Aalto-yliopisto
Sähkötekniikan
korkeakoulu

Master's Programme in Data Science

Automating Information Extraction from Non-Standard Financial Reports Using Large Language Models

Enhancing Efficiency through Format-Aware Extraction with Large Language Models

Gabriel Gomes Ziegler

Master's Thesis
2024

A" Aalto-yliopisto
Sähkötekniikan
korkeakoulu

Author	Gabriel Gomes Ziegler
Title	Automating Information Extraction from Non-Standard Financial Reports Using Large Language Models — Enhancing Efficiency through Format-Aware Extraction with Large Language Models
Degree programme	Data Science
Major	ICT Innovation
Supervisor	Prof. Bo Zhao
Advisor	MS Manne Larsson (MSc)
Collaborative partner	Datia

Date 21 September 2023   **Number of pages** 17+15   **Language** English

Abstract

The abstract is a short description of the essential contents of the thesis, usually in one paragraph: what was studied and how and what were the main findings.

For a Finnish thesis, the abstract should be written in both Finnish and English; for a Swedish thesis, in Swedish and English. The abstracts for English theses written by Finnish or Swedish speakers should be written in English and either in Finnish or in Swedish, depending on the student's language of basic education. Students educated in languages other than Finnish or Swedish write the abstract only in English. Students may include a second or third abstract in their native language, if they wish.

The abstract text of this thesis is written on the readable abstract page as well as into the pdf file's metadata via the \thesisabstract macro (see comment in this TeX file above). Write here the text that goes onto the readable abstract page. You can have special characters, linebreaks, and paragraphs here. Otherwise, this abstract text must be identical to the metadata abstract text.

If your abstract does not contain special characters and it does not require paragraphs, you may take advantage of the \abstracttext macro (see the comment in this TeX file below).

Keywords For keywords choose, concepts that are, central to your, thesis

**Aalto-yliopisto
Sähkötekniikan
korkeakoulu**

Tekijä Gabriel Gomes Ziegler	
Työn nimi Opinnäyteen otsikko — Opinnäytteen mahdollinen alaotsikko	
Koulutusohjelma Elektroniikka ja sähkötekniikka	
Pääaine Sopiva pääaine	
Työn valvoja Prof. Pirjo Professori	
Työn ohjaajat TkT Alan Advisor, DI Elsa Expert	
Yhteistyötaho Yhtiön tai laitoksen nimi (tarvittