applied sciences

MDPI

Article

Improvement of Business Productivity by Applying Robotic Process Automation

Younggeun Hyun [1], Dongseop Lee [2], Uri Chae [1], Jindeuk Ko [1,*] and Jooyeoun Lee [1,*]

[1] Department of Industrial Engineering, Ajou University, Suwon 16499, Korea; hyunyg@ajou.ac.kr (Y.H.); chaeuri@ajou.ac.kr (U.C.)
[2] Global AI Center, Samsung Research, Samsung Electronics, Seoul 06765, Korea; ds.james.lee@samsung.com
* Correspondence: jdko@ajou.ac.kr (J.K.); jooyeoun325@ajou.ac.kr (J.L.)

Abstract: Digitalization has been bringing about various changes and innovations not only in our daily life but also in our business environment. In the manufacturing industry, robots have been used for automation for a long time, resulting in innovation in terms of the faster operation process and higher product quality. Robotics Process Automation (RPA) can be said to have brought this innovation in the productivity improvement of many industries into the business office. The purpose of this study is to improve business productivity by applying RPA named CoPA. It is based on Domain-Specific Languages (DSLs) and Model-Driven Engineering (MDE) coupled with MS Office. CoPA has been replaced to perform the repetitive patterned tasks (especially document work) done by many people in an office. For the applications of business productivity, CoPA has been implemented to revise five government project proposals requiring quite strict writing standards. The improvement of business productivity obtained by CoPA has been compared to the performance of 10 employees who are familiar with MS Office. The paper explains the method of CoPA coupled with MS Office as well as the agile method of human collaboration. It is clearly shown that CoPA as a business RPA can improve business productivity in terms of time consumption and document quality.

Keywords: business automation; RPA; document automation; correction process automation; software bots

Citation: Hyun, Y.; Lee, D.; Chae, U.; Ko, J.; Lee, J. Improvement of Business Productivity by Applying Robotic Process Automation. *Appl. Sci.* **2021**, *11*, 10656. https://doi.org/10.3390/app112210656

Academic Editor: Emanuele Carpanzano

Received: 22 September 2021
Accepted: 8 November 2021
Published: 12 November 2021

1. Introduction

RPA (Robotic Process Automation) regarding Digital Transformation (DT), refers to the conversion of human jobs in terms of repetitive and patterned tasks to an automatic computer process. In this sense, "robotic" is not a physical robot, but it is a 'computer process' in the sense that it does the same work like a human's cognitive work. The purpose of RPA is to make employees to be concentrated on higher priority work and creative tasks besides low-value-added tasks [1,2]. With the advantage of RPA, many companies are applying RPA throughout the value chain to improve productivity. RPA has been implemented on various tasks to maximize productivity in the fields of production, sales, purchasing, finance, and employee management [3]. The financial sectors are one of the most actively adopted RPA fields where chatbots are utilized to automate simple repetitive tasks. The chatbots can manage customer response and customer service using personalized RoboAdvisor [4–7]. In addition, the medical and legal circles are also actively introducing RPA and producing desired results [8]. One of the representative examples of RPA in the medical field is automated medical appointments according to various variables including doctor availability based on location and treatment attainability as well as financial statements and insurance information. In addition, the RPA robot was applied to the task of providing patient data to the medical analysis service (3rd party) for more accurate and improved patient treatment [9]. The introduction of RPA in the legal industry has been applied to achieve two main improvements: structural improvement of a huge

amount of case law databases and search for similar legal cases. These have dramatically reduced the time and effort required to search vast amounts of existing data [10].

As such, RPA is introduced to improve productivity in various fields, but it may be neglected to improve productivity in the fields of business operations especially documentation. One of the main reasons is that all business filing tasks are assumed to require the cognitive judgment of humans. In fact, human judgment is only necessary to determine the direction and logic of the business document. Furthermore, there are still many repetitive works especially on document revision in terms of both editing and design. The productivity of business document work is the one of core tasks for a company's business. It is due to that document automation using RPA will reduce the time and costs corresponding to manpower for correcting documents. It is expected that both work productivity and document quality will be improved if an alternative RPA is introduced to document work with such patterned work. In addition, it will produce a more thorough document with a lower cost if an alternative RPA system is applied for documentation jobs.

The paper is organized as follows; Section 2 goes over the history of business automation, Section 3 presents the design rationale of document automation, Section 4 presents the method of document automation, the numerical results obtained by CoPA(Correction Process Automation) and manpower are compared in terms of computational efficiency and document quality in Section 5, and Section 6 discusses the limitation and future work of CoPA. Section 7 concludes the business document automation using CoPA.

2. Background of Business Automation

The following section will review the background of business automation. The concept of business document automation will be explained, the background of the rise of RPA and previous research works, and various cases in which digital labor replaced human labor were analyzed.

Major global companies have focused on sustainable cost reduction using ERP (Enterprise Resource Planning) during the 1990s. For labor cost savings in call center and accounting, Offshore BPO (Business Process Outsourcing) is adopted by companies since the year 2000. As of the year 2014, they recently have been using SW-Robot for business automation such as RPA. The main reason for this trend moving from BPO and to RPA is that there is still a lot of business process to improve work efficiency by eliminating non-cognitive working processes repeating patterned tasks. PwC forecasted that the business automation will eliminate the repeated work by 70% and can be reduced 45% of business operation costs which is approximately US$2.0 trillion dollars [8,11,12].

2.1. Emergence of RPA

Digital labor is replacing employees in the process of using various technologies of the 4th industrial revolution, such as big data and Artificial Intelligence (AI), in overall corporate management. This could mean that digital labor communicates with employees in natural language as a new form of labor [1,13]. A German company named 'Zurich Insurance' is a prime example of the introduction of RPA to significantly improve productivity. In 2014, employees of this company had to work on dozens of documents such as hospital certificates and traffic accident statements submitted by customers to pay insurance payments, and 37 out of 81 employees were in charge of repetitive and patterned tasks. Zurich Insurance introduced RPA to handle 51 processes for insurance contract confirmation and compensation payment. As a consequence, twenty-seven human workforces have been relocated to improve service quality work that can increase customer satisfaction as a human labor role in enhancing added value [3,14]. These advantages have become a trigger that RPA tools play a large role in the establishment of business automation. One of US market research organizations called Forrester^{TM} analyzed the functions of 12 leading vendors on seven evaluation standards ('current offering', 'bot development and core functions', 'control room, system management, reporting and resilience', 'RPA analytics', 'architecture', 'breadth of use case', 'deployment, governance and security'). Based on the analysis, RPA

solutions from three companies; UiPath, Blue Prism, Automation Anywhere are evaluated to be the world's best. However, just because the number corresponding to these seven evaluation standards is high that does not mean a good business solution for all companies. It will be the best way to select the RPA solution by adjusting the weight accordingly on whether the solution has considered the characteristics and business strategies of the company or not, and also comparing the characteristics of vendors, performing detailed product evaluation [15–17].

2.2. Spectrum of Digital Labor

Many companies are promoting automation in production, sales and purchasing, finance, HR, and IT by using RPA across the value chain to improve productivity. For example, data input, e-mail reception/sending, report writing, etc., as well as sales reports linked to the computer system, and market trend collection can be performed [3]. In the financial sector, RPA is most widely used; it is applied to banks' non-face-to-face customer response services, and insurance companies' customer relation management and contract management. For repetitive inquiries, chatbots and RoboAdvisors are used for customer services such as retirement pension design, investment portfolios, funds, money laundering monitoring, cybersecurity, etc. [4,5]. In particular, RoboAdvisors uses natural language processing (NLP) and social recognition algorithms of external distribution information based on internal corporate data. It delivers personalized customer services including easy account setup, robust goal planning, account services, portfolio management, and security features, attentive customer service, comprehensive education, and low fees in investment decision-making and asset management fields [4,5]. As a result of using RPA, financial companies could reduce the cost reduction by 20–30%.

Furthermore, SW-Robot coupling with artificial intelligence, innovation is being pursued to improve productivity and service quality in expert work areas. For example, the Anderson Cancer Center in the U.S. used SW-Robot for cancer diagnosis, significantly lowering approximately 20% of the false diagnosis rate by general doctors. It has increased diagnostic accuracy to 98% for colorectal cancer, 91% for bladder cancer, and 94% for pancreatic cancer. Baker & Hostetler, a US law firm, used an SW-Robot named Ross to replace lawsuit case analysis (30% of the total work) that is the most time-consuming work. Helpme, a law firm in Korea, also automates the creation of customer debt lawsuits that offer a 90% reduction cost to US$39 originally cost US$400 [8].

Robots are continuously used for patterned and repetitive tasks to improve productivity in every field. It can be noticed a trend that RPA is to minimize human action for automation of simple and repetitive tasks while IPA (Intelligent Process Automation) is increasingly used for the minimization of human judgment work [8,18,19]. Table 1 shows the comparison of RPA and IPA.

Table 1. Characteristics of RPA and IPA [8,18,19].

Category	RPA	IPA
concept	mimic human action	mimic human judgment
application area	routine, repetitive & rules-based	non-routine, pattern recognition etc.
major role	follow instruction	come to conclusion
market status	mature	emerging
cost of construction (period)	relatively low (weeks)	relatively high (months)

It is mentioned that RPA is not only a tool for the automation of repetitive business processing in the financial sector but also in the medical and legal fields. It has been becoming a trend that the implementation of RPA and/or IPA could reduce the cost of business operation. However, there are not many cases in both academic research and industries where RPA and/or IPA have been implemented to improve the productivity of business document work. Since it is assumed that human judgment is required for business

documentation tasks even though human judgment is only necessary to determine the direction and logic of the document. The business document works are the most common and frequent office tasks in the business operation. This automation will change daily office work and paradigm in the business operation.

The use of business documents is to share business strategies and information including creative ideas and human cognitive judgment. Therefore, the business document, especially project proposals, is thoroughly created and completed by substantial labor forces under an agile working environment. This working system can cause frequent human errors and also the quality of the document is eventually degraded. The project proposals especially for public government require a strict rule including structure, term/word correction, etc. It will be promptly rejected that the documents do not comply with the writing standards shown in RFP (Request for Proposal). It can be one of the solutions that can reduce the risk of rejection.

From the review of prior studies, three distinguished limitations are found; the limit of spelling verification for specific words mentioned in RFP, the limit of general implementation, and the limit of documentation tools as discussed detailly in Section 2.2. Therefore, alternative methods will be required to make automated project business document corrections in both editing and design layout. This paper proposes an alternative IPA named CoPA to improve the productivity of business documents for both editing and design of documentation (Typo Verification, Table of Content Verification, Font Standard Verification, Table Layout Verification, Schematic Diagram Layout Verification). The company can expect to reduce the cost for the preparation of the project proposal using CoPA.

3. Rationale of BDA (Business Documentation Automation)

Based on the technical concept of RPA discussed earlier, this paper will focus on improving the productivity of business documents. In this section, the type of business documents including reports, planning, and project proposals are analyzed for BDA (Business Documentation Automation). The review of advanced BDA research is presented and also the contribution of this research is described.

3.1. Type Analysis of Business Document

The use of business documents is to share business strategies and information including creative ideas and human cognitive judgment and understanding. These business documents generally have a certain format according to their purpose and a type of similar pattern in the flow of content. In this paper, it is largely divided into three types of business documents based on past experiences: reports, planning proposals, and project proposals.

1. Reports

This document is prepared when reporting the progress of a specific task in the company or the result of the analysis. It is a fact-oriented document rather than individual opinions. The core of this documentation refers to the fact based on the numerical result of the analysis.

2. Planning Proposals

This is a document to be prepared when establishing a future plan before pursuing a specific task. It is prepared based on a fact-oriented forecast. The core of this documentation refers to the facts based on the trends or results of survey analysis. These documents are used to deliver marketing plans and promotion plans.

3. Project Proposals

This is a document created when presenting a specific method for solving a problem described in a client's RFP. This document is used to propose a method and/or solution to minimize disadvantages while maximizing advantages. The core of these documents is that they are written based on adequate judgments and analysis about situations, cognitively diagnostic assessment, and creative ideas (counterplan).

Table 2 shows the characteristics of business documents such as the number of pages, required human resources, writing standards, and patterns. Report and planning proposal show similar characteristics including the number of pages, required HR, content, writing standards, and quality inspection method. Above all, the biggest difference between project proposal and report & planning proposal lies in the content and writing standards. For instance, the content of a report & planning proposal is based on the facts obtained through the analysis of situation/numerical results, while the project proposal shows a new idea under the predefined business scope mentioned in the RFP.

Table 2. Characteristics of business documents *.

Category	Report	Planning Proposal	Project Proposal
number of pages	1~10 pages	1~20 pages	100~3000 pages
human resource	1~3 people	1~3 people	2~50 people
content	fact-based status information	fact analysis based future plans	Suggesting business idea based on RFP
writing standard	freestyle without writing standards		the style complying writing standards shown in RFP
quality check	logical validity		document standard compliance

* These characteristics can be different based on the specific company and circumstance.

The project proposal strictly needs to comply not only with correct words/terms but also many items in RFP. For example, font type, table style, schematic style, overall layout/design, and table of contents should be based on the guidelines. In addition, there are many parts of project proposals to be corrected since relatively many people work simultaneously under an agile environment when compared to reports and planning proposals. For a project proposal, it is difficult to keep the consistency of document contents due to the frequent human errors. Among business documents, project proposals have many tasks patterned for repetitive works lowering the work productivity when compared to the report and planning proposals. In this paper, RPA named CoPA will be developed and applied to project proposals to improve business productivity.

The typical correction procedure of business proposals is shown in Figure 1. Many designers are working together to correct the project proposal. First of all, all of the designers need to understand the rules shown in RFP and then designers revise recheck the document. There can be two main correction steps for quality improvement to maintain consistency and make sure compliance with the writing standards; the first step is to correct non-standard words/terms and context errors. The second step is to revise the design/layout of the document including font, tables, and diagrams. Otherwise, the company will use outsourcing which will cost a lot more. It can be seen that all of the correction processes are manually done by people even though the process can be replaced by RPA.

3.2. Review of Advanced Research for BDA

Prior research on RPA based on the 4th industry technology (e.g., AI, IoT, Robot, Blockchain, etc.) has been described only in its maturity and trend of each core element technology [20]. For BDA (Business Document Automation), the prior studies have been conducted in four directions; spelling checking, document generation, applying document standards, and document classification work based on text recognition.

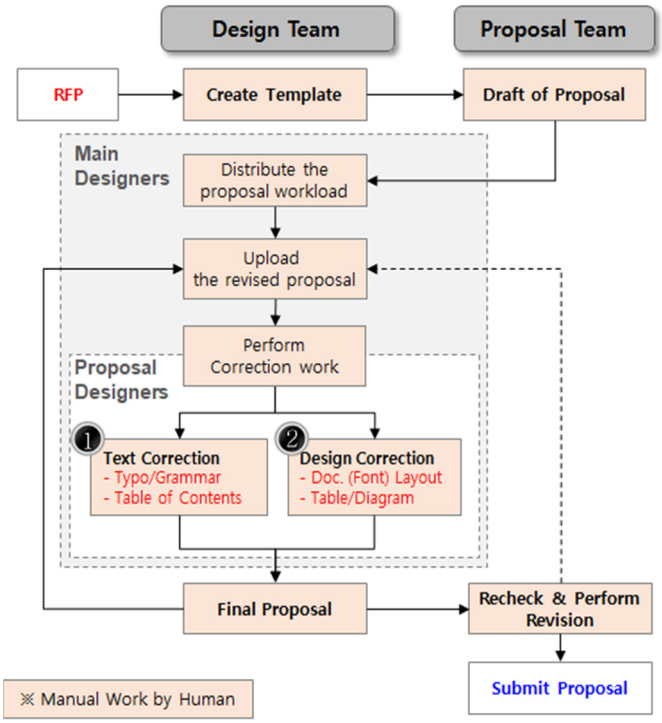

Figure 1. Process of project proposal correction by manpower.

First of all, there are two main studies by Wood Kevin (2020) and Lee et al., (2020) [21,22]. Frank has developed RGL (Restricted Grammar Language) for cloud-based agile document work. RGL consists of a writing rule set that performs real-time spelling corrections. The research of Lee proposed a spell correction method using deep learning based on the evaluation of word relations between surrounding situations. The method can verify two types of spelling errors in terms of word spelling errors and contextual spelling errors.

The second case is an automatic document generation, which means that it automatically creates a court document with standard form besides automatically generating texts with a free form like a novel. Thomas Northoff and Klaus Gresbrand (2020) used the case of a law firm to automatically generate documents first, even if they are of poor quality, and then manually correct them. They emphasize that it is better to improve manually the quality of paper rather than manually creating high-quality business documents [23]. It was possible since most legal and court documents have standardized rules.

Regarding the application of web document standards as the third research case, Liegl et al., (2010) and Wu et al., (2019) reported that the standard of business documents has changed from the existing EDI (Electronic Data Interchange) to XML (eXtensible Markup Language). In addition, a lot of unnecessary manual work occurs in creating XML-based invoices, contracts, and various graphs. Therefore, they proposed an RPA applying document standards to improve the productivity of document work are proposed [24,25]. Chae et al., (2019) analyzed the standardization compliance of Web (World Wide Web) documents assumed as a business document and suggested a way to effectively correct the errors with RPA [26].

The fourth research case is the automation of document classification work through detection and recognition of text described in business documents using BDA. Seraogi et al., (2017) describe a method for recognizing text described in graphic drawing documents [27], and Ma et al., (2020) describe a method for text detection and recognition in the process of digitizing classical documents [28]. In addition, research was conducted to recognize the meaning of the recognized text and automatically classify business documents. Harold Smith (2002) proposes the method recognizing the abstract of a patent and then classifying the technology type by IPC (International Patent Classification). Cho et al. (2020) describe a method of automatically classifying the types of documents through a neural network

for scanned business documents [29], Lee et al., (2018) presents the automatic classification according to the KSIC (Korea Standard Industry Code) [30], and Yun et al., (2018) describes the automatic classification method for business documents for which document classification is not defined [31]. In addition, Tien et al., (2020) describe a method for deriving meaning for unlabeled documents and then performs classification [32].

This paper is highly related to the first to third previous studies for BDA. From the review of prior studies, three distinguished limitations are found: limit of spelling verification for synonyms, limit of general implementation, limit of documentation tools. The method proposed by Wood and Tajeran can only correct typos and contextual wrong words (e.g., some and sum) but does not correct synonyms to keep standard word consistency. For the second case, there is a difficulty in automating project proposals created using the method of Northoff and Gresbrand, as the document standards and rules are various according to the project governing organization. For the implementation of document standards, the methods suggested by Wu's studies are only for XML-based web business documents since XML has already been standardized and can be easily automated. In other words, the methods of Wu cannot be implemented in MS Words or PowerPoint. Therefore, alternative methods will be required to make fully automated project proposal corrections for both editing and design layout. This paper proposes an alternative IPA named CoPA (Correction Process Automation) to improve the productivity of business documents for both editing and design of documentation (Typo Verification, Table of Content Verification, Font Standard Verification, Table Layout Verification, Schematic Diagram Layout Verification).

4. Method of CoPA

In this section, the method of CoPA applied to correct the project proposal is described. It is explained how CoPA is designed and able to be coupled with Microsoft Office, especially with PowerPoint. In addition, five correction features of CoPA are explained with examples.

4.1. System Environment of CoPA

CoPA is currently based on Model-Driven Engineering (MDE) and Domain-Specific Languages (DSLs). The nature of business document correction is very domain-specific where document editing programs such as MS Word, PowerPoint, and Excel are used. DSLs and MDE are alternative solution to overcome various problems faced while developing Cyber-Physical Systems (CPS) of a specific domain. They offer substantial gains in expressiveness and ease of use compared with general-purpose programming languages and development methodologies. As such, DSLs and MDE become an emerging popular area of research within the field of Software Engineering [33–35]. Microsoft Office is one of the representative documentation tools and it is in C# language in .NET Framework [36]. Therefore, CoPA has been developed in C# to couple with MS Office. The main reason for using C# is that the program source code in such as C# or VB written in Visual Studio, will be recompiled by CLR (Common Language Runtime) module in .NET, and then the program can be run independently without any limitations or restriction of the operating system as shown in Figure 2. This is possible because the exe file in the .NET program is not a complete machine language, but an IL (Intermediate Language), which allows the operating system to control applications such as MS Office PowerPoint, Excel, Word, Outlook, OneDrive, etc.

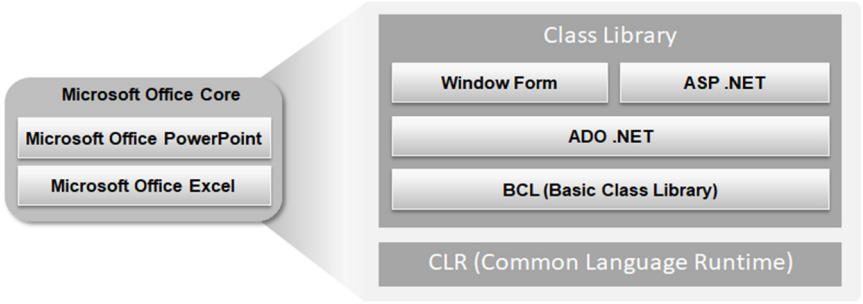

Figure 2. Components of the .NET framework.

Especially for PowerPoint automated correction, CoPA has been developed using BCL (Basic Class Library) and PowerPoint API guided in MSDN (Microsoft Developer Network). In addition, an application of Window Forms has also been developed to check the results of CoPA automatically besides the use of the console. The process of project proposal corrections CoPA is shown in Figure 3. The process of correction by CoPA is much simpler and easier since the designers just need to set the writing standards when compared to the correction process by manpower as shown in Figure 1. CoPA will perform the correction based on the writing standards. The main differences between manpower work shown in Figure 1 and CoPA process are that primarily all of the correction tasks including editing and design with CoPA are automated. The second difference is that the designer only needs to establish the writing standard. It can be expected that CoPA will reduce the cost by time consumption and improve the quality of documents. In addition, document automation with CoPA will improve the daily working experience.

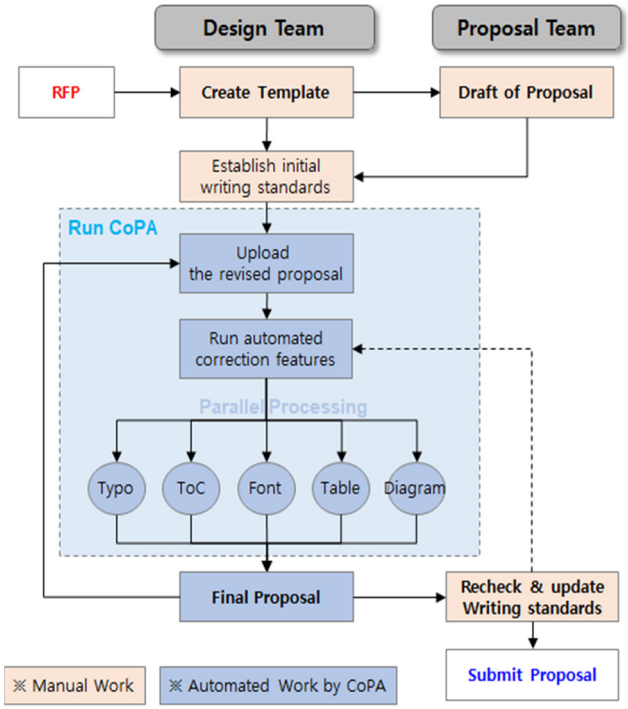

Figure 3. Process of project proposal correction by CoPA.

4.2. Agile Methods of CoPA Development

CoPA has been developed based on the agile methodology since agilely by proceeding each step in parallel is faster than the waterfall method developing procedurally. Agile methodology can be selected according to the characteristics of the ongoing project and the degree of participation and there are many cases including Scrum, XP (eXtreme Programming), FDD (Feature Driven Development), DSDM (Dynamic System Development), Lean

Software Development, Crystal Clear, and ASD (Adaptive Software Development) [11,37]. Among them, Scrum has been chosen to develop CoPA since it is the most effective and used method for the development and maintenance of high-quality software without the dependency of program language and/or project management method. For this Scrum method, Product Backlog has been based on Table 3 where five correction functionalities were treated as Sprint. Sprint has a repetitive development cycle of 30 days, working 8 h a day based on the Product Backlog. A list of required tasks and daily progress are drawn through Sprint Meeting.

Table 3. Research documents related to BDA as a comparison of this paper.

Reference No.	Category	Authors	Research Content
[21]	Spelling checking	Wood, F.; Tajeran, K	Real-time verification of spelling when multiple users work on business documents at the same time
[22]		Lee, J.H. et al.	Detection method for two spelling errors (word errors, context errors)
[23]	Document generation	Northoff, T.; Gresbrand, K	Methods and limitations of automatic document creation based on case studies of law firms
[24]	Applying document standards	Liegl, P.; Huemer, C.; Pichler, C	Document standard method in invoices and contracts based on XML
[25]		Wu, H.T. et al.	
[26]		Chae, S.Y.; Lee, C.Y.	How to analyze standardization compliance for Web (World Wide Web) documents
[27]	Document classification (Detection & Recognition)	Seraogi, B. et al.	Recognition method of text described in graphic drawing document
[28]		Ma, W. et al.	Text detection and recognition method in processing of digitized classic documents
[29]		Cho, D.H. et al.	How to automatically classify types of documents scanned through Neural Network
[30]		Lee, J.S. et al.	Automatic classification method of documents according to Korea Standard Industry Code
[31]		Yun, S.H. et al.	Automatic classification method for documents for which document classification is not defined
[32]		Nguyen, M.T. et al.	How to derive meaning and classify unlabeled documents
-	CoPA	Hyun Y.G. et al.	CoPA of the document automation method presented in this paper is superior in terms of performance and business applicability compared to existing studies

4.3. Development of CoPA Functionality

For CoPA applications, repetitive correction activities according to a certain pattern will be the target of automation and is applied after the draft version of proposal is written. RPA for correction activities consist of five correction functionalities: ① typo verification, ② table of contents verification, ③ font standard verification, ④ table layout verification, ⑤ schematic diagram layout verification. Five correction features are described in Table 3. Correction activities ①~② are mainly correcting the context violation against of writing standards while ③~⑤ automated corrections are for the design violation.

4.3.1. Typo Verification

It has been developed using MS-Word Object to check the typos collectively by comparing the words in the project proposal with the standard words registered in the MS-Office Dictionary. Typos are common errors by human mistakes especially when several members work at the same time. The major errors are from the words that have exactly the same meaning with the different pronunciation as shown in Figure 4. For instance, "Ministry Of the Interior and Safety" which is formally "행정안전부" in Korean, pronounced as "haengjeonganjeonbu". The word can be written in different ways as follows; "행안부" pronounced as "haenganbu", or "행정부" pronounced as "haengjeongbu", or called MOIS, or client. In this case, those have to be searched in all documents and revised to a guided single word to keep consistency if all working members wrote in different ways. This searching and revision works are developed as the first functionality of CoPA.

Figure 4. Examples of the word "Ministry of the Interior and Safety" written in different writing format in both English and Korean. Korean terms, "행안부", "행정안전부", "행정부", "행정 안전부" are all variations of the same word "Ministry of the Interior and Safety".

4.3.2. Table of Contents (ToC) Verification

The ToC of the project proposal is always located at the top of the document (Header area) and has a characteristic that it is written in a predetermined pattern (section numbers + section title, e.g., '1.2.2 Composition Plan'). First of all, the ToC correction logic starts extracting both section numbers and titles and then checks whether the section numbers are in the ordered list or not. Secondly, the correction logic checks whether each section number and title presented at the indicated pages are the same as shown in the table of contents. CoPA will perform ToC corrections if there are any wrong numbers or titles. In this paper, the numbering of ToC is set as follows "1 > 1.1 > 1.1.1 > A > 1) > A))" according to the writing standard. Figure 5 shows the example of ToC Correction.

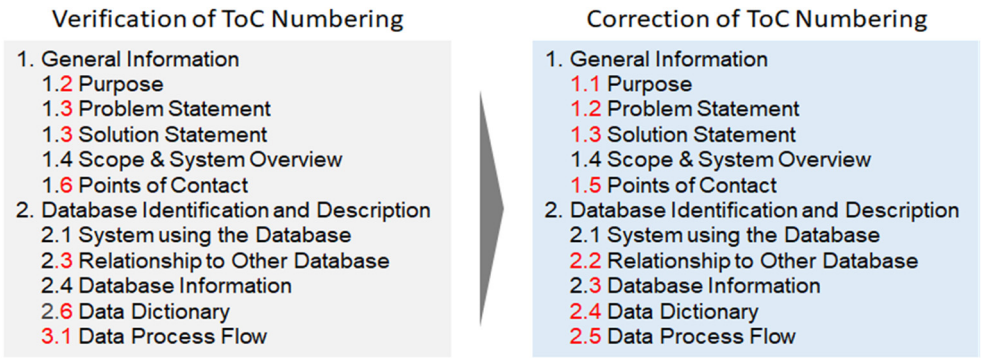

Figure 5. Example of ToC correction.

4.3.3. Font Verification

The font correction logic compares the standards and written front specification, which are font type, size, color, and other effects applied to text such as bold, italic, tilt.

The correction logic will automatically revise the text if the fonts are not complying the standards. For the writing standard, the font with 'clear gothic', 24 size, and RGB (28, 28, 28), without any special effects are considered. Figure 6 shows the example of font standard correction.

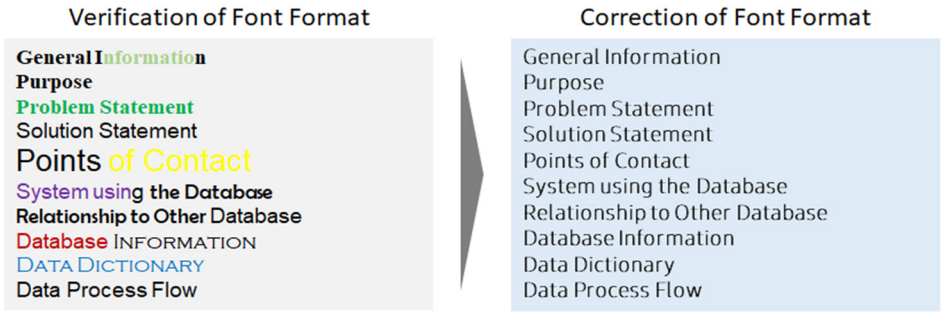

Figure 6. Example of text standard correction.

4.3.4. Table Verification

Table formats are expressed in various ways depending on the content, so it is practically impossible to apply one standard at a time. For example, the title may or may not exist on the above table and consist of 1~3 depth in the upper part of the table. In this paper, the correction logics for tables are as follows; the title format is set with 'Gulim' font, 12 font sizes, RGB (255, 255, 255) and 1 depth. Inside the table, the standards for each cell contents are set with 'Gulim' font, 10 font sizes, RGB (88, 88, 88), and located in the middle of the cell. For the table design, the standards consist of thickness = 0.5 pt, color = RGB (67, 1118, 79) and line type = solid. All table standards are shown in Figure 7.

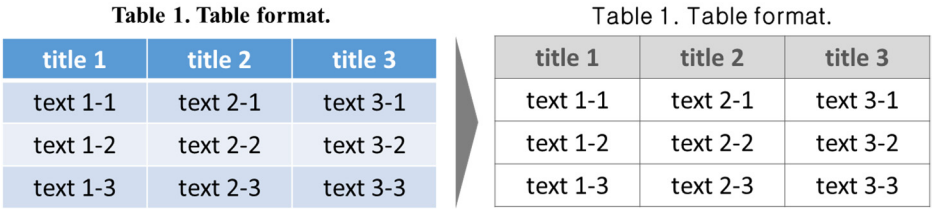

Figure 7. Example of table standard correction.

4.3.5. Schematic Diagram Verification

The correction logics are applied to the lines and arrows in all schematic diagrams. The schematic diagram correction standards are as follows; line type = solid, thickness = 1.5 pt, RGB (127, 127, 127), the starting point = circle, and the ending point = arrow-style arrow as shown in Figure 8.

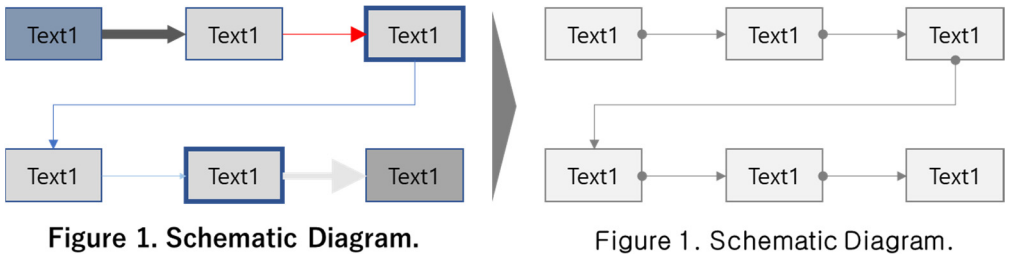

Figure 8. Example of schematic diagram correction.

5. Applications of Business Document Automation

In this section, test cases and environment are explained. CoPA's performance in terms of elapsed time and document quality are compared with the performance obtained by human correction.

5.1. Test Environment and Applications

CoPA is applied to correct five project proposals written in Korean using MS PowerPoint 2016. The project proposals consist of more than hundred pages (average 208 pages) as shown in Table 4. In these test cases, public project proposals are considered which are in a vertical format and less than 300 pages. For manpower, ten people (5 males, 5 females) aged between 20s to 40s manually correct these project proposals as shown in Table 5. All of the candidates (average 33 years old) get used to using the MS Office. The revision of project proposals has been proceeded by 2 people in 1 team (total five teams: one project proposal per one team) under scrum method.

Table 4. Automation area and plan.

Automated Correction Area	Content of Automated Correction
① Typo verification	• Automatic extraction of typographical error and non-registered words in the standard dictionary shown in MSDN • Maintain consistency of specific predefined words (synonyms) • Batch editing of the same typographical and non-registered words in the document
② Table of Contents verification	• Automatic check of the table of contents described on each page - Automatic extraction of the table of contents - Automatic checking of the numbering sequence and whether the table of contents name is exactly the same
③ Font verification	• Applying specified font standards - Automatically check whether the standard is applied (e.g., font type, color, size and other effects) - Automate modification according to the standard
④ Table verification	• Applying specified table standards - Automatically modify table lines (thickness, solid/dotted line, color) - Automate the editing of text and background color in each table cell
⑤ Schematic Diagram	• Applying specified schematic diagram standards - Automatic check whether the line standard is applied - Automate editing of line (thickness, line, color, head)

Table 5. Specifications of test project proposals.

Test Cases	Number of Pages	Number of Words	Number of Tables	Number of Diagrams
Case 1	252	195,547	189	339
Case 2	267	207,631	213	360
Case 3	272	212,484	216	377
Case 4	289	225,615	231	397
Case 5	299	233,831	239	414

The simulation using CoPA has been run on Samsung Galaxy Book Flex in a system environment; Intel® Core™ i7-1065G7 CPU, 16.0 GB LPDDR4x Memory, 12 GB NVMe SSD, under Windows 10.

5.2. Numerical Results of BDA Using CoPA

Five project proposals have been revised by five manpower teams and CoPA, and their performance in terms of time consumption and error count are summarized in Tables 6–12. Table 6 compares the average performance obtained by manpower and CoPA through Test cases 1–5. Error reduction column from test cases 1–5 as in Tables 7–12 describes the percentage of the errors reduced for work done from humans to CoPA.

Table 6. Specification of employees for revision.

Specification	C1 [1]	C2	C3	C4	C5	C6	C7	C8	C9	C10
Ages	29	35	33	26	45	38	37	35	27	49
MS Office	O	O	O	O	O	O	O	O	O	O

[1] C1 represents the first candidate.

Table 7. Test cases 1–5: Mean performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	55.4 min	23	2.0 min	6	**96.5%**	**76.2%**
② Table of Contents Verification	31.8 min	0	0.6 min	0	**98.0%**	**100.0%**
③ Font Verification	1517.5 min	108	1021.0 min	0	**32.7%**	**100.0%**
④ Table Verification	167.1 min	7	109.9 min	0	**34.3%**	**100.0%**
⑤ Schematic Diagram Verification	584.4 min	15	420.5 min	0	**28.0%**	**100.0%**
Total Cost (①–⑤)	**2356.1 min**	**153.8**	**1554.0 min**	**5.6**	**57.9%**	**95.2%**

Table 8. Test case 1: Performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	55.4 min	15	1.7 min	3	**97.0%**	**80.0%**
② Table of Contents Verification	33.4 min	0	0.7 min	0	**97.9%**	**100.0%**
③ Font Verification	1393.0 min	64	931.1 min	0	**33.2%**	**100.0%**
④ Table Verification	157.9 min	8	99.8 min	0	**36.8%**	**100.0%**
⑤ Schematic Diagram Verification	544.4 min	17	382.9 min	0	**29.7%**	**100.0%**
Total Cost (①–⑤)	**2184.1 min**	**104.0**	**1416.2 min**	**3.0**	**58.9%**	**96.0%**

Table 9. Test case 2: Performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	64.4 min	27	2.1 min	5	**96.8%**	**81.5%**
② Table of Contents Verification	34.2 min	0	0.8 min	0	**97.6%**	**100.0%**
③ Font Verification	1477.5 min	115	989.0 min	0	**33.1%**	**100.0%**
④ Table Verification	163.0 min	6	107.1 min	0	**34.3%**	**100.0%**
⑤ Schematic Diagram Verification	585.2 min	14	408.1 min	0	**30.3%**	**100.0%**
Total Cost (①–⑤)	**2324.2 min**	**162.5**	**1507.1 min**	**5.0**	**58.4%**	**96.3%**

Table 10. Test case 3: Performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	44.4 min	22	1.9 min	4	**95.7%**	**82.2%**
② Table of Contents Verification	26.7 min	0	0.8 min	0	**96.9%**	**100.0%**
③ Font Verification	1486.0 min	112	1004.4 min	0	**32.4%**	**100.0%**
④ Table Verification	164.0 min	6	108.4 min	0	**33.9%**	**100.0%**
⑤ Schematic Diagram Verification	557.7 min	14	419.5 min	0	**24.8%**	**100.0%**
Total Cost (①–⑤)	**2278.9 min**	**154.5**	**1535.1 min**	**4.0**	**56.7%**	**96.4%**

Table 11. Test case 4: Performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	51.8 min	28	2.0 min	7	**96.1%**	**75.0%**
② Table of Contents Verification	28.7 min	0	0.5 min	0	**98.3%**	**100.0%**
③ Font Verification	1583.5 min	122	1056.3 min	0	**33.3%**	**100.0%**
④ Table Verification	169.3 min	9	114.3 min	0	**32.5%**	**100.0%**
⑤ Schematic Diagram Verification	599.7 min	17	439.4 min	0	**26.7%**	**100.0%**
Total Cost (①–⑤)	**2433.1 min**	**176.5**	**1612.5 min**	**7.0**	**57.4%**	**95.0%**

Table 12. Test case 5: Performance comparison between manpower and CoPA.

Automation Area	Human		CoPA		CoPA Productivity Improvement	
	Elapsed Time	Errors	Elapsed Time	Errors	Time Efficiency	Error Reduction
① Typo Verification	60.8 min	25	2.1 min	9	**96.6%**	**63.9%**
② Table of Contents Verification	35.9 min	0	0.4 min	0	**98.9%**	**100.0%**
③ Font Verification	1647.5 min	126	1124.3 min	0	**31.8%**	**100.0%**
④ Table Verification	181.4 min	6	119.6 min	0	**34.1%**	**100.0%**
⑤ Schematic Diagram Verification	634.9 min	15	452.7 min	0	**28.7%**	**100.0%**
Total Cost (①–⑤)	**2560.5 min**	**171.4**	**1699.1 min**	**9.0**	**58.0%**	**92.8%**

In overall, CoPA reduces the total working time by more than 57.9% while maintaining 95.2% accuracy. It can be seen that CoPA improves the revision time efficiency by more than 96% for editing tasks such as ① and ② when compared to the manpower. For design tasks including ③, ④, and ⑤, the time consumption of CoPA took approximately 70%. It can be also noticed that CoPA generates zero-error for correction tasks ②–④. However, CoPA produces 5 or 9 errors for ① Typo Verification since the sub-dictionary of CoPA was not synchronized with the standard dictionary of MS office during the test time. This issue will be discussed in Section 5. In overall, it is clearly shown that the benefits of using CoPA for business document automation are not only reducing correction working time but also improving the quality of documents.

6. Discussion

Even though CoPA produced more than 98% lower errors while spending only half of working time when compared to the manpower, it is still great to discuss CoPA's limitations and how to improve its productivity. Through the simulations, it is found the important discussion points for five automated correction features including typo, ToC, font, table, and schematic diagram verifications. For each correction feature, it will be discussed its functional limitations and how it can be improved.

6.1. CoPA Logic Limitation and Future Improvements

For automated typo correction, CoPA was not able to produce 100% accurate corrections through five project proposals as shown in Tables 6–10. The main reason is that the standard dictionary registered in MS-Office does not include the newly emerged words such as BigData, BlockChain, IoT, Covid, etc. So CoPA tried to correct the new emerging words. To avoid this circumstance, it is necessary to register an emerging word data set to the standard dictionary in MS-Office before executing CoPA. This process will improve the accuracy of CoPA.

In the simulations, only one style of table of contents (1 > 1.1 > 1.1.1 > A > 1) > A)) was applied that cannot cover all demands from all users. So it would be better to consider various styles of table of contents to offer variety if CoPA is considered as a commercial solution in the near future.

The predefined font correction logic was perfectly applied to the body text of proposals with constant font style, size, color, and effect of the text (Tilt/Bold/Italic). In the near future, it will be effective if various font logics are considered for the governing messages in the header of each page.

The biggest problem in applying automated correction to tables is the titles consisting of a wide variety of forms as mentioned in Section 4. It was working well with a single standard. However, it should be manually corrected if the title consists of 3 depth components. Another manual process was applying bold and/or colored fonts for specific cells which is depending on the users. In the near future, the content with tags will be applied so CoPA can verify whether bold and/or font color and background color of the cell will be applied or not.

For the schematic diagram correction, it can be seen that CoPA successfully applied the standards described in Section 4. However, it can be found that there were some manual works for drawing dotted lines meaning off-line and cloud connection after the simulations. As a result, the use of one single standard is expected to cause another unnecessary work. For the next version of CoPA, it will only apply the dotted line when the line has the predefined line thickness.

Furthermore, it can be more efficient and user friendly if the errors are visible using MS Excel. For the next version of CoPA, it will apply TreeNode and TreeView under tree structure to provide an intuitive error treatment.

6.2. CoPA Physical Limitation and Future Improvements

It has been proven that CoPA can improve the business productivity in terms of the time consumption and document quality when compared to manpower. However, it is still required physical improvement of CoPA especially for correction features ③–⑤ that are approximately 30% faster when compared to manpower. The main reason for this is that the system of CoPA has been run in the laptop described in Section 5.1. For the future, CoPA will be located on the cloud including AWS, MS Azure and Google Cloud, to speed up the correction process that is one of the advantages of 5G technology. In other words, the users just need to upload files to the cloud and CoPA will perform the correction tasks online without installation process on the local machine. As a result, CoPA can be used as long as there is internet connection without any machine dependencies.

7. Conclusions

RPA named CoPA (Correction Process Automation) with five features including typographical error, table of contents, font style, tables, and schematic diagram verification, are demonstrated and implemented to solve the project proposal correction problems. Numerical results obtained by CoPA and employees are compared in terms of the business productivity in terms of computational efficiency and document quality. The paper clearly shows the benefits of using CoPA which produces error-free for table of contents, font, tables, and schematic diagram.

Current research focuses on the government public institution where uses formal reports using Hangul software similar to MS word. Working productivity will be improved significantly for public officials to inspect document quality. CoPA coupled with Hangul SDK is under investigation.

In near future, CoPA will be upgraded using VPL (Visual Programming Language). Therefore, a user who does not know programming will be able to easily create and implement a customized correction algorithm.

Author Contributions: Conceptualization, Y.H. and J.K.; Data curation, Y.H., D.L. and U.C.; Formal analysis, Y.H.; Funding acquisition, J.L.; Investigation, D.L.; Methodology, Y.H. and D.L.; Project administration, D.L., J.K. and U.C.; Resources, D.L. and U.C.; Software, Y.H. and J.K.; Supervision, J.L.; Validation, J.L.; Visualization, D.L. and J.K.; Writing—original draft, Y.H. and D.L.; Writing—review & editing, J.K. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kim, G.S.; Lee, G.Y.; Jo, M.J. *RPA Implement and Service Innovation (Focusing on Financial Industry Cases)*; Samjong KPMG: Seoul, Korea, 2017.
2. Schatsky, D.; Muraskin, C.; Iyengar, K. Robotic process automation. In *A Path to the Cognitive Enterprise*; Deloitte Consulting: New York, NY, USA, 2017.
3. Lee, K.Y. The Simple Repetition Task Is Entrusted to the Robot, but Focuses on Work Smart Rather Than Work Less. Available online: https://dbr.donga.com/article/view/1206/article_no/8868/ac/magazine (accessed on 22 September 2021).
4. Ahn, S.M. Chat-Bot. Accelerates Introduction of RPA such as Robot Advisor. *METRO*. Available online: https://www.metroseoul.co.kr/article/2018081200037 (accessed on 22 September 2021).
5. Lim, E.Y. Software Robots Simulating Human Tasks. Available online: https://jmagazine.joins.com/forbes/view/309997 (accessed on 22 September 2021).
6. AI Multiple. Available online: https://research.aimultiple.com/robotic-process-automation-use-cases/#banking (accessed on 22 September 2021).
7. Langmann, C.; Turi, D. *Robotic Process Automation (RPA)—Digitalisierung und Automatisierung Von Prozessen: Voraussetzungen, Funk-Tionsweise und Implementierung am Beispiel des Controllings und Rechnungswesens*; Springer: Berlin/Heidelberg, Germany, 2020. [CrossRef]
8. Jung, J.H. *The Fourth Industrial Revolution Winds at Office (the Era of Robots in Biz. Operation)*; POSCO Economic Research Institute: Seoul, Korea, 2017.
9. AI Multiple. Available online: https://research.aimultiple.com/robotic-process-automation-use-cases/ (accessed on 22 September 2021).
10. CiGen. Available online: https://www.cigen.com.au/cigenblog/4-ways-leverage-robotic-process-automation-legal-industry (accessed on 22 September 2021).
11. Kim, Y.D. Requirement Tracing Method Using Matrix in Agile Development Process. Ph.D. Thesis, Junnam University Graduate School, Gwangju, Korea, 2013.
12. Reinkemeyer, L. Process Mining, RPA, BPM, and DTO. In *Process Mining in Action*; Springer: Cham, Switzerland, 2020; pp. 41–44. [CrossRef]
13. Syed, R.; Moe, W. How to Trust a Bot: An RPA User Perspective. In Proceedings of the International Conference on Business Process Management, Seville, Spain, 13–18 September 2020; Springer: Cham, Switzerland, 2020; pp. 147–160. [CrossRef]

14. AI Multiple. Available online: https://research.aimultiple.com/robotic-process-automation-use-cases/#insurance (accessed on 22 September 2021).
15. Park, J.H. *White-Color Worker's Tool to Improve Productivity*; Institute for Information & Communication Technology Promotion: Seoul, Korea, 2017.
16. Le Clair, C.; Cullen, A.; King, M. *Robotic Process Automation, Q1 2017*; The Forrester Wave: Cambridge, UK, 2017.
17. Desai, P. Robotic process automation: RPA pre-requisite and pivotal points: Special issue: Special issue: IAISCT (SS4). In Proceedings of the 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 9–10 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 446–451. [CrossRef]
18. Chakraborti, T.; Isahagian, V.; Khalaf, R.; Khazaeni, Y.; Muthusamy, V.; Rizk, Y.; Unuvar, M. From Robotic Process Automation to Intelligent Process Automation. In Proceedings of the International Conference on Business Process Management, Seville, Spain, 13–18 September 2020; pp. 215–228. [CrossRef]
19. Agostinelli, S.; Marrella, A.; Mecella, M. Research challenges for intelligent robotic process automation. In Proceedings of the International Conference on Business Process Management, Seville, Spain, 26–28 June 2019; pp. 12–18. [CrossRef]
20. Cho, H.J.; Jeong, K. A Study on the Connective Validity of Technology Maturity and Industry for Core Technologies based on 4th Industrial Revolution. *J. Korea Converg. Soc.* **2019**, *10*, 49–57. [CrossRef]
21. Wood, F.; Tajeran, K. Error Detection and Correction for Multiple Document Editor. U.S. Patent Application No. 16/213,795, 2020.
22. Lee, J.H.; Kim, M.; Kwon, H.C. Deep Learning-Based Context-Sensitive Spelling Typing Error Correction. *IEEE Access* **2020**, *8*, 152565–152578. [CrossRef]
23. Northoff, T.; Gresbrand, K. Writing It up Right: Which Document Automation Tool Is Best for Me? In *Liquid Legal*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 393–410. [CrossRef]
24. Liegl, P.; Huemer, C.; Pichler, C. A bottom-up approach to build XML business document standards. In Proceedings of the 2010 IEEE 7th International Conference on E-Business Engineering, Shanghai, China, 10–12 November 2010; pp. 56–63. [CrossRef]
25. Wu, H.T.; Guo, L.H.; Wang, M.J.; Yang, J. Optimization Algorithms Study and Implementation on Graph Drawing Based on XML Document. *Procedia Comput. Sci.* **2019**, *154*, 33–39. [CrossRef]
26. Chae, S.Y.; Lee, C.Y. Analysis and Correction of Web Documents' Non-Compliance with Web Standards. *J. Korean Phys. Soc.* **2019**, *74*, 731–743. [CrossRef]
27. Seraogi, B.; Das, S.; Banerjee, P.; Majumdar, H.; Mukkamala, S.; Roy, R.; Chaudhuri, B.B. Automatic Orientation Correction of AEC Drawing Documents. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017; Volume 2, pp. 9–10. [CrossRef]
28. Ma, W.; Zhang, H.; Jin, L.; Wu, S.; Wang, J.; Wang, Y. Joint Layout Analysis, Character Detection and Recognition for Historical Document Digitization. In Proceedings of the 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund, Germany, 8–10 September 2020; pp. 31–36. [CrossRef]
29. Cho, D.H.; Kim, J.H.; Park, J.M.; Park, T.H.; Park, B.J.; Kim, Y.H. Automatic classification of scanned business documents by page/type. *Korean Inst. Inf. Sci. Eng.* **2020**, *41*, 1655–1657.
30. Lee, J.S.; Jun, S.P.; Yoo, H.S. A Study on Automatic Classification Model of Documents Based on Korean Standard Industrial Classification. *Korea Intell. Inf. Syst. Soc.* **2018**, *24*, 221–241.
31. Yun, S.H.; Lee, S.H.; Son, C.S.; Kang, W.S. A techniques of document classification for documents without well-defined classification scheme. *Korea Inst. Commun. Sci.* **2018**, *41*, 1154–1155.
32. Nguyen, M.T.; Le, D.T.; Le, L. Transformers-based information extraction with limited data for domain-specific business documents. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104100. [CrossRef]
33. Mohamed, M.A.; Challenger, M.; Kardas, G. Applications of model-driven engineering in cyber-physical systems: A systematic mapping study. *J. Comput. Lang.* **2020**, *59*. [CrossRef]
34. Kosar, T.; Bohra, S.; Mernik, M. Domain-Specific Languages: A Systematic Mapping Study. *Inf. Softw. Technol.* **2016**, *71*, 77–91. [CrossRef]
35. Shen, L.; Chen, X.; Liu, R.; Wang, H.; Ji, G. Domain-Specific Language Techniques for Visual Computing: A Comprehensive Study. *Arch. Comput. Methods Eng.* **2020**. [CrossRef]
36. Microsoft MSDN. Available online: https://docs.microsoft.com/en-us/dotnet/standard/components (accessed on 22 September 2021).
37. Kim, H.M. A Case Study of Android Applications Development with Agile Methodology. Ph.D. Thesis, Soongsil University, Seoul, Korea, 2013.