An Analysis of Platforms for Scholarly Publication

Student Name: Yun Li
Student ID: 41017919
Supervisor: Robert Dale
Department of Computing
Macquarie University
Sydney, Australia
yun.li1@students.mq.edu.au

Table of Contents

Abstract

In the journal publishing world, it involves various text processing tasks: authors create manuscripts, copy-editors edit and correct them, and typesetters format the results for presentation either online or in hard copy. Modern text processing technologies, and in particular the kinds of techniques developed in natural language processing and language technology, make it possible to automate some of these editing and revision tasks that are currently performed by authors and editors, such as copy-editors add comments to manuscripts or typesetters change the style of headings. However, it is difficult for these technologies to integrate well with the variety of different text processing platforms and wolkflows of publishing process. The aim of this paper is to help us understand the range of ways in which authors and publishers are currently work and tools that they use, and then to identify some possible automation which could be integrated into these tools and to indicate places where these integrations might be introduced in the different publication workflows. We summarise a general editorial workflow from a web observation. We also conduct a survey to gather information of production process in which relevant data is hard to find. Base on the 100 responses from the survey, we outline 3 most common production workflows. For each workflow, as well as the editorial workflow, we finally indicate the target audience and some possible automated editing assistance which might be introduced to them.

1. Introduction

1.1 What is a Scholarly Journal?

Currently, most academic work is published in journal article, book or research paper form. A scholarly journal is a periodical publication, ranging in frequency from monthly to quarterly or sometimes only twice a year. It is published in every academic discipline and contains articles written by and for scholars in a chosen discipline. These articles usually focus on a specific topic or a particular research interest. Experts in the same academic discipline review the articles to decide whether or not they are sufficient to be published in an academic manner. Scholarly journals are essential resources for researchers in terms of doing academic research. A researcher

will review the available literature on a topic to get a better understanding of the current state of research in a particular academic field. The process of the literature review helps the research to discover gaps in knowledge and to raise new questions that could be discussed within the field.

In order to distinguish scholarly journals from other periodicals, we compare some characteristics between scholarly journals and popular magazines like *Time* or *Fortune*. Table 1 outlines some common differences:

	Scholarly Journals	**Popular Magazines**
Audience:	Written for experts in a particular academic discipline	Written for the general public
Author:	Scholars who are experts on their topics	May be written by authors who have no expertise on the topic
Editorial process:	Have a peer review process	May not have a peer review process
Content:	Articles follow a scientific method; always include bibliographies of works cited; use technical language or jargon	Articles in a normal format; includes no or very brief bibliographies; avoid specialized terminology
Appearance:	Articles are lengthy with in-depth analysis; usually contain no advertising; very few photographs or diagrams	Articles are often short; contain more advertising; very visual, often use exciting pictures
Format:	Besides publication in paper format, many articles are also available on the Web	Most are in paper format

Table 1 the Difference between Scholarly journal & Popular Magazines

In short, scholarly journals have a serious, reasonable look. It assumes some scholarly background on the part of readers. Many scholarly journals are 'peer-reviewed' to ensure their articles are sufficient to be published in a specific academic discipline. Such kinds of "quality control" are applied in the process of scholarly journal publishing.

1.2 How is a Scholarly Journal Cooked?

Scholarly journal publishing is usually described as the distribution of academic research and scholarship. It can be summarised in a liner flow as showed in figure 1:

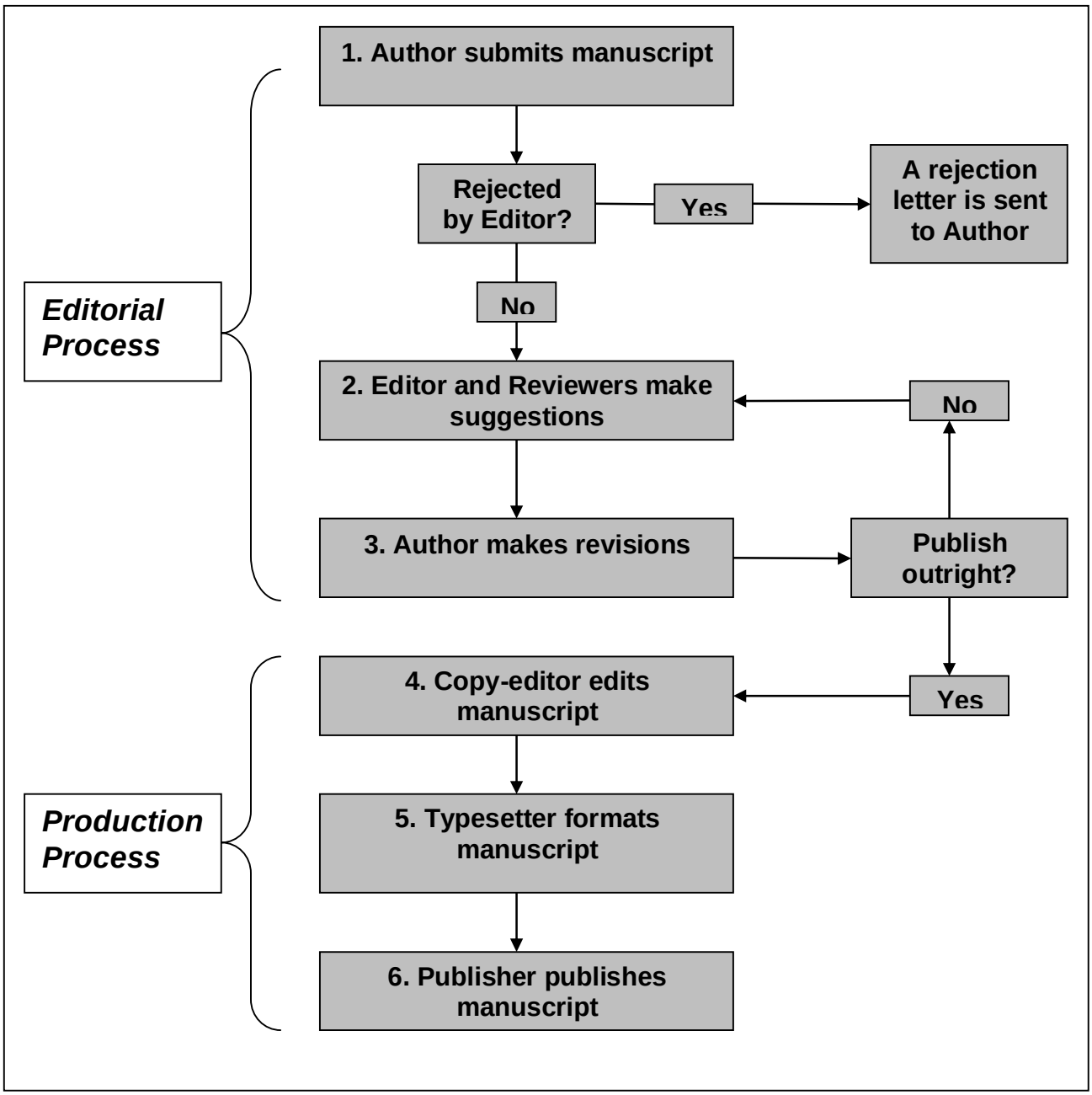

Figure 1 the Overview of the Scholarly Publishing Process

In the liner flow, steps 1-3 belong to the editorial process, while steps 4-6 belong to the production process. The details of these processes will be discussed in section 2; we firstly introduce the difference between them. The information of the editorial process regarding submitting articles and peer-review can be acquired from the journals themselves or the journals' web sites. The editorial process is usually controlled by the journal editor-in-chief who is responsible for assigning reviewers to each article and for passing articles among different parties and by the author who is responsible for revising the article. But once the article is accepted and passed into the production process, the information about this process becomes ambiguous since most journals do not publish the detailed production process in the journals nor on the web sites; it is difficult for us to understand the process as clearly as the editorial process.

1.3 What Challenges Exist in Scholarly Publishing?

Generally the whole scholarly publishing process involves various text processing tasks such as peer-reviewing, copy-editing, typesetting or proofing. Although some of these tasks have been streamlined by some modern technologies, such as manuscript management systems for managing peer-reviewing workflow, a number of editing and revision tasks are still labor-intensive as they are performed by authors, editors and publishers.

In 2003, Mabe (2003, pages191-197) indicated that there were approximately 1.4 million articles published by about 21000 journals. As the number of researchers is increasing, Steele (2006, page 42) showed the amount of journals has accrued to 23000. On behalf of these journals, it will be genuinely useful if some modern text processing technologies could be applied to automate some of the labor-intensive tasks in journal publication; Dale (1997, pages 235-237) also indicates that speech and natural language processing technologies are likely to be an Information Technology solution which could assist in these tasks. As noted by Dale (1997, pages 235-237), currently some language technologies have already been applied for text revision: for example language sensitivity has been utilized in developing spelling correction technology and grammar checking technology. As machines become more powerful, we expect to see more tasks to be automated. However, the variety of different text processing platforms used by authors, editors and publishers and the different workflows of publishing process make it difficult to determine what kinds of automation would be likely to be most successful. So the purpose of this paper is to carry out an

analysis of the workflows of scholarly journal publishing process and the tools that are being used in these workflows, then to identify places where automated editorial assistance might be introduced and recommend some kinds of possible automation, for instance a plug-in grammar checker for Microsoft Word.

The approach we use for information gathering is a survey of journal publishing process. We select 400 journal editors[1] from 360 A+, A and B ranked computer science journals on the CORE journal list[2]; most of them are editors-in-chief and production editors. The survey does not consider the journal editorial process as the information of this process can be acquired by the web observation. Instead, the survey has been undertaken to better understand the production process in which the information is hard to find.

Section 2 outlines some general journal publishing processes in which authors and editors currently work, and the tools and text processing platforms that are normally used. In section 3, we introduce the survey of journal publishing process along with its results. Then an analysis of these results is carried out in section 4. In section 5, we finally suggest some possible integrations of automated editorial assistance that might be introduced in scholarly journal publishing process.

[1] Some journals have co-editors-in-chief, which means more than one editor-in-chief.
[2] http://www.core.edu.au/

2. General Publishing Process

In the recent years, scholarly publishing is undergoing major changes as it is transferring from paper-based to electronic format. The paper-based publishing has persisted for a long time as it is cheap for authors to produce submissions (they do not need expansive hardware); for editors and reviewers, there is no special technology required to view submissions. During these years, since the increasing costs in handling paper submissions and the increasing use of Internet, paper-based publishing starts to go electronic. However, the main processes involved in traditional scholarly publishing and electronic scholarly publishing are almost the same. As mentioned by Campbell *et al.* (1997, page 24), the key processes can be described as editing, production, marketing and distribution:

- Editing: it generally involves peer-reviewing and revision.
- Production: it includes revision, copy-editing, typesetting, and proofreading.
- Marketing and distribution: they are concerned with selling the journal in the marker.

In this paper, we only focus on the two earliest stages: editing and production, since they involve most of the text processing procedures.

2.1 Journal Editorial Process

Editorial process, which is mainly described as the process of peer-review in scholarly journal publishing, covers the following common stages as indicated by Ciesielski (2005, pages 2-4):

- Editors collect submissions from authors
- Editors and reviewers carry out peer-review
- Editors make decisions (reject, accept, correct)
- Authors correct and resubmit (if asked to correct)

There are three main parties participate in the editorial process: authors, editors and peer-reviewers:

- Authors: also known as scholars, create the work that is published.
- Editors and Peer-reviewers: They provide "quality control" of the work that is created by authors, including screening submissions,

reviewing manuscripts, suggesting revisions, corresponding with authors and overseeing the final copy.

The process of peer-review is usually organized by the journal editor, then conducted by editors and peer-reviewers, and completed when the article is accepted for publishing. In order to better understand the editorial process, we have performed a web observation with journals on the CORE list. Approximately 100 A+ ranked journals' websites, which contain direct information about journal editing process, are selected for observation. The summary of results is quite similar to the workflow in figure 2: first of all, an author submits an article to a journal editor, then the editor checks whether the article is of sufficient quality to go through the peer-review. If the article is appropriate to the journal scope, then the editor assigns 3-5 reviewers who have the similar academic background as the author to each article for reviewing. Each review who agrees to evaluate the article needs to produce a report. The report may contain a set of assessment guidelines proposed by the journals.

After reviews are finished, reviewers send their reports on the article along with their comments back to the editor. The reviewers' comments generally fall into four categories: publish outright (though some further changes may be called for), revise and resubmit (usually with minor or major changes), and reject outright (without any revision). Very few articles are accepted outright, most fall into the middle group.

Once the author receives a comment like "revise and resubmit", the article will be changed base on the suggestions from the reviewers. If the revised and resubmitted article is approved by the editor, then it will be passed into the production process for publishing; otherwise it will be reviewed again until it meets the journal's sufficient academic standard for publishing.

