JOURNAL IN-HOUSE STYLE BASED ON MENDELEY'S METADATA EXTRACTION

Dhevi Enlivena Irene Restia Mahelingga*

Direktorat Repositori, Multimedia, dan Penerbitan Ilmiah BRIN

Email: dhevlingga@gmail.com

ABSTRAK: Manajer referensi seperti Mendeley Desktop membantu memudahkan penulisan dalam hal pengutipan. Namun, tidak semua metadata dapat diekstraksi dengan baik. Kesalahan dalam menghasilkan metadata akan merugikan pengarang artikel dan jurnal yang dikutip. Penelitian ini berusaha menemukan gaya selingkung yang dapat secara akurat terdeteksi oleh Mendeley Desktop. Penelitian menggunakan sampel 27 artikel dari 27 jurnal yang berbeda yang diterbitkan oleh LIPI. Sampel diteliti keakuratan ekstraksi metadata menggunakan Mendeley Desktop versi 1.19.4 dan mengacu pada metadata extraction pipeline Mendeley dengan delapan variabel penilaian. Analisis data menggunakan pendekatan deskriptif. Berdasarkan ekstraksi metadata dari sampel dokumen PDF artikel 19.4 didapatkan hasil bahwa seluruh jurnal tidak dapat terekam variabel judul jurnal sehingga persentase keakuratan maksimal adalah 87,5%. Jurnal yang mendapat persentase tersebut hanya empat jurnal. Terdapat sembilan indikator yang bisa diikuti untuk mencapai angka persentase tersebut. Bagi pengelola jurnal, mengikuti tampilan artikel jurnal berbasis ekstraksi metadata Mendeley Desktop atau konvensi Google Scholar perlu dilakukan di era digital. Bagaimanapun kemajuan teknologi informasi memberikan pengaruh pada tampilan visual dan gaya selingkung KTI.

Kata kunci: Jurnal; Mendeley; teknik pengutipan; layout jurnal

ABSTRACT: A reference manager such as Mendeley Desktop is needed to make it easier for scientific paper writers to automatically generate bibliography. However, metadata of several refereed journal articles cannot be extracted properly by Mendeley. Errors in generating a journal's metadata will certainly disadvantage the authors of the cited article and journal. Mistakes in citation can also threaten the citer with plagiarism. This study seeks to find an in-house style that can be accurately detected by Mendeley Desktop as a reference for journal managers to make their journals' metadata can be properly extracted. The study investigates samples of 27 articles from 27 different journals published by LIPI. The accuracy of metadata extraction of the sample articles is examined on Mendeley Desktop version 1.19.4 and refers to Mendeley's metadata extraction pipeline with eight assessment variables. The accuracy percentage from Mendeley Desktop's metadata extraction is analyzed using a descriptive approach. It was found that variable of journal titles in all journals cannot be extracted which gives the maximum accuracy percentage of 87.5%, with only four journals that get this percentage. There are nine indicators that can be followed to achieve this percentage. For journal managers, setting the layout of journal articles based on Mendeley Desktop's or Google Scholar's convention in extracting metadata needs to be considered in the digital era. However, advances in information technology have an influence on the visual layout and the in-house style of scientific papers.

Keywords: Journal; Mendeley; inhouse style; journal layout

110

1. INTRODUCTION

Every journal manager has a distinctive in-house style in the journal publishing process. The in-house style is designed to uniform and ensure the quality of the publications of a publisher (Helmi et al., 2019). The in-house style in journal management is often summarized in the form of author guidelines or templates. The writing guidelines provide guidance on the format or layout of writing, how to cite journal articles or books, how to write datasets, how to mention tools used, and several materials that are directly related to the display and presentation of data in a journal manuscript (Kemenristekdikti, 2018).

The writing guidelines is important because it is referred to by all persons who play a role in the publishing process. Authors write their manuscripts referring to an in-house style from the writing guidelines of a target journal to make their manuscripts accepted and published. Manuscript verifier in the submission process will select and verify the incoming manuscripts to conform to the format and manuscript systematics of a specific in-house style (LIPI Press, 2019). Subsequently, during a journal production process, designers will refer to the in-house style to design the layout and font used in the final layout of a journal article.

In Indonesia, there are several national standards for periodical publications that apply, including those for magazines, news, bulletins, and annual reports in national standard number SNI 19-1950-1990. The national standard refers to the international standard: ISO 8-1977 Documentation-Presentation of Periodicals. The standard contains guidance on the information that should be embedded on the cover page, title page, table of contents, page text, magazine title, magazine number, magazine volume, and so on (SNI 19-1950-1990 Terbitan Berkala, 1992; Purnomowati, 2003).

In addition to the national standard, journal management has also been standardized through the Guidelines for Accreditation of Scientific Journals published by the Ministry of Research, Technology and Higher Education as a follow-up to Permenristekdikti Number 9 of 2018 concerning Accreditation of Scientific Journals (Kemenristekdikti, 2018; Permenristekdikti Nomor 9 Tahun 2018 Tentang Akreditasi Jurnal Ilmiah, 2018). However, there is no specific mention of the standard of the in-house style of a journal article. The standards and accreditation only focus on what should be included in a journal and the aspect of consistency. This indeed seems to not limit the creativity of the manager in setting the in-house style that becomes a distinctive feature of a journal layout.

The adaptation of in-house style is carried out when journal managers want their journal articles to be indexed by scientific writing indexing engines, such as Google Scholar. Google Scholar is an indexing engine that harvests metadata from open access journal publishing sites, university repository pages, document sharing sites, or any other sites that contain any scientific rich files (Mahelingga, 2020a). Google Scholar crawls the web and indexes any document with an academic-looking structure (Martín-Martín et al., 2018). The crawler then follows links to metadata which is then evaluated by the *Google Scholar* algorithm to determine whether or not the information is added to the Google Scholar index (Kenning & Patrick, 2012; Mahelingga, 2021; Williamson & Mirza, 2015). Google Scholar itself has its visual arrangement limitations in categorizing document structures that seem academics through a convention (Google, 2020).

Apart from the adaptation to the indexing engines, there is one other standard that can be referred to in setting the in-house style of a journal, it is reference manager. A reference manager, by the journal managers, is often used as a prerequisite for manuscript submission. The use of a reference manager assists in writing citations in a journal article, such as generating automatic bibliography (Granitzer et al., 2012). The bibliography can be arranged to match the bibliography writing style referred to by the in-house style of the journal. The reference manager not only makes it easier for authors to avoid accidental plagiarism, but it also makes journal managers easier to check citations with the references used (Mahelingga, 2020b) One of the most popular

reference managers, that has also been a prerequisite of journal submission by several journal managers nowadays, is *Mendeley*.

In principle, by uploading a PDF into the Mendeley application, it then can recognize the reference data and extract the metadata automatically which is then enriched with catalog metadata (Gooch & Jack, 2015). However, the writers need to be careful in using Mendeley because sometimes the metadata extraction from a PDF journal uploaded to Mendeley is not very good. This is because Mendeley also has a kind of convention like Google Scholar in classifying metadata from the visual layout of a journal. However, unlike Google Scholar which provides information about the inclusion guidelines on its website, Elsevier as Mendeley's developer does not provide a specific guideline of the convention used in Mendeley to classify metadata of a journal on its website. Even though the guideline is important to be known so that the journal articles can be detected accurately when uploaded to Mendeley which has an impact on the accuracy of the bibliography produced when writing papers.

Most researchers download and collect a lot of research papers in PDF to their desktops for later reading and reference, and to make their work easier, they use a reference manager (Kiran & Reddy, 2018). Inaccurate metadata extraction will affect citations because incorrectly written references will result in citations from authors of articles and journals not recorded in the indexing engine (Garfield, 1990; Kratochvíl, 2017). This will certainly be a disadvantage to those who expect their works to be cited by as many academics as possible. Errors in citing bibliography also result in a plagiarism-alike problem, because such errors can prevent the scientific recognition of deserving work and disrupt the award system for scientific publications, which is the citation (Garfield, 1990).

This study seeks to find an in-house style or journal writing guidelines that can be accurately detected by Mendeley. The study uses 27 sample articles from 27 different journals published by the Indonesian Institute of Sciences (LIPI). The articles are then examined for their accuracy of metadata extraction using Mendeley Desktop version 1.19.4 referring to the Mendeley metadata extraction pipeline. This study aims to find a layout structure that can be perfectly classified through metadata extraction by Mendeley Desktop. As with metadata research in general, it seeks to better describe metadata resources so that they can be more easily targeted to specific uses and can be more easily integrated (Sicilia, 2013) An accurate metadata extraction also plays an important role in supporting the automation of digital library management (Lipinski et al., 2013).

The study is expected to find an in-house style that can accommodate an accurate metadata extraction as a reference for the journal managers to design a journal layout based on a specific standard and knowledge, not just based on personal taste or aesthetics. Through a revision of the writing guidelines, the accuracy of citations is improved to prevent authors' disadvantages due to citation errors generated during *Mendeley*'s metadata extraction. In addition, this research can also be used as a reference for journal managers to assess their journal layout based on the accuracy of Mendeley Desktop's metadata extraction using the same method.

2. LITERATURE REVIEWS

Journal In-house Style

Scientific journals play an important role in the advancement of science, acting as a medium of scientific communication. As a form of communication media, adaptation to technological development and globalization is one of the consequences that should be considered to make the media stay relevant. The development of globalization and information technology requires the standardization of scientific journal publications that can be referred to in designing the in-house style as outlined in the writing guidelines. Indonesian government has published a magazine layout standard, through a national standard number SNI 19-1950-1990, which refers to the international standard ISO 8-1977. The standard contains several guidance on the information that must be embedded on the cover page, title page, table of contents, page text, the title of the magazine, magazine number, magazine volume, and others (Purnomowati, 2003). However, the

guidance about the font's type, size, and placement is not explained in detail. On the other hand, the development of information technology that leads to the automation of metadata extraction needs to be adapted to make it easier for readers and writers of scientific papers to have good access to metadata from the journals they read.

Google's Journal Convention

The guiding convention about the limiting standard for visual layout of PDF journals in Google Scholar version is accessible for journal managers since it is stated on the Google Scholar website clearly (Google, 2020). This convention is important because Google Scholar uses automated software, known as a parser, to identify the bibliographic data of papers, as well as references between papers. Some conventions help the identification process by the Google Scholar parser, which is as follows.

First, the PDF manuscript has the title of the article in the form of text with the largest font size at the top of the page using a minimum font size of 24 pt. All titles are expected to use the same font. It should also be ensured that all other text on the page, particularly the repository or journal name, is set in smaller font size than the paper title, otherwise, this other larger text may be misinterpreted as the document title. In Google convention, it is recommended that article titles use Sentence case to avoid confusion with author names using Title case (Capitalize Each Word). The use of sentence cases for article titles is also adapted by the American Psychological Association in writing bibliography for journals. (American Psychological Association, 2019b) Author names are separated by commas or semicolons and affiliation and title are removed from the author's line. Explicit formatting can be used such as adding the words "by" or "Author:" to emphasize the author's name if needed.

Second, the author of the paper must be listed right before or after the title, in a slightly smaller but still larger font than normal text. It is recommended that the author's name in the PDF use a 16–23 pt font. It is also recommended to use the same font for all author names. The name of the repository or journal is in a smaller font than the author of the paper, otherwise, this other larger text may be mistaken for the author's name.

Third, include the bibliographic citation to the published version of the paper on one line and place it in the header or footer of the first page in the PDF file using an explicit citation format in writing.

Mendeley

Mendeley is a tool for organizing research citations and annotating accompanying PDF articles (Zaugg et al., 2011). One of Mendeley's products is Mendeley Desktop which is compatible with Windows, Mac, and Linux operating systems (MacMillan, 2012). Files maintained for reference by Mendeley are files in PDF format so all metadata information in the PDF file (author name, year, title, journal name, publisher name, volume and issue number, article pages, and other metadata) will be read automatically (Kemenristekdikti, 2018). Regarding the reference PDF format uploaded to Mendeley, based on the Ministry of Research, Technology, and Higher Education guidelines, users are advised to collect PDF reference sources with high resolution so that all metadata information can be read by Mendeley accurately. However, if only low-quality or not-so-good reference sources are obtained, then Mendeley will also not be able to read the metadata information accurately. This will result in wrong data and thus need to be corrected manually (Kemenristekdikti, 2018).

Metadata

Metadata is etymologically defined as data about data (Sicilia, 2013). Metadata is key in the functionality of content storage systems, enabling users to find items of interest, record important information about them, and share that information with others. Metadata is structured information that enables a specific data search more easily. Metadata is an electronic form attached to data or stored in a database. Metadata is called data about data or information about

information (Riley, 2017). Metadata in the form of author's name, year of publication, book titles harvested by the Google Scholar indexing engine or extracted by Mendeley Desktop are categorized as descriptive metadata. More details about metadata types can be seen in Table 1.

Table 1. Metadata Type (Riley, 2017)

Metadata Type	Example Properties	Primary Uses
Descriptive metadata	Title Author Subject Genre Publication date	Discovery Display Interoperability
Technical metadata	File type File size Creation date/time Compression scheme	Interoperability Digital object management Preservation
Preservation metadata	Checksum Preservation event	Interoperability Digital object management Preservation
Rights metadata	Copyright status License terms Rights holder	Interoperability Digital object management
Structural metadata	Sequence Place in hierarchy	Navigation
Markup languages	Paragraph Heading List Name Date	Navigation Interoperability

3. RESEARCH METHODOLOGY

The research is limited to articles from 27 LIPI journals, namely (1) *Journal of Mechatronics, Electrical Power, and Vehicular Technology;* (2) *Oseanologi dan Limnologi di Indonesia;* (3) *Jurnal Kimia Terapan Indonesia;* (4) *BACA: Jurnal Dokumentasi dan Informasi;* (5) *RISET Geologi dan Pertambangan;* (6) *STIPM Journal;* (7) *Teknologi Indonesia;* (8) *Metalurgi;* (9) *LIMNOTEK Perairan Darat Tropis di Indonesia;* (10) *Marine Research Indonesia;* (11) *Oseana;* (12) *REINWARDTIA, A Journal on Taxonomy Botany, Plant Sociology and Ecology;* (13) *Treubia;* (14) *Berita Biologi;* (15) *Anales Bogorienses;* (16) *Journal of Microbial Systematics and Biotechnology;* (17) *Journal of Lignocellulose Technology;* (18) *Widyariset;* (19) *Jurnal Penelitian Politik;* (20) *Jurnal Ekonomi dan Pembangunan;* (21) *Jurnal Elektronika dan Telekomunikasi (JET);* (22) *Jurnal Masyarakat dan Budaya;* (23) *Jurnal Kependudukan Indonesia;* (24) *Jurnal Masyarakat Indonesia;* (25) *Jurnal Kajian Wilayah;* (26) *Journal of Indonesian Social Sciences and Humanities (JISSH);* and (27) *Buletin Kebun Raya.* The sample articles are then examined for the accuracy of metadata extraction using *Mendeley* Desktop version 1.19.4.

The research refers to the Mendeley metadata extraction pipeline (Figure 1) described by Kris Jack, Mendeley's chief data scientist, consisting of five stages, starting from when the PDF is uploaded to the Mendeley Desktop until the metadata is recorded (Gooch & Jack, 2015). The stages include (1) uploaded PDF; (2) PDF converted to text, pdf to XML extracts text from PDF into a format that provides information about size, font, and position of each character on the page; (3) Extracting metadata, information is converted into features that can be understood by the classifier who decides the order of characters representing the title, author list, abstract, or else; (4) Metadata is enriched with a metadata catalog, the extraction results are used to generate queries to the *Mendeley* metadata search API, if a match is found in the Mendeley catalog then the

metadata will be used or if not, the metadata will appear without being enriched; and (5) Mendeley users get metadata.
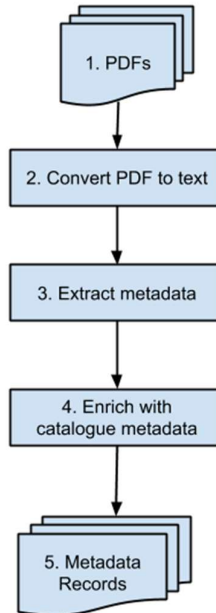

Figure 1. Metadata Extraction Pipeline (Gooch & Jack, 2015)

In this study, the stages of enriching the results of metadata extraction with the Mendeley metadata catalog are intentionally not conducted so that the resulting metadata recording is purely the result of metadata extraction from PDF. This method is taken by not activating the internet or being offline when extracting metadata so that Mendeley Desktop cannot generate queries to the Mendeley metadata search API as shown in Figure 2. Thus, it will be possible to find a good standardized PDF journal layout according to Mendeley.
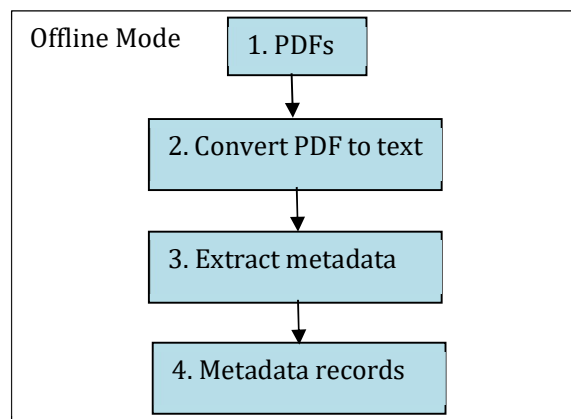

Figure 2. Metadata Extraction Pipeline without Metadata Catalog Enrichment

The assessing indicator is the ability of Mendeley to accurately detect journal descriptive metadata through the accuracy of 8 variables, including (1) article title, (2) author, (3) journal title, (4) year, (5) volume, (6) number, (7) pages, and (8) DOI. The eight variables are information that must be provided in the preparation of a journal bibliography. (American Psychological Association, 2019a) Of the eight metadata variables, a correct metadata extraction gets a value of 2, while the wrong extraction gets 1, and the blank extraction gets a value of 0. The accuracy percentage of each metadata variable is calculated by dividing each total value with a maximum

divisor of 16 as shown in Figure 3. The result of the evaluation is then described to get an overview of journal layout and in-house style, both to follow and to avoid, to get an accurate Mendeley Desktop metadata extraction.

$$\text{Percentage Accuracy} = \frac{\text{Total value}}{\text{Maximum Value}} \times 100$$

Figure 3. Metadata Extraction Accuracy Percentage Formula

The study uses samples from 27 articles in 27 journals organized by LIPI work unit. It cab accessed through LIPI journal portal here ejournal.lipi.go.id. The 27 articles are open access articles published in the last period of 2020. The articles are then uploaded to Mendeley Desktop version 1.19.4 in an offline network to be evaluated their accuracy of descriptive metadata extraction based on the in-house style of each journal respectively. The display results of the extraction are screenshotted as evidence and the accuracy of the resulting metadata extraction is recorded.

4. RESULTS AND FINDINGS

The extraction of PDF metadata using Mendeley Desktop on 27 sample articles from 27 journals managed by work units within LIPI environment is described in Table 2. The extraction of metadata from PDF sample articles using Mendeley Desktop version 1.19.4 resulted in the maximum percentage of accuracy of 87,5%. The 100% accuracy is not achieved because Mendeley Desktop can not extract the journal title variable from any sample of the uploaded articles. It rises a tendency that journal titles are filled only from the metadata catalog on Mendeley's servers.

Table 2. Metadata Extraction Results from 27 PDF Articles

Name	Article Title	Author	Journal Title	Year	Volume	Number	Page	DOI	Total Value	Percentage Accuracy	Description
Journal of Mechatronics, Electrical Power, and Vehicular Technology	1	2	0	2	2	0	2	0	9	56,25	The title of the article is detected incorrectly because the font size of journal title is too large, the issue number is not embedded.
Oseanologi dan Limnologi di Indonesia	1	0	0	2	2	1	2	2	10	62,5	Wrong article title is detected because the font size of the journal title is too large and the issue number is wrongly detected.
Jurnal Kimia Terapan Indonesia	2	2	0	2	2	1	2	0	11	68,75	One author's name is not recorded, the month of publication is detected as issue number.
BACA: Jurnal Dokumentasi dan Informasi	2	2	0	2	1	1	2	0	10	62,5	The title of the article and the author's name are detected only 2 lines respectively, DOI is still using the domain.
RISET Geologi dan Pertambangan	2	2	0	2	2	2	2	2	14	87,5	The title of the article and the author's name are detected only 2 lines respectively.
STIPM Journal	1	2	0	2	0	0	0	0	5	31,25	Licensing rules fill the first page.
Teknologi Indonesia	1	1	0	2	2	2	0	0	8	50	Journal cover is detected as the first page.
Metalurgi	2	1	0	2	0	0	2	0	7	43,75	The title font is small capped and only recorded 1 line, the author's name is detected as the publishing agency.
LIMNOTEK Perairan Darat Tropis di Indonesia	2	2	0	2	2	2	2	0	12	75	DOI is not embedded
Marine Research Indonesia	2	2	0	2	2	2	2	2	14	87,5	
Oseana	1	1	0	2	2	0	2	0	8	50	The spacing between the title and author lines is too wide.

Name	Article Title	Author	Journal Title	Year	Volume	Number	Page	DOI	Total Value	Percentage Accuracy	Description
REINWARDTIA, A Journal on Taxonomy Botany, Plant Sociology and Ecology	1	1	0	2	2	2	0	0	8	**50**	Journal cover is detected as the first page.
Treubia	1	1	0	2	2	2	0	0	8	**50**	Journal cover is detected as the first page.
Berita Biologi	1	1	0	2	2	2	0	0	8	**50**	Journal cover is detected as the first page, issue number using letters.
Anales Bogorienses	2	2	0	2	0	0	0	0	6	**37,5**	Title, volume, issue, and year are placed below and the DOI still uses the domain.
Journal of Microbial Systematics and Biotechnology	2	2	0	2	0	0	2	2	10	**62,5**	The writing of volume and issue using brackets, the author's name is detected with left alignment.
Journal of Lignocellulose Technology	1	1	0	2	2	2	0	0	8	**50**	Journal cover is detected as the first page.
Widyariset	2	1	0	2	2	2	2	0	11	**68,75**	Wrong author name is detected because the title is bilingual, DOI is still using the domain.
Jurnal Penelitian Politik	1	1	0	2	2	2	0	0	8	**50**	Journal cover is detected as the first page.
Jurnal Ekonomi dan Pembangunan	2	2	0	0	0	0	2	0	6	**37,5**	The title of the article is 4 lines, there is no title, volume, issue, and DOI on the first page.
Jurnal Elektronika dan Telekomunikasi (JET)	2	2	0	2	2	2	2	2	14	**87,5**	
Jurnal Masyarakat dan Budaya	2	2	0	2	2	2	2	2	14	**87,5**	Wrong article title is detected due to bilingual title.
Jurnal Kependudukan Indonesia	2	2	0	2	2	2	0	0	10	**62,5**	Wrong article title is detected due to bilingual title, page is separated by "|".
Jurnal Masyarakat Indonesia	1	1	0	0	0	0	0	0	2	**12,5**	Journal cover is detected as the first page, volume is separated with number and year.
Jurnal Kajian Wilayah	1	1	0	2	2	2	2	0	10	**62,5**	Wrong article title is detected because the title is bilingual, the author's name and affiliation has no different, DOI is still using the domain.
Journal of Indonesian Social Sciences and Humanities (JISSH)	2	2	0	0	0	0	0	0	4	**25**	Volume, issue, and year in the header are in JPG format.
Buletin Kebun Raya	1	1	0	2	2	2	2	0	10	**62,5**	Wrong article is detected because the journal title is too large, the article title is wrongly detected as the author's name, DOI is still using the domain.

Out of 27 samples of journal articles, there are only 4 journals received 87.5% accuracy percentage: (1) *RISET Geologi dan Pertambangan*; (2) *Marine Research Indonesia*; (3) *Jurnal Elektronika dan Telekomunikasi (JET);* and (4) *Jurnal Masyarakat dan Budaya*. The metadata of the four journals successfully extracted by Mendeley Desktop is (1) article title, (2) author, (3) year, (4) volume, (5) number, (6) page, and (7) DOI. The four journals have different layouts respectively, but they have similarities in several aspects in general so that the metadata can be extracted properly by *Mendeley* Desktop.

The analysis of metadata extraction, both those correctly extracted and those failed to extract, yields several findings related to indicators that need to be considered during journal layout design, so that the metadata of the journal articles can be extracted properly by *Mendeley*. The indicators are general so they can be implemented in an in-house style without changing the

entire layout of the journal to be very different from the previous one. The nine in-house styles are as follows.

First, it is necessary to apply the conventions of *Google Scholar* because the well-extracted journals generally follow the standards of *Google Scholar*. One of the conventions states that article title has the largest font size, followed by the author's name with a smaller font size but larger than the font size of normal text.

Second, journal title uses smaller font size than the author's name does. Title is followed by volume, number, year, and page. The writing of volume, number, and year is written clearly. For example: "Jurnal Marine Research Indonesia Volume 44 Number 2 Year 2019 pages 42-62" can be written "Mar. Res. Indonesia Vol.44, No.2, 2019: 42-62" or it can also be written in full as " Jurnal Masyarakat dan Budaya, Volume 22 No. 3 Tahun 2020". Volume, number, year, and page should be written in sequential order, separating any of them can cause one of the metadata to fail to be extracted. In some cases, the use of brackets "()" or separator "|" leads to improper extraction. In some cases, the addition of month of publication is misinterpreted as journal number.

Third, DOI writing does not include the domain address. DOI written with a domain address result in DOI unable to be extracted by *Mendeley* Desktop. For convenience, displayed DOI is taken from the prefix or starting from the number 10 backward. A good example of writing for a DOI is "DOI: 10.14203/mri.v44i2.552" instead of "DOI: https://dx.doi.org/10.14203/j.baca.v41i2.563".

Fourth, Mendeley Desktop extracts most of its metadata from the first page of journal articles. The use of the first page of the article for journal cover, or license explanation, as well as French page produces incorrect metadata extraction.

Fifth, all information on the first page, such as journal title, volume, number, year, page, and DOI is in text format, not in image format with JPG or PNG extension. The use of image for captions in header or footer results in inability of *Mendeley* to extract the metadata.

Sixth, article title should not be in two languages with the same font size. The use of article titles in two languages with the same font size results in incomplete extraction of article title.

Seventh, the font size for author's affiliation and/or contact must be different and spaced from the author's name. The contiguous placement between affiliation and author's name and the use of the same font size can cause Mendeley to mistakenly take author's affiliation or contact for the author's name. It occurs because the author's name and affiliation tend to use the same writing format, using Title Case capitalization or capitalize each word.

Eighth, too long article title with more than two lines and the too wide spacing between lines has potential to cause extraction errors. In some cases, the very bottom line of article title that is too far separated can be mistaken for author's name.

Ninth, the use of 'Sentence case' capitalization for article title is highly recommended, to distinguish it from author's name that uses 'Title Case' capitalization. In some cases, the use of 'UPPERCASE' or 'Title Case' capitalization for article title causes the extraction result to be less perfect and is only effective if the article title is in English.

5. DISCUSSION

For journal managers, designing layout of journal articles based on Mendeley Desktop metadata extraction or Google Scholar convention needs to be considered in this era of information technology. However, advances in information technology have influenced the visual layout and in-house style of scientific literature. The writing of references based on the American Psychological Association that omits publisher's location and the use of 'Sentences case' capitalization for article title is also indicated as several compromises to the development of electronic scientific literature, information technology, as well as the automation of parsing, extracting, and metadata classifying technologies.

The use of Mendeley metadata extraction-based in-house style provides the benefit of journals being better extracted in the Mendeley Desktop application. The better metadata extraction practically prevents Mendeley Desktop users from writing a wrong-cited journal bibliography. The accuracy of the cited bibliography will certainly have an impact on the citations obtained from published journal articles which finally becomes an urgency, both for authors of the journal articles and for the reputation of the journal itself.

6. CONCLUSION

In principle, the guiding conventions on journal layout between Mendeley reference manager and Google Scholar indexer are complementary. There are no conflicting indicators between Mendeley and Google Scholar convention. However, unlike Google Scholar as an indexing engine that constantly updates its metadata base by crawling websites that have scientific rich files, Mendeley's metadata catalog collects databases using metadata records extracted from PDFs and user's manual input. This is what makes the layout of journal articles that support the accuracy of the extraction results important and needs to be considered.

Some reputable journals, such as those of Elsevier's, have a good metadata catalog on Mendeley's servers. Meanwhile, pioneering journals generally still have no adequate metadata catalog on Mendeley's serves yet. Therefore, some compromises in the design of journal layout following Mendeley Desktop's metadata extraction convention need to be carried out. Finally, a more accurate metadata recording helps build a better journal database in Mendeley metadata catalog.

REFERENCES

American Psychological Association. (2019a). *Concise guide to APA style, seventh edition*. American Psychological Association.

American Psychological Association. (2019b). *Sentence case capitalization*. https://apastyle.apa.org/style-grammar-guidelines/capitalization/sentence-case

SNI 19-1950-1990 Terbitan berkala, (1992). http://sispk.bsn.go.id/SNI/DetailSNI/2309

Garfield, E. (1990). Journal editor awaken to the impact of citation errors. How we control them at ISI. *Current Contents*, *13*(41), 367–375.

