consumer on a new level. As mentioned in the beginning, the shift from retailers to end consumers as main customers challenges the marketing structure of book publishers. Real-time personalized marketing efforts could be able to strengthen the online performance of book publishers and help transforming their marketing and sales into a new era. *Dynamic Yield* developed a software for this purpose explicitly for publishers. Key features are customer data management, personalized customer targeting and engagement (Dynamic Yield, 2018), which are highly valuable features for book publishers to face the challenges of digitalization. In addition, e-books can be viewed as one channel to get in touch with the customer and offer personalized ads. Based on profile information, purchase history, and browsing behavior tailored offers can be made. This is already being used by publishers to some extent. However, there is a significant growth potential (Abel, 2011).

Evaluation: Personalized marketing applications present an enormous potential for book publishers to reach and engage with the end consumer. Furthermore, it can strengthen their online occurrence and lead more traffic to them directly (e.g. a book publisher's website). Eventual-

ly, this could create more independence from retailers, first and foremost amazon, and the technique does already exist today, is well-tested, and available for an immediate implementation.

Propensity Modeling: The underlying questions behind this marketing technique as a form of predictive analytics are what, when, and why the customer is going to buy. It aims to understand and predict customer behavior. Based on customer characteristics like tendency to open a marketing email, participate in a feedback survey, and type and number of purchases customers are segmented, and their behavior is anticipated. Exemplary segments for book publishing could be frequent book buyers, occasional book buyers, and book buyers at-risk (Lark, 2016). Derived from that information retention activities can be implemented and marketing campaigns' success predicted. This approach rests on a huge, accurate, and up to date data basis and as such the predictions are only as good as the data the systems receives to learn from (Allen, 2017). Book publishers would be capable of aligning their marketing campaigns on the made predictions including analysis of what kind of books certain consumers are interested in, how likely they are to buy again a book from a specific author, or if they might be lost for good.

Evaluation: Propensity Modeling and predictive analytics are a key feature for a future oriented marketing process. However, an outstanding data basis is necessary for an accurate prediction. Due to the years of relying mainly on retailers to sell books to the end consumer, interaction with the latter has been limited. Consequently, customer data information is rare. Thus, gaining valuable data from customers plays a pivotal role in order to benefit from propensity models. A prompt implementation is not feasible (Davenport, 2017).

Automated metadata creation: Metadata is data that describes other data. For example, pictures, documents, or in case of the book publishing industry books and stories. Metadata may include information about the length of a book, the author, genre, and certain information

about content. The latter is often represented in form of specific keywords. Metadata in form of keywords is used by search engines like Google and online retailers like amazon to determine which content is being displayed as a response to a given search query (SEO marketing). Consequently, the discoverability of a publisher's books highly depends on the metadata which influences directly a book's sales and success (Johnson, 2011). At this point, artificial intelligence comes into play. It is not always easy to identify suitable keywords and as of today this process is mainly done manually. However, this harbors the danger of neglecting important aspects of a book that might be able to boost sales. Artificial intelligence solutions can help identifying and extracting the right metadata from books. The software *StoryFit* is specialized on metadata creation for movies and books. Based on storytelling metrics, search engines activity, readers reviews etc. the software extracts keywords that increase the visibility of a given book (StoryFit, 2018).

Evaluation: The importance of metadata cannot be overstated. With the rise of amazon and Google as the number one interaction point with end consumers, book publishers are dependent on sophisticated metadata to gain visibility and discoverability, which is crucial in today's online world. Artificial intelligence can support editors in achieving a higher degree of it through already existing metadata generating software.

Demand forecasting: Inventory management is an important task to perform by book publishers. In the manufacturing step a decision must be made about how many books should be printed initially and a continuous review of stocks is necessary to ensure a smooth and ongoing delivery to the different retailers and end consumers, which is a complex and challenging task mainly based on the expertise of industry professionals. Book publishers make advance payments in this process without knowing with certainty that the ordered number of books will be sold or not, which presents the risk of unsold inventory. Furthermore, booksellers have the right to return unsold books to the publishers, thereby increasing the mentioned risk and

emphasizing the importance of accurate demand planning (Cullen, 2006). Demand forecasting is one of the most promising applications of artificial intelligence. An agent learns from past internal data (CRM and other systems) and external data like social media, market intelligence, and data received by partners. Again, the accuracy of the data is crucial. Based on sophisticated algorithms the agent can now predict an expected demand including real-time adaption with regard to changes in the environment. Artificial intelligence software is able to bring together all the information needed to successfully forecast demand (Glass, 2018).

Evaluation: Demand forecasting with artificial intelligence software presents an outstanding opportunity for book publishers supporting industry professionals in gathering and analyzing available data in a way humans just might not be capable of. However, the technology can only play a supportive role. A human's evaluation and experience are still crucial to make an optimal decision about expected sales and the number of books that should be printed initially. Especially since artificial intelligence-based demand forecasting solutions are still in an early stage of development and most certainly have to be adapted to the book publishing industry (Feizabadi and Shrivastava, 2018).

(c) Customer Service:

Conversational Commerce: The technology enables companies to engage with customers via chat or voice interfaces automatically. It can support and improve the whole customer journey including automatically processing a first order, question answering, and handling inquiries. Conversational commerce is usable on any platform that supports chat or voice interactions (Messaging apps, chatbots on websites, Alexa etc.) and enables a two-way communication. This results in the ability of responding to one's customers' needs and inquiries, while learning from them and interacting immediately, and thus, building a strong relationship. This is very important in today's world where messaging apps like Facebook Messenger and WhatsApp are omnipresent in nearly everyone's life. Therefore, these messenger services

are channels with a very high potential, mostly unexploited by companies (Schlicht, 2018). Especially for book publishers with a missing direct line to their end consumers this technology offers a huge opportunity. Conversational Commerce can work proactively, for example sending a thank you message after a customer bought a book, or reactively (e.g. a customer requests information about new fantasy books). Furthermore, a single customer can be contacted or a specific segment, e.g. sending an announcement to loyal science-fiction readers about the publication of a new science fiction book. As one can see, the possible applications are numerous. If a customer needs to talk to a real human because of a special question or inquiry, the system connects him or her automatically with an employee. All features aim to increase customer experience and satisfaction (Schlicht, 2018).

Evaluation: Conversational commerce is a very promising technology with a bright future ahead. It is most likely to shape marketing and customer services significantly. As of today, it is used by only a few companies. Book publishers should exploit this form of artificial intelligence, as it can result in a higher customer engagement and loyalty, which is highly valuable for book publishers. It offers a new way of customer interaction and opens up new possibilities for publishers. An implementation already makes sense at this point and can be recommended.

Summarizing evaluation: derived from the analysis and evaluations above, one can identify four applications which should be focused. In comparison, the chosen ones are relatively easy to implement in the short-term. (1) *Content analysis* applications present a very promising tool to help editors to fasten and improve the profitability of the editing process, and software does already exist. Book publishers that implement software to support the content analysis process could gain a competitive advantage in the near future, since it takes time to adapt the software to a book publisher's individual needs and to learn how to take advantage of it.

(2) *Personalized advertising* is especially important in face of the missing end consumer connection and can help publishers to gain direct access and data from the end consumers, which will have a considerable importance in the future. In addition, the implementation of propensity models and other marketing applications of artificial intelligence could benefit from it. Personalized marketing seems to be a promising way to lead book publishers into the future of marketing. (3) Improved *metadata* extraction is crucial to increase discoverability and visibility of book publishers and their content. Especially when considering the high dependency on online retailers like amazon. Professional and outstanding metadata might decide about the future of book publishers. (4) Even though *conversational commerce* might not be fully developed yet, it can be seen as the future of end consumer communication. Book publishers should exploit it rather sooner than later, as it strengthens digital appearance and enables data collection, which could increase customer loyalty and sales.

6. Conclusion

The research question of this work is *how artificial intelligence can support the book publishing value chain*. This question has been answered. Artificial intelligence turns out to be a very interesting and promising field for book publishers. Numerous opportunities and potential applications have been identified, nearly in every step of the value chain. Obviously, some of the introduced techniques are more promising and more likely to play a significant role in the future. Especially support applications in content analysis, marketing and sales, and customer service seem to contain the capability to shape book publishing's future positively. However, artificial intelligence is a very new and highly dynamic field and as such additional interesting applications might exist and there will most likely come up even more. The future impact of artificial intelligence cannot be underestimated (Russell and Norvig, 2010). Book publishers must face the changing environment they are operating in and adapt to it. Artificial intelligence applications can help them to successfully master those present or future devel-

opments. At this point, it is of significant importance to clarify that artificial intelligence solutions are far away from replacing humans in book publishing and they might not be able to at any time. Book publishing deals with a very valuable and cultural important asset, which often implies creative writing. As of today, the creation of a program or software that can deal with the complexity of human languages in a similar way humans do is not foreseeable. Consequently, the introduced applications can only take over a supportive role and should not be seen as anything else. They are valuable and promising developments which are necessary to be able to deal with upcoming challenges. In the future, book publishing will neither survive relying solely on people's capabilities nor through automatization of whole processes without human's experience and way of thinking. Professionals must exploit artificial intelligence solutions to support them in their work. This results in the necessity of human machine collaboration trainings (True Anthem, 2018). Employees must be trained in the new ways of work and book publishers can only survive if the latter is done successfully. The industry faces numerous obstacles and changes. Artificial intelligence should be understood as a chance to face those threats and secure the future.

Bibliography

- Abel, Scott.** 2011. "Advertising in Ebooks: Heresy or Genius?" *EContent*. <http://www.econtentmag.com/Articles/Column/Flexing-Your-Content/Advertising-in-Ebooks-Heresy-or-Genius-77412.htm> (accessed November 15, 2018).
- Allen, Robert.** 2017. "15 Applications of Artificial Intelligence in Marketing." <https://www.linkedin.com/pulse/15-applications-artificial-intelligence-marketing-robert-allen/> (accessed November 20, 2018).
- Aggarwal, Charu C., and Zhai, ChengXiang.** 2012. "Introduction to Text Mining." In *Mining Text Data*, ed. Charu C. Aggarwal, and ChenXiang Zhai, 1-10. New York: Springer.
- Alpaydin, Ethem.** 2010. *Introduction to Machine Learning*. Cambridge, Massachusetts: The MIT Press.
- Baird, Henry S., Coates, Allison L., and Fateman, Richard J.** 2003. "PessimPrint: a reverse Turing test." *International Journal on Document Analysis and Recognition*, April 2003, Volume 5, 158–163.
- Banou, Christina.** 2016. *Re-Inventing the Book. Challenges from the Past for the Publishing Industry*. Oxford: Chandos Publishing.
- Boersenblatt.** 2017. "Thalia zieht Bilanz". https://www.boersenblatt.net/artikel-thalia-zieht-bilanz_.1379866.html (accessed December 27, 2018).
- Boersenverein.** 2018. "Geschätzte Umsätze buchhändlerischer Betriebe zu Endverbraucherpreisen." http://www.boersenverein.de/sixcms/media.php/976/Gesch%C3%A4tzte_Umsätze_und_Ver%C3%A4nderung_gesamt_1976-2017.pdf (accessed December 27, 2018).
- Clement, M., and Blömeke, E.** 2009. *Ökonomie der Buchindustrie. Herausforderungen in der Buchbranche erfolgreich managen*. Wiesbaden: Gabler.
- Cullen, Cheryl Dangel.** 2006. "10 Tips for Inventory Management." *Book Business*. <https://www.bookbusinessmag.com/article/book-inventory-management-tips-industry-executives-41637/all/> (accessed November 20, 2018).
- Davenport, Thomas H.** 2014. "Book Publishing's Big Data Future." *Harvard Business Review* <https://hbr.org/2014/03/book-publishings-big-data-future> (accessed November 20, 2018).
- Dynamic Yield.** 2018. "Capabilities." <https://www.dynamicyield.com/media/> (accessed November 20, 2018).
- Eaglen, Audrey.** (1983). "Book Distribution: Part 2", *Collection Building*, Vol. 5 Issue: 3, pp.51-52
- Emerging technology from the arXiv.** 2016. "Deep Neural Network Learns to Judge Books by Their Covers." *MIT Technology Review*. <https://www.technologyreview.com/s/602807/deep-neural-network-learns-to-judge-books-by-their-covers/> (accessed November 28, 2018).

English Oxford Living Dictionaries. 2018. Artificial intelligence. https://en.oxforddictionaries.com/definition/artificial_intelligence (accessed November 17, 2018)

Feizabadi, Javad, and Shrivastava, Apurv. 2018. “Does Artificial Intelligence Enabled Demand Forecasting Improve Supply Chain Efficiency?” [https://www.supplychain247.com/article/does ai enabled demand forecasting improve supply chain efficiency](https://www.supplychain247.com/article/does_ai_enabled_demand_forecasting_improve_supply_chain_efficiency) (accessed November 29, 2018).

Fogden, Tom. 2018. “Could Google’s AI-Powered Grammar Suggestions Change Proofreading?” <https://tech.co/google-ai-grammar-suggestions-spellcheck-proofread-2018-07> (accessed November 28, 2018).

Forsyth, Patrick. 1997. *Marketing in Publishing*. London: Routledge.

Glass, Stephanie. 2018. “AI and the Evolution of Demand Forecasting.” *A Medium Corporation*. https://medium.com/@Aera_Technology/ai-and-the-evolution-of-demand-forecasting-147dd4e783aa (accessed November 22, 2018).

Glock, Sebastian. 2018. “Automated Content Creation Using Natural Language Generation.” <https://www.e-spirit.com/us/blogs/marketing/content-unlocked/2018/08/14/automated-content-creation-with-natural-language-generation.html> (accessed November 27, 2018)

Gobinda, G. Chowdhury. 2005. “Natural language processing.” *Annual Review of Information Science and Technology*. January 31, 51-89.

Greco, Albert N., Milliot, Jim, and Wharton, Robert M. 2014. *The book publishing industry*. New York: Routledge.

Harrington, Peter. 2012. *Machine Learning in Action*. Shelter Island: Manning Publications Co.

Heath, Nick. 2018. “What is AI? Everything you need to know about Artificial Intelligence.” *ZDNet*. <https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/> (accessed November 30, 2018).

Heinold, Spiller & Partner. 2018. “KI Report 2018. Marktüberblick zum Einsatz von künstlicher Intelligenz in Medien und Publishing.“ http://www.hspartner.de/jsp292/source/site/content/veranstaltungen/crossmediaforum/vortrag2018/CMF_20_2018_KI_Studie.pdf (accessed December 1, 2018).

Johnson, Hannah. 2011. “How to sell books with metadata.” *Publishing Perspectives*. <https://publishingperspectives.com/2011/01/sell-more-books-with-metadata/> (accessed November 15, 2018).

Kravariti, Alexandra. 2017. “How good are AI proofreading services tools in 2017?” <https://www.translateplus.com/blog/good-ai-proofreading-services-tools-2017/> (accessed November 28, 2018).

- Lark, Jason.** 2016. "Propensity Modeling: How to Predict your Customer's next Move." *Destination CRM* https://www.destinationcrm.com/Articles/Web-Exclusives/Viewpoints/Propensity-Modeling-How-to-Predict-Your-Customers-Next-Move--114721.aspx?utm_content=buffer2f46d&utm_medium=social&utm_source=plus.google.com&utm_campaign=buffer (accessed November 27, 2018).
- Laube, Steve.** 2012. "The Editorial Process." <https://stevelaube.com/the-editorial-process/> (accessed November 11, 2018)
- Liddy, Elizabeth D.** 2001. "Natural Language Processing." In *Encyclopedia of Library and Information Science*, ed. Miriam A Drake, 2126-2136.. New York: Marcel Decker, Inc.
- Marr, Bernard.** 2018. "The Key Definitions of Artificial Intelligence (AI) That Explain Its Importance." <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/#68a77c254f5d> (accessed November 17, 2018).
- Merriam-Webster.** 2018. Law of thought. <https://www.merriam-webster.com/dictionary/law%20of%20thought> (accessed November 17, 2018).
- Moster, Aléta.** 2006. "Corporate publishing in South African banks: Focus on formal, external publications." Diss. University of Pretoria
- Nigel, Dana.** 2018. "The writer in the machine: Automatic story generation." <http://machineloveus.com/the-writer-in-the-machine-automatic-story-generation/> (accessed November 27, 2018)
- Noguchi, Yuki.** 2011. "Why Borders Failed While Barnes & Noble Survived." <https://www.npr.org/2011/07/19/138514209/why-borders-failed-while-barnes-and-noble-survived> (accessed December 27, 2018).
- Petersen, Valerie.** (2018). "Publishing and Book Marketing." <https://www.thebalancecareers.com/about-book-marketing-2799981> (accessed November 12, 2018).
- Porter, Michael E.** 1985. *Competitive Advantage. Creating and Sustaining Superior Performance*. New York: Free Press.
- Porter, Michael E.** 2008. *Competitive Advantage. Creating and Sustaining Superior Performance*. New York: Free Press. (Orig. pub. 1985)
- QualiFiction.** 2018. <https://www.qualifiction.info/> (accessed November 10, 2018).
- Raschka, Sebastian.** 2015. *Python Machine Learning*. Livery Place: Birmingham.
- Retresco.** 2018. "About us." <https://www.retresco.de/en/about-us/> (accessed November 27, 2018).
- Russel, Stuart, and Norvig, Peter.** 2010. *Artificial Intelligence. A modern approach*. Upper Saddle River: Pearson Higher Education

SAS Institute Inc. 2018. “Artificial Intelligence – What it is and why it matters.” <https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/> (accessed November 30, 2018).

Schlicht, Matt. 2018. “The Complete Guide to Conversational Commerce.” Chatbots Magazine, <https://chatbotsmagazine.com/the-complete-guide-to-conversational-commerce-e47059293efa> (accessed November 30, 2018).

Smith, Jeremy. 2018. “Does Personalized Advertising Boost Conversions? Here’s the Research-Based Answer” <https://www.jeremysaid.com/blog/personalized-advertising-boost-conversions-heres-research-based-answer/> (accessed November 14, 2018).

StoryFit. 2018. “For Publishing:Metadata.” <https://storyfit.com/products/metadata/> (accessed November 20, 2018).

The Writing Center, University of North Carolina at Chapel Hill. 2018. “Editing and Proofreading” <https://writingcenter.unc.edu/tips-and-tools/editing-and-proofreading/> (accessed November 11, 2018)

True Anthem, 2018. “How Artificial Intelligence can make Publisher more profitable.” <https://www.trueanthem.com/true-anthem-ai/> (Accessed December 1, 2018).

UKSG. 2018. UKSG eNews. Connecting the knowledge community. Growth and maturity in the self-publishing industry. <https://www.uksg.org/sites/uksg.org/files/Editorial412.pdf> (accessed November 11, 2018).

Vetrov, Jury. 2017. “Algorithm-Driven Design: How Artificial Intelligence Is Changing Design.” *Smashing Magazine*. <https://www.smashingmagazine.com/2017/01/algorithm-driven-design-how-artificial-intelligence-changing-design/> (accessed November 28, 2018).

Walters, David, and Rainbird, Mark. 2007. *Strategic operations management. A value Chain approach*. New York: Palgrave Macmillan.

Weber, Stefan. 2018. “Eine Million Bücher mit automatisch erzeugten Texten.“ <https://www.heise.de/tp/features/Eine-Million-Buecher-mit-automatisch-erzeugten-Texten-4209972.html>. (accessed November 27, 2018).

Wodecki, Andrzej. 2018. *Artificial Intelligence in value creation. Improving competitive advantage*. Cham: palgrave macmillan