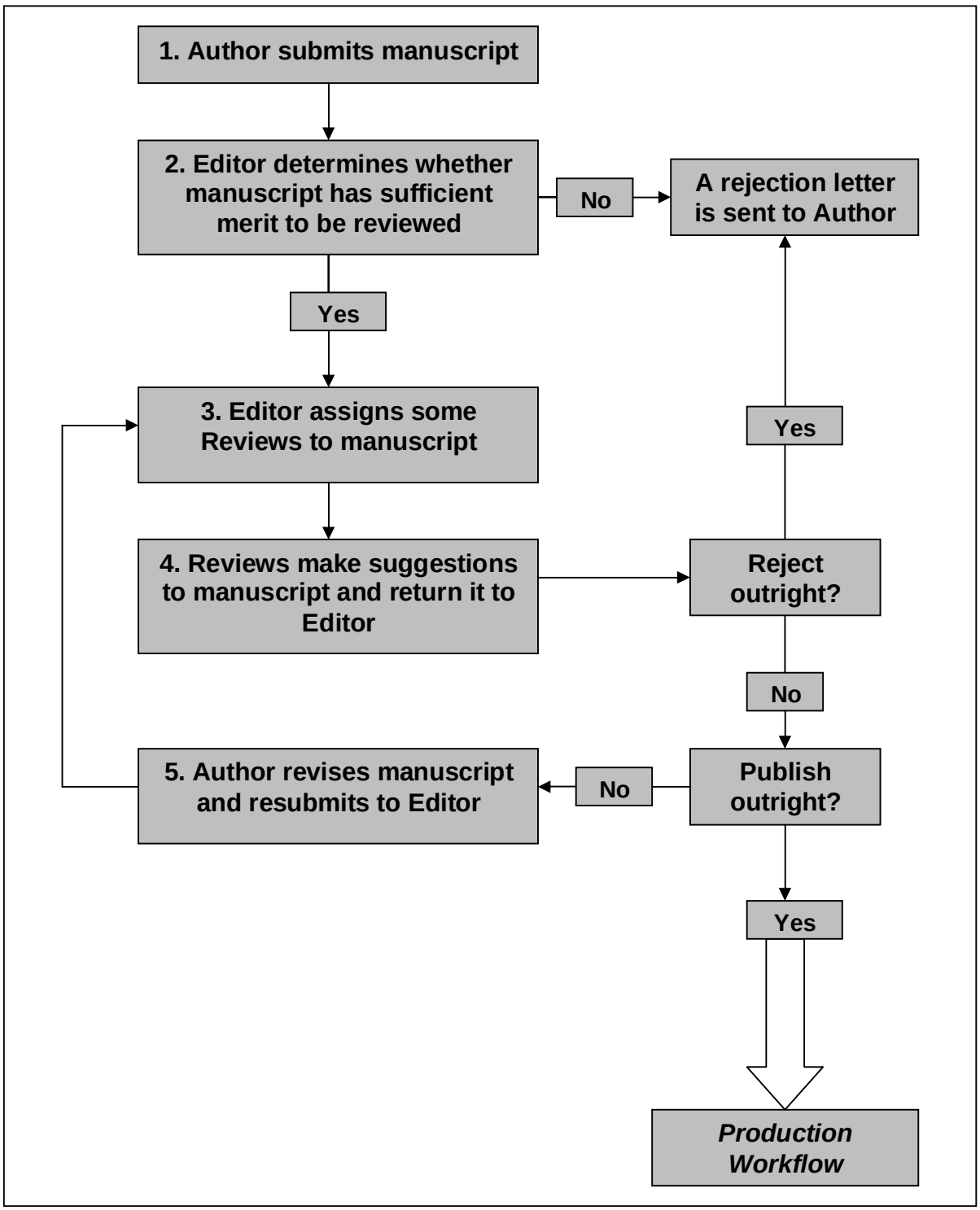

Figure 2 General Editorial Workflow

The electronic peer-review process can be automated through the use of online manuscript management systems such as ScholarOne ManuscriptCentral[3], Aries Editorial Manager[4], and EJournalPress[5]. These systems have been used to manage the submission and peer-review process in an effective way: authors can submit their manuscripts to the system; editors can review the submissions and assign reviewers directly in the system. The manuscript management systems assist editors to manage peer-review rapidly and comfortably; they automate or eliminate many manual steps in editorial process, such as distributing the article to the reviewers or submitting review reports to the editor.

2.2 Journal Production Process

Once the submissions are finally accepted by the editors after the editorial process, the production process happens and it is controlled by a production editor or publisher. The production process covers the following common stages: copy-editing, typesetting, inclusion in a specific issue of a journal, and then printing and publication.

Copy-editing is performed by copy-editors who make revision to the article and raise any questions or comments regarding content, sources and features. There are three levels of copy-editing: light, medium and heavy:

- Light Copy-editing: Copy-editors check and correct spelling, grammar, and punctuation errors, correct incorrect word usage, and correct cross-references. They also ensure the consistency in spelling, hyphenation, numerals, fonts, and capitalization. For the manuscripts fall into the light copy-editing level, they just require less editorial involvement than others.
- Medium Copy-editing: In the medium copy-editing, copyeditors perform all tasks in light copy-editing with additional checks for style consistency and relationship between text and graphics, as well as table-of-contents entries and organizational problems.
- Heavy Copy-editing: The main difference between medium and heavy copy-editing is the level of judgment and rewriting involved. In heavy copy-editing, besides all the tasks in medium copy-editing, copy-

[3] http://scholarone.com/products/manuscript
[4] http://www.editorialmanager.com/homepage/home.htm
[5] http://www.ejpress.com/index.shtml

editors check more serious errors in syntax, rewrite portions to enforce a uniform level, and rearrange sentences to improve readability.

Typesetting is about the presentation of an article: typesetters deal with layouts, fonts, headings etc to make sure the article is formatted in a manner that is consistent with the style of the journal. Proof-reading is involved for the author to review and correct proofs in one or more stages of the production process. As Dale (1990, pages 59-67) indicated, because both copy-editing and proof-reading are time-consuming and error-prone processes, automated assistance would be most likely to happen there. However, information regarding the details of the production process is inadequate and hard to find through the web observation. We need to understand the production process well enough before deploying any automated editorial assistance.

2.3 Existing Editing Assistance Tools

Almost each one of the publishing processes we mention above involves the using of text processing tools: authors may write their articles using Microsoft Word or LaTeX base on the requirements of the journal they submit to; copy-editors may enter corrections on the submissions using Microsoft Word; typesetters may format the submissions using Plain TeX, LaTeX or other tools. Microsoft Word and LaTeX both are standard word processing software. Microsoft Word is a "What You See Is What You Get" word processor. This means any content you can see in the editor is the same when it gets printed out. The user must concern with not only the content to write, but also the layout of them. It is difficult to get professional layout in Word.

LaTeX, on the other hand, is much more powerful in producing professional layout. The user just needs to type in some syntax and content then let LaTeX to generate a formatted document (usually in PDF format). With LaTeX, the user only needs to focus on the content without boring the layout. However, for those who have never used LaTeX before, they may need to invest some time to learn the syntax before they can actually do some tasks. But with Word they may also need a similar amount of time at the end to get the formatting right or move figures. So, LaTeX on large documents will be slower at the start and get faster at the end, and probably will overtake Word since formatting is really a time-consuming work.

As the variety of different platforms exists and moreover, the information about the production process is hard to find through the web observation, it is difficult for us to determine what kinds of possible automation which could be integrated into these tools would be most beneficial.

In order to better understand the production process and existing editing tools which have been utilized in scholarly publishing, we conducted a survey of journal publishing process which has been distributed to around 360 computer science scholarly journals and for each journal, we select one or two editors, including editor-in-chief and production editor. In the following section, we discuss the survey and its results in detail.

3. Survey of Journal Publishing Process

3.1 Overview of the Survey

The survey of journal publishing process acquired information from journal editors in the field of computer science (the reason for choosing this discipline is because the author currently is a master of Information Technology student studying in the computer science background). A 10-question web survey was e-mailed to approximately 400 editors from 30th Mar 2009 to 8th May 2009. These editors, including editors-in-chief and production editors, are selected from 360 A+, A and B ranked computer science journals on the CORE journal list. Of these 400 editors contacted, 100 responses (or a response rate of 25.0%) are aggregated for the analysis of this study.

The object of this survey is to help us understand the journal publishing process, especially the production stage. The 10 questions in the survey are designed to find out possible text processing platform(s) and workflow(s) that might exist in the production process. Below are the questions asked, corresponding results and verbatim comments from all who responded by 29th may, 2009 deadline.

3.2 Summary of the Survey Results

The first question is a very general question which asks each participant to provide the name of his/her journal. 97 of the participants provided the information while only 3 of them skipped the question. These results are useful when we analyze on a journal by journal basis in section 4.

Question 2 and 3 acquire information about how accepted manuscript is processed. Question 2 asks about the role that is responsible for copy-editing in the production process. The results from this question are very important for us to find out our target audience of possible automation in copy-editing. As showed in table 2, approximately 47% of the journals have in-house copy-editors who are responsible for the copy-editing work of accepted manuscripts. From the responses in the "Others" option, we find that in about 9% of the journals, the editors do the copy-editing.

Similar as question 2, the third one aims to find out the role that incorporates copy-editor's revisions into manuscripts for return to authors for checking. Table 3 indicates that 39% of copy-editor's revisions are done by copy-editors themselves, and around 20% are done by typesetters (both in-house and external). Only A small percentage (<8%) of revisions is incorporated by authors. 4 participants in the others say that minor changes are made by the editors, while major changes are made by the authors.

Q2: How is copy-editing of accepted manuscripts carried out for your journal?		
	Response Percent	**Response Count**
Copy-editors on the fulltime staff	**46.3%**	**38**
External freelance copy-editors	15.9%	13
No copy-editing is carried out	11.0%	9
Others: 1. Copy-editing is done by editors (8 responses) 2. Copy-editing is outsourced (3 responses) 3. All changes and revisions are made by authors (2 responses) 4. Copy-editing is undertaken by volunteer academics/copy-editors (2 responses) 5. Done by production manager/publisher (3 responses) 6. No clue (4 responses)	26.8%	22
answered question		**82**
skipped question		**18**

Table 2 Results of Question 2

Q3: How are the copy-editor's revisions incorporated into the manuscript for return to the author for checking?		
	Response Percent	**Response Count**
Copy-editors	**39.0%**	**30**
In-house typesetters	11.7%	9

External typesetters	9.1%	7
Author	7.8%	6
Others: 1. No copy-editing(4 responses) 2. Minor changes are made by the editor; major changes are return to the author(s) (4 responses) 3. Changes are incorporated by typesetters, but a marked proof is returned to the corresponding author, who can make further changes, etc. (2 responses) 4. Editorial revisions by editors or copy-editors are not communicated to the author. (3 responses) 5. Don't know (12 responses)	32.5%	25
answered question		**77**
skipped question		**23**

Table 3 Results of Question 3

Q4: In which format(s) do you provide the accepted versions of manuscripts to your copy-editors? (Answer all that apply.)		
	Response Percent	**Response Count**
Hard copy	6.8%	5
Electronic-word	**63.0%**	**46**
Electronic-PDF	41.1%	30
Others: 1. LaTeX (12 responses) 2. Depends on the submittal (4 responses) 3. Don't know (6 responses) 4. PDF with source files (4 responses)	35.6%	26
answered question		**73**
skipped question		**27**

Table 4 Results of Question 4

Q5: In which format do copy-editors indicate the changes to be made to articles? (Answer all that apply.)		
	Response Percent	**Response Count**
Hard copy	5.5%	4
PDF with added embedded comments	**39.7%**	**29**
Word with comments and changes	26.0%	19
Others: 1. In a variety of electronic formats, but authors just get PDF (5 responses) 2. E-mail (3 responses) 3. the copy editor takes only care of the formatting and this is done directly on the digital copy - no check by the author necessary - a last check it is done by the editor-in-chief (2 response) 4. LaTeX (1 response) 5. Don't know (10 responses)	28.8%	21
answered question		**73**
skipped question		**27**

Table 5 Results of Question 5

Question 4 and 5 request information regarding the formats in which the documents are passed between authors and copy-editors. Copy-editors, authors or typesetters (depending on who makes copy-editing changes) need to be using a document format that they can make changes to. Question 4 can tell us what exact platforms are being used. In table 4, we can see that the majority of these journals are using Microsoft Word format for accepted version of manuscripts to copy-editors, and approximately 41% of them turn the document into a PDF format then provide it to copy-editors (if copy-editors make changes). We can also find that 12 participants from the others are providing LaTeX format to copy-editors (if copy-editors make changes). LaTeX files are often used to produce PDF files that are sent to copy-editors.

Table 5 shows about 40% of copy-editors are using PDF with added embedded comments to indicate changes to be made to articles, while 26%

of them are using Word. This question along with question 4 provides us the information in terms of platforms: they can indicate the most possible places in which automations might happen in copy-editing process.

Q6: Who makes the copy-editing changes to the manuscript? (Answer all that apply.)		
	Response Percent	**Response Count**
Author	40.9%	27
Copy-editor	**53.0%**	**35**
Typesetter	21.2%	14
Others: 1. The editor (7 responses) 2. Don't know (8 responses)	22.7%	15
answered question		**66**
skipped question		**34**

Table 6 Results of Question 6

Q7: In which format(s) do you provide the final revised versions of manuscripts to your typesetters? (Answer all that apply.)		
	Response Percent	**Response Count**
Hard copy	4.6%	3
Electronic-word	**60.0%**	**39**
Electronic-PDF	38.5%	25
Others: 1. LaTeX (7 responses) 2. HTML (3 responses) 3. Any source files (3 responses) 4. Don't know (11 responses)	36.9%	24
answered question		**65**
skipped question		**35**

Table 7 Results of Question 7

Q8: What text processing platform does typesetter use to create the final published article? (Answer all that apply.)		
	Response Percent	**Response Count**
Microsoft Word	21.5%	14
LaTeX	24.6%	**16**
Don't know	**53.8%**	**35**
Others: 1. DeGruyter's in-house system (1 response) 2. Home-made system (1 response) 3. We use the final Word version and turn it into a PDF (4 responses) 4. InDesign (1 response) 5. Almost always using LaTeX (2 responses) 6. Not applicable - the author produces an XHTML file, which is the final format for the Website (1 response) 7. They handle the majority/all of the commonly used programs (1 response) 8. Don't know (4 responses)	23.1%	15
answered question		**65**
skipped question		**35**

Table 8 Results of Question 8

Question 6-8 are concerned with how copy-editors' and author' changes are provide to typesetter. From table 6, we can see that in over half of the journals, copy-editors make the change directly to manuscripts. These results tell us copy-editors are most likely to be the target audience to which we could introduce automations.

Table 7 indicates the format of final revised versions of manuscripts that is provided to typesetters after copy-editing is done. The majority of them are in Microsoft Word, while a smaller percentage of them are in PDF (38.5%) and a much smaller percentage in LaTeX (9%). The results of this question correlate with the results of question 8 provide us information that helps us to determine possible places where automations could happen in typesetting

process. For example, if most of the typesetting jobs are performed with Microsoft Word, then probably we may introduce a plug-in formatting tool to work with Word and automate the typesetters to format the manuscripts.

Table 8 summarizes the text processing platforms which are being used by typesetters. Almost 54% of the participants do not know the details. There might be some different platforms being used out there since some participants in the others indicate that they are using some commercial products or home-made systems rather than Microsoft Word or LaTeX.

Question 9 asks participants to provide comments on any other relevant aspects of their production process that are not adequately covered in the previous questions. 22 participants give the comments: 11 of them point out that some editors-in-chief are not directly involved with the production process. Since in the first half period of the survey, we only considered the editor-in-chief of each journal. After notice the information from this question, we change the survey targets to not only the editors-in-chief, but also the production editors from the relevant publishers.

The last question asks the participants to leave their contact addresses only if they are interested in the results.

Base on all these survey results, now we can get a better understanding about the progress of production process. In the next section, we outline some possible production workflows in scholarly journal publishing.

4. Outcome of the Survey

In this section, a further analysis on a journal by journal basis is presented. We use Microsoft Excel to help us automate the analysis: the 100 responses, along with the journal's name, are firstly placed in a spreadsheet, then we use the sort function provided by Excel to come up with some general groups of similar data.

As we have mentioned in section 2.2, automation would be most likely to happen in copy-editing process (Dale, 1990). Since the survey results in table 2 show that 38 journals (46.3%) use in-house copy-editors to carry out copy-editing, we especially focus on these journals in terms of automation because it would be more productive for them. In the following paragraph, we outline 3 most common workflows of the production process.

1. Workflow 1 (11 journals out of 38):

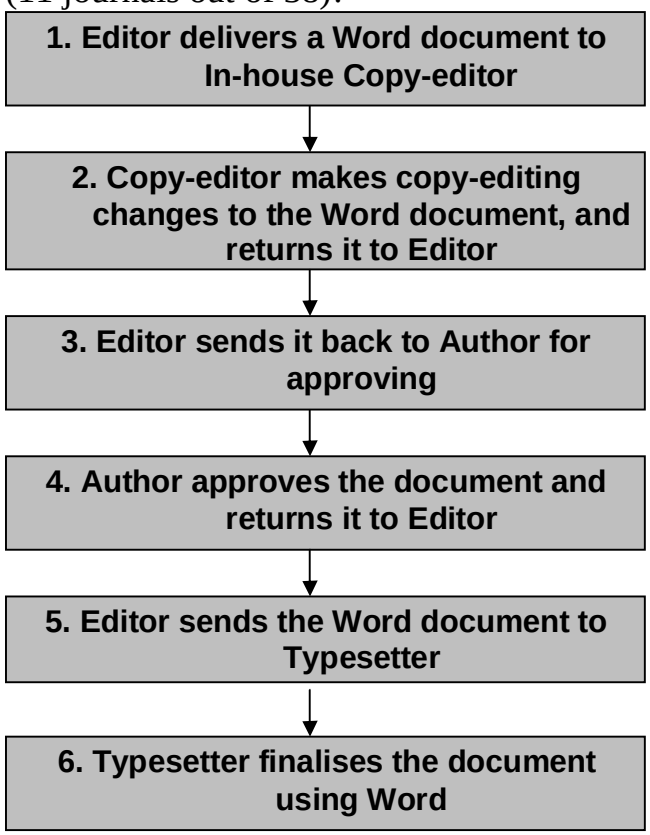

Figure 3 Production Workflow One

As showed in figure 3, in the first workflow, the copy-editor does copy-editing and corrects the manuscript using Microsoft Word platform which is the same one used by the typesetter to perform typesetting.

2. Workflow 2 (8 journals out of 38):

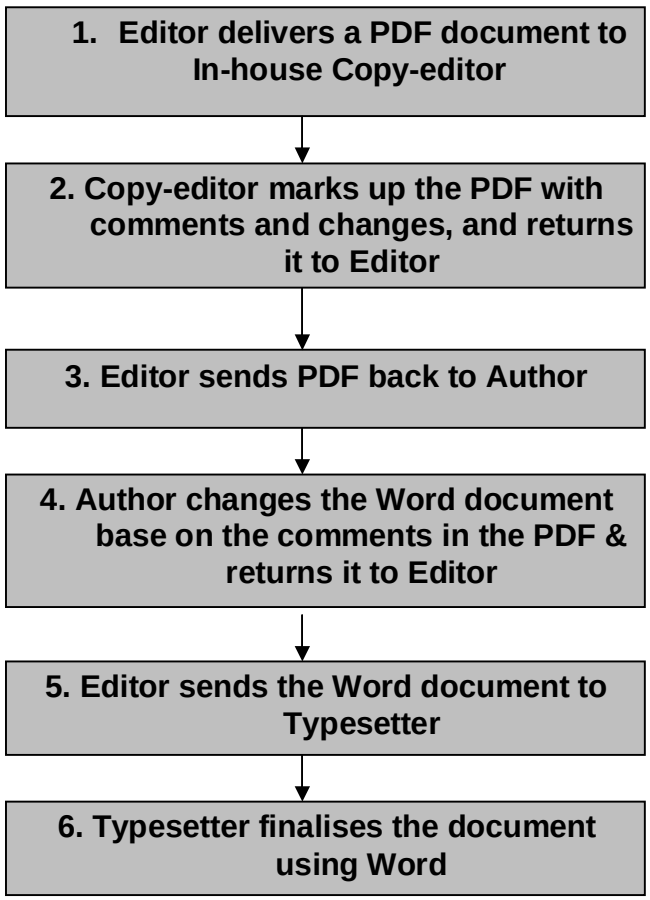

Figure 4 Production Workflow Two

In figure 4, the copy-editor only indicates copy-editing comments on the PDF format manuscript; it is the author who changes the manuscript base on these comments. The typesetter uses the same platform as the first workflow to perform typesetting.

3. Workflow 3 (7 journals out of 38):

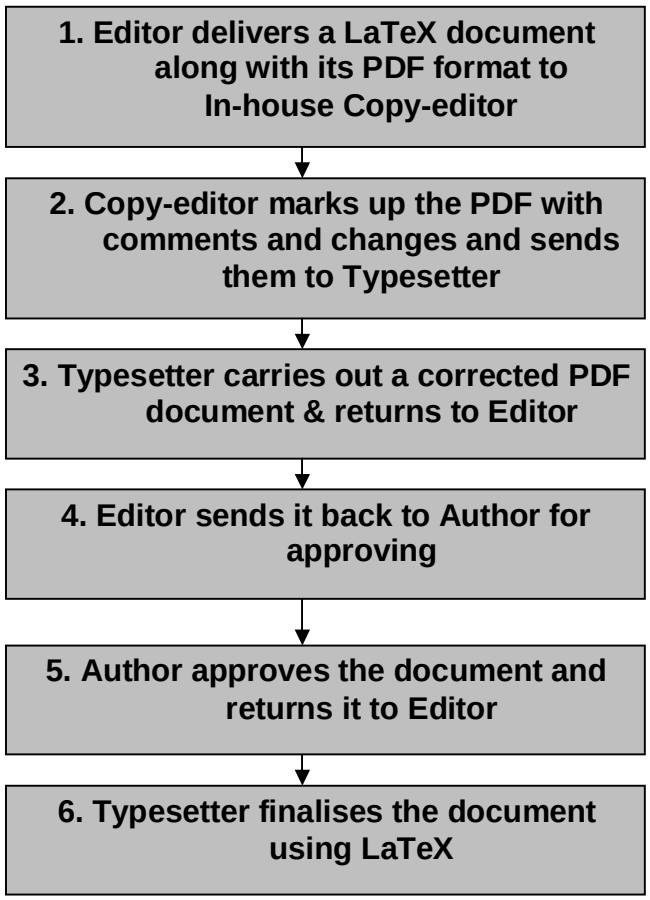

Figure 5 Production Workflow Three

In the third workflow in figure 5, it is the typesetter who indicates the copy-editing comments provided by the copy-editor, and then uses LaTeX to perform typesetting work.

After we have gathered adequate information of the production process, it is possible for us to identify places in which automation might be introduced. In the next section, we suggest some kinds of automation which could benefit the publishing process.

5. Automation in the Publishing Process

5.1 Possible Automation in the Editorial Process

In the editorial process, because the process of peer-review is the most frequent one, the target audience to be introduced with automation might be reviewers and authors. Reviewers might need an automation tool which could be fitted in to Word or LaTeX to help them mark up comments on the manuscripts rapidly and comfortably. Authors might need an automation tool which could help them easily correct their manuscripts.

5.2 Possible Automation in the Production Process

In the production process, for the first workflow, the target audience to be introduced with automation might be copy-editors and typesetters (see figure 6). Copy-editors might need a plug-in grammar/styling checker with Microsoft Word to help them easily mark up and correct grammar errors. Typesetters might need a plug-in formatting tool which could be fitted into Word as well but automates them with some typesetting tasks like changing the style of headings.

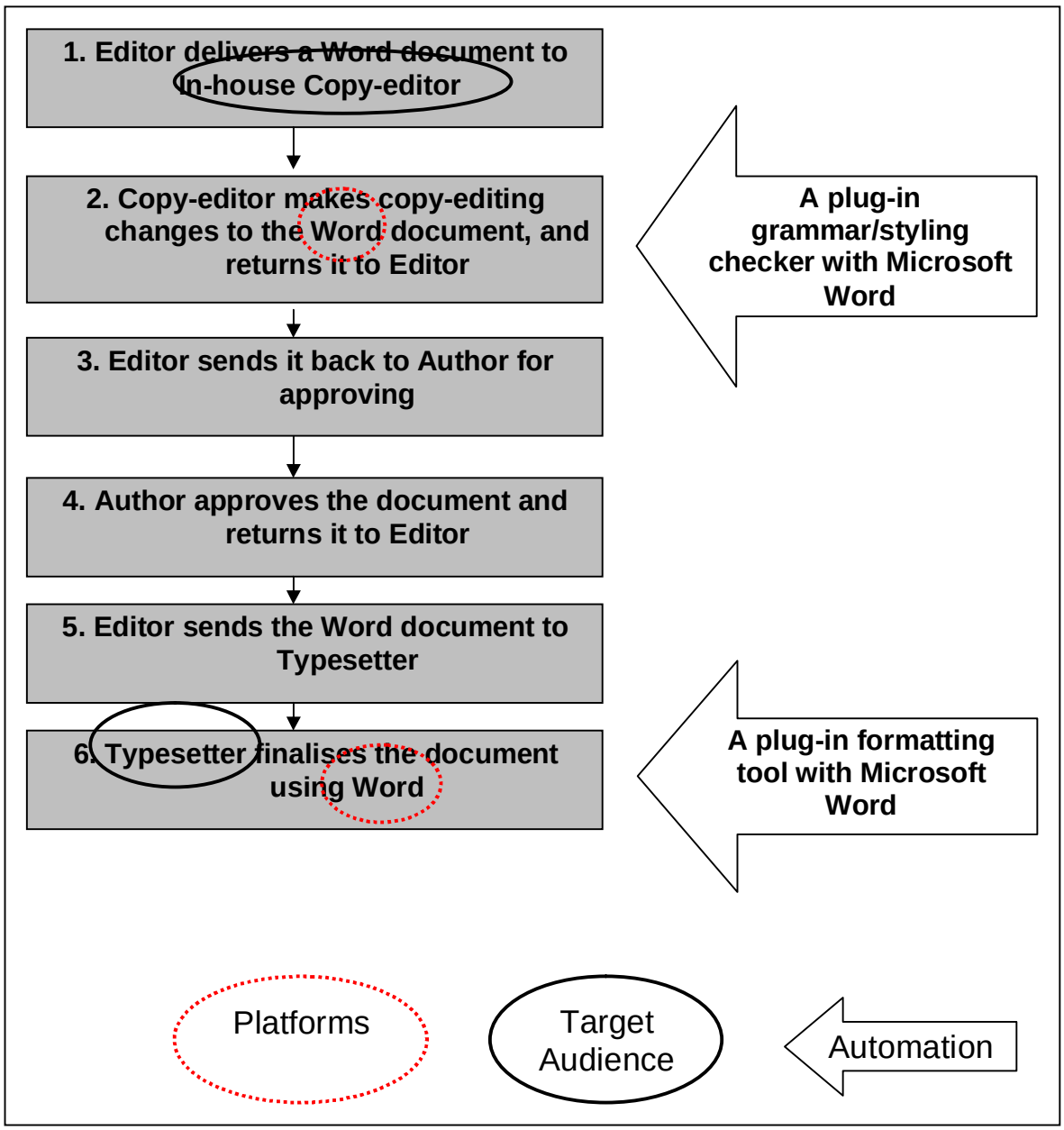

Figure 6 the Automation in Production Workflow One

As we can see in figure 7, the possible automation is similar in the second workflow, whereas copy-editors might need additional automation tools that could be integrated with Adobe Acrobat and provide rapid mark up assistance for PDF files.

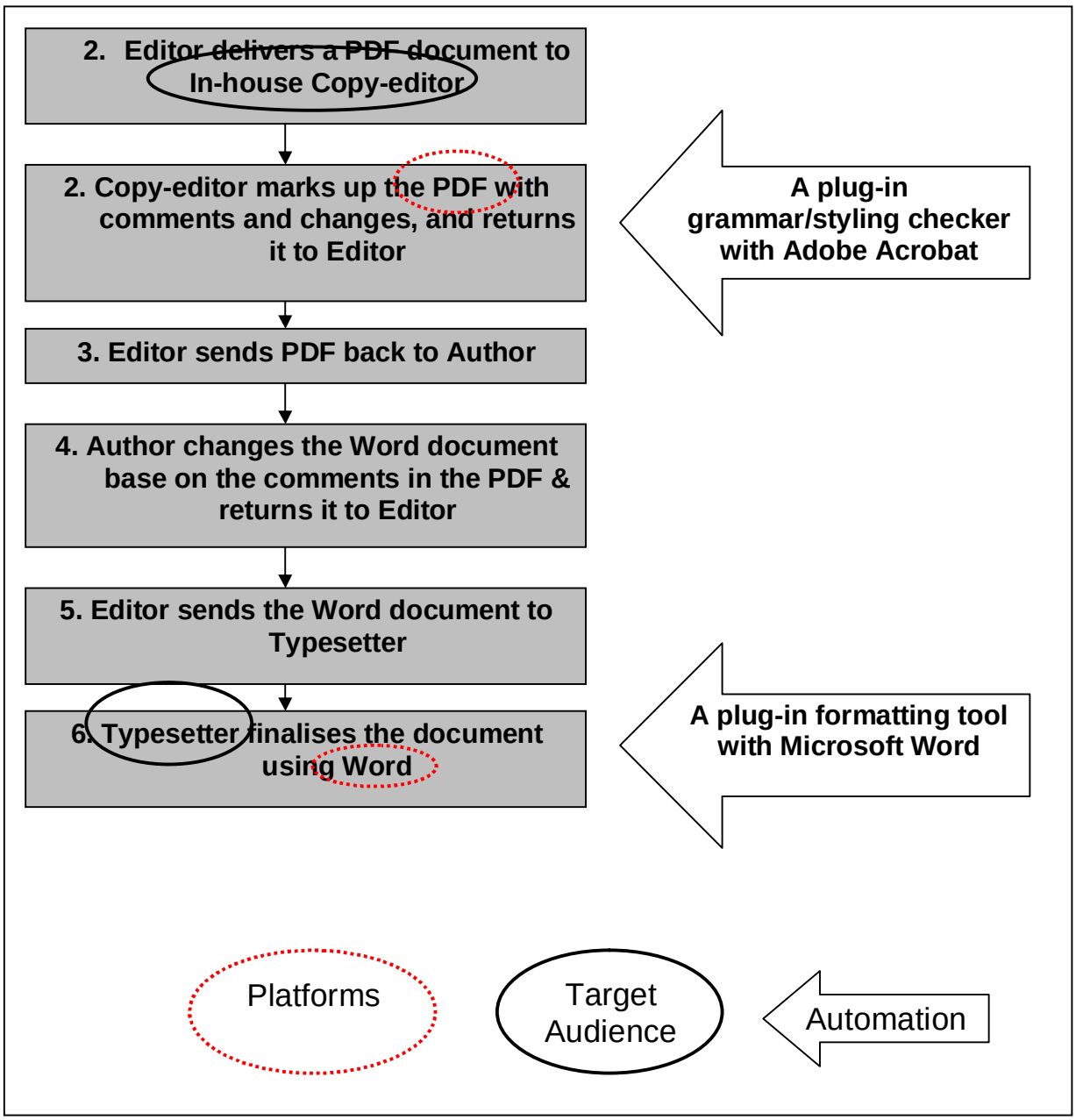

Figure 7 the Automation in Production Workflow Two

For the third workflow, the target audience might also be copy-editors and typesetters. As showed in figure 8, cope-editors might need a plug-in grammar/styling checker with Adobe Acrobat to help them easily correct grammar errors. Typesetters might need a plug-in formatting tool which could be fitted into LaTeX to provide some typesetting assistances.

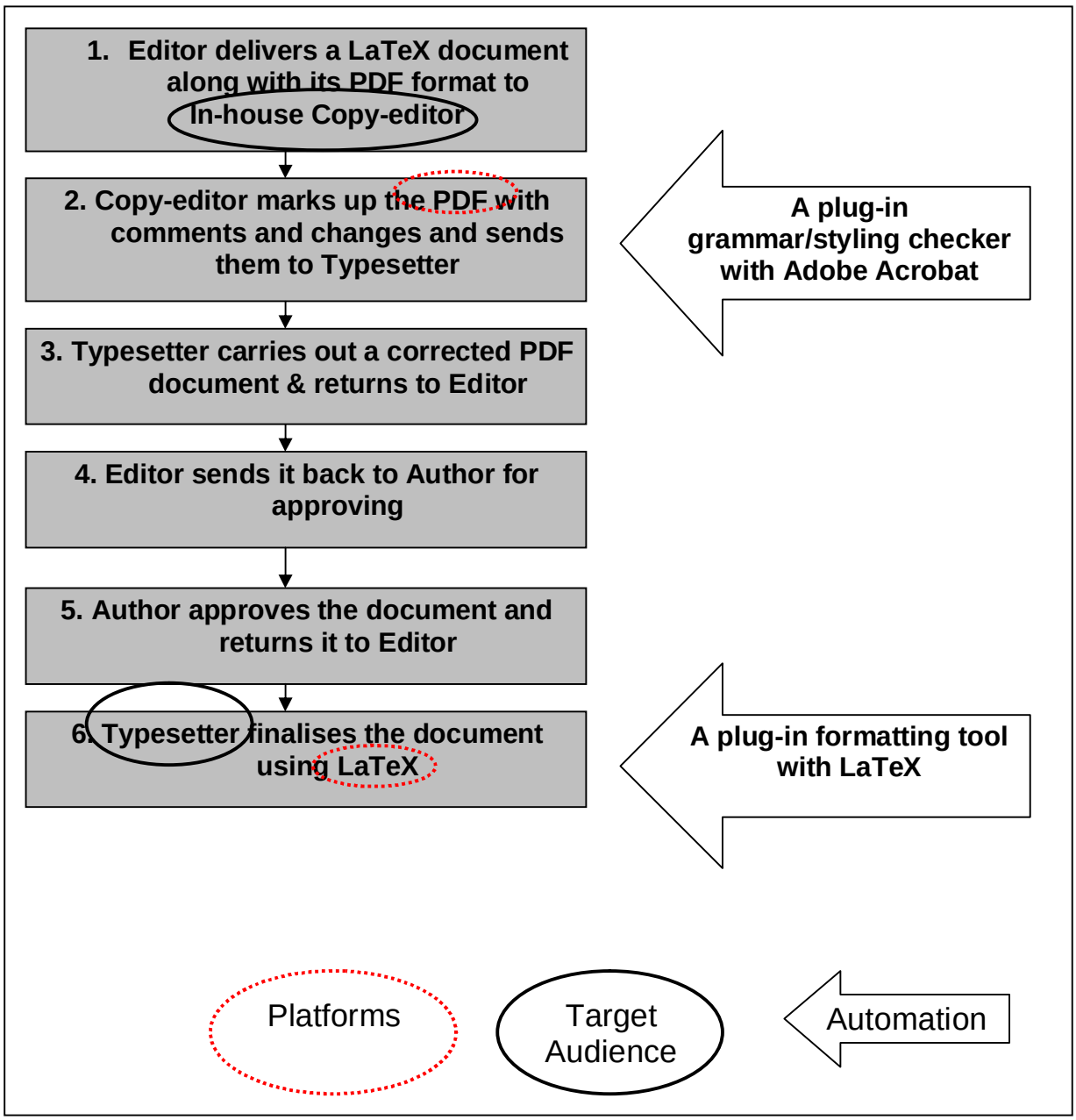

Figure 8 the Automation in Production Workflow Three

There are also some automation tools that already exist in the market. For the editorial process, we can employ those manuscript management systems which are mentioned in section 2.1. We can also introduce some PDF tools, which can annotate, edit and comment PDF files, to peer-reviewers.

For the production process, automated copyediting can be realized with some tools like Cadmus: Rapid Edit[6], Allen: TurnStyle[7] or Editors: Editors Toolkit[8]. Rapid Edit is developed by Cadmus Communications, and can be customized to meet the specific standards of scholarly publishers. It can automate many repetitive style changes, corrections and tagging for copy-editors. Allen: TurnStyle is another smart tool which is also designed to automate repetitive and time-consuming copy-editing tasks. One of its best features is to cross-check in-text citations against reference list in a short time. Because these repetitive tasks such as reference formatting, standardizing abbreviations and eliminating extra spaces have been automated by the smart tools, copy-editors are able to focus on more substantive editorial tasks.

For typesetting, tools like Quark[9] or InDesign[10] are employed to help typesetters compose and prepare articles visually for print production. These tools are reasonably priced, e.g. InDesign costs US$699, and they are designed to process large volumes of manuscripts; so the possible automation proposed in this paper would be a cost performance solution to those scholarly journals which might not have enough budgets.

[6] http://www.cadmus.com/products_and_services/
[7] http://www.allenpress.com/allen_press/gen/allen_press_generated_pages/Allen_TurnStyle_m101.html
[8] https://usd.swreg.org/soft_shop/47578/shopscr5.shtml
[9] http://www.quark.com/
[10] http://www.adobe.com/products/indesign/

6. Conclusion

Scholarly publication is a process involves a lot of text processing tasks. Before we attempt to deploy any automated editorial assistance, we need to understand the variety of different workflows and text processing platforms involved in the publishing process. In this paper, we discuss the workflows of editorial and production processes since they contain most of the text processing procedures. We investigate the production process in detail base on the results from the survey of journal publishing process. We also suggest some possible automation that might be introduced in both editorial and production processes.

Finally, it should be noted that these efforts might not cover the whole picture of journal publishing process since the survey is restricted by its amount of questions and is limited to the computer science discipline: we can see that there are only a small amount of questions regarding typesetting and the results from these questions are not as sufficient as those from questions regarding copy-editing in terms of deploying any further automation. In the future, a larger survey with more number of questions could be conducted to get more specific data of journal production process, especially for typesetting, and may from more disciplines.

Acknowledgement

I would like to thank my supervisor, Professor Robert Dale, for his constant guidance and encouragement. He has offered me a lot of suggestions through this semester. Without these instructions, this paper could not have reached its present form.

I would also like to thank all the editors who have participated in the survey of journal publishing process. They provide significant information for this project to be successful.

References

Campbell, R., Meadows, J. & Page, G., 1997, *Journal Publishing*, Cambridge University Press, UK, Feb 1997.

Ciesielski, J., 2005, "How are electronic journals cooked", 11th Conference on Professional Information Resources, Prague, May 24-26, 2005.

Dale, R., 1990, "A rule-based approach to computer-assisted copy-editing", *Computer Assisted Language Learning*, vol.2.

Dale, R., 1997, "Computer assistance in text creation and editing", *Survey of the state of the art in human language technology*, Cambridge University Press, NY.

Mabe, M., 2003, "The growth and number of journals", *Serials*, vol.16.

Steele, C., 2006, "Research with purpose", *The Australian Higher Education Section*, Sydney.

Appendix A

The Survey of Journal Publishing Process

Question 1: What is the name of your journal?

Question 2: How is copy-editing of accepted manuscripts carried out for your journal?

- o **Copy-editing is carried out by copy-editors on the fulltime staff of your journal's publisher**
- o **Copy-editing is carried out by external freelance copy-editors**
- o **No copy-editing is carried out**
- o **Other (please specify)**

Question 3: How are the copy-editor's revisions incorporated into the manuscript for return to the author for checking?

- o **The copy-editor edits the manuscript directly**
- o **The changes are integrated into the manuscript by in-house typesetters**
- o **The changes are integrated into the manuscript by external typesetters**
- o **The changes are returned to the author for the author to incorporate**
- o **Other (please specify)**

Question 4: In which format(s) do you provide the accepted versions of manuscripts to your copy-editors? (Answer all that apply.)

- o **Hard copy**
- o **Electronic – Word**
- o **Electronic – PDF**
- o **Other (please specify)**

Question 5: In which format do copy-editors indicate the changes to be made to articles?

- o **Hard copy**
- o **PDF with added embedded comments**
- o **Word with comments and changes**
- o **Other (please specify)**

Question 6: Who makes the copy editing changes to the manuscript? (Answer all that apply.)

- **The author**
- **The copy-editor**
- **The typesetter**
- **Other (please specify)**

Question 7: In which format(s) do you provide the final revised versions of manuscripts to your typesetters? (Answer all that apply.)

- **Hard copy**
- **Electronic – Word**
- **Electronic – PDF**
- **Other (please specify)**

Question 8: What text processing platform does your typesetter use to create the final published article? (Answer all that apply.)

- **Microsoft Word**
- **LaTeX**
- **Don't know**
- **Other (please specify)**

Question 9: Please provide comments here on any other relevant aspects of your production process that you do not think are adequately covered above.

Question 10: If you would like to receive a copy of the results of this study, please provide the email address to which it should be sent here.

Appendix B

The List of Participated Journals

ACM Computing Surveys	Journal of Artificial Intelligence Research
ACM TODS	Journal of Bimolecular Screening
ACM Transactions on Asian Language Information Processing	Journal of Community Informatics
ACM Transactions on Modeling and Computer Simulation	Journal of Database Management
Advances in Mathematics of Communications	Journal of documentation
Annals of Mathematics and Artificial Intelligence	Journal of Global Information Technology Management
Anonymous 1	JOURNAL OF INFORMATION ETHICS
Anonymous 2	Journal of Information Science
Artificial Intelligence Journal	Journal of Librarianship and Information Science
Australian Academic and Research Libraries	Journal of Strategic Information Systems
Australian Library Journal	Journal of Systems and Software
Cataloging & Classification Quarterly	Journal of the Medical Library Association
Computational Intelligence	Journal of the Medical Library Association
Computer Speech and Language	Libraries and the Academy
Computers & Education: An International Journal	Library Resources & Technical Services
Constraints Journal	LIBRES
Discrete Applied Mathematics	Libri
Discrete Mathematics	Machine Translation
D-Lib Magazine	Medical Reference Services Quarterly
EBLIP	Multiagent and Grid Systems
Electronic Journal of Information Systems in Developing Countries	Multimedia Systems Journal

Electronic Markets - The International Journal on Networked Business	Neural Computing and Applications
Env. Modeling and Software	Online Information Review
Environmental Modeling & Software	Personal and Ubiquitous Computing
European journal of information systems	Reference & User Services Quarterly
Evidence Based Library and Information Practice	Requirements Engineering
Formal Methods System Design	Research on Language and Computation
Human IT	Science of Computer Programming
IEEE Transactions on Information Theory	Software Practice and Experience
IMAVIS	Software Testing, Verification, and Reliability
Information and Organization	Software: Practice and Experience
Information Research	Speech communication
Information Systems	Telecommunication systems Journal
Information Systems Journal	The Electronic Library
Information, Communication & Society	The Journal of Collaborative Computing
Information & Management	The Journal of Supercomputing
Int. J. of Computational Intelligence Systems	Theoretical Computer Science
International Journal of Computational Geometry and Applications	
International of Pattern Recognition and Artificial Intelligence	
IT and People	