Gooch, P., & Jack, K. (2015). *How well does Mendeley's Metadata Extraction Work?* Making Tools for Researchers. https://krisjack.wordpress.com/2015/03/12/how-well-does-mendeleys-metadata-extraction-work/

Google. (2020). *Google Scholar Help*. https://scholar.google.com/intl/id/scholar/inclusion.html#indexing

Granitzer, M., Hristakeva, M., Jack, K., & Knight, R. (2012). A comparison of metadata extraction techniques for crowdsourced bibliographic metadata management. *Proceedings of the ACM Symposium on Applied Computing*, 962–964. https://doi.org/10.1145/2245276.2245462

Helmi, R. L., Safitri, M., Fazi, R., Mahelingga, D. E. I. R., Suhendra, M. F., Hartiningsih, R. W., Indrasari, N. P., Kusuma, S. H., Taslima, R. H., Sasiwi, P., Dewi, S. P., Helmiawan, M., & Puspita, A. D. (2019). *Pedoman Penerbitan Buku LIPI Press*. LIPI Press. https://doi.org/10.14203/press.52

Kemenristekdikti. (2018). *Pedoman akreditas jurnal ilmiah.* Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi.

Permenristekdikti Nomor 9 Tahun 2018 tentang Akreditasi Jurnal Ilmiah, 1 (2018).

Kenning, A., & Patrick, S. O. (2012). Invisible institutional repositories: addressing the low indexing ratios of IRs in Google. *Library Hi Tech*, *30*(1), 60–81. http://dx.doi.org/10.1108/07378831211213210

Kiran, M. K., & Reddy, K. T. (2018). An approach towards establishing reference linking in desktop reference manager. *Journal of Information and Knowledge Management*, *17*(3), 1–22.

https://doi.org/10.1142/S021964921850034X

Kratochvíl, J. (2017). Comparison of the Accuracy of Bibliographical References Generated for Medical Citation Styles by EndNote, Mendeley, RefWorks and Zotero. *Journal of Academic Librarianship*, *43*(1), 57–66. https://doi.org/10.1016/j.acalib.2016.09.001

LIPI Press. (2019). *Panduan Penelaahan dan Penilaian Naskah Buku Ilmiah*. LIPI Press.

Lipinski, M., Yao, K., Breitinger, C., Beel, J., & Gipp, B. (2013). Evaluation of header metadata extraction approaches and tools for scientific PDF documents. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, 385–386. https://doi.org/10.1145/2467696.2467753

MacMillan, D. (2012). Mendeley: Teaching scholarly communication and collaboration through social networking. *Library Management*, *33*(8), 561–569. https://doi.org/10.1108/01435121211279902

Mahelingga, D. E. I. R. (2020a). Penerbitan buku ilmiah daring berbasis open monograph press (OMP). *Berkala Ilmu Perpustakaan*, *16*(2), 155–169. https://doi.org/10.22146/bip.v16i1.265

Mahelingga, D. E. I. R. (2020b). *Pengenalan Manajemen Referensi Mendeley*. https://doi.org/10.31219/osf.io/g2yrj

Mahelingga, D. E. I. R. (2021). Indeks dan Sitasi Buku Ilmiah di Google Scholar. *Pustakaloka*, *13*(1), 22–39. https://doi.org/10.21154/pustakaloka.v13i1.2379

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google scholar, web of science, and scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, *12*(4), 1160–1177. https://doi.org/10.1016/J.JOI.2018.09.002

Purnomowati, S. (2003). Penampilan Majalah Ilmiah: Standar Dan Penerapannya. *Baca: Jurnal Dokumentasi Dan Informasi*, *27*(1), 20–27. https://doi.org/http://dx.doi.org/10.14203/j.baca.v27i1.68

Riley, J. (2017). Understanding Metadata: What is Metadata, and What is it For? In *Cataloging & Classification Quarterly* (Vol. 55, Issues 7–8). National Information Standards Organization (NISO). https://doi.org/10.1080/01639374.2017.1358232

Sicilia, M. A. (2013). Metadata research: Making digital resources useful again? *Handbook of Metadata, Semantics and Ontologies*, 1–8. https://doi.org/10.1142/9789812836304_0001

Williamson, P. O., & Mirza, R. (2015). Does Google Scholar help or hurt institutional repositories? *The Complete Guide to Using Google in Libraries: Instruction, Administration, and Staff Productivity*, 211–222. https://books.google.com/books?hl=en&lr=&id=Pv3bBgAAQBAJ&pgis=1

Zaugg, H., West, R. E., Tateishi, I., & Randall, D. L. (2011). Mendeley: Creating Communities of Scholarly Inquiry Through Research Collaboration. *TechTrends*, *55*(1), 32–36. https://doi.org/10.1007/s11528-011-0467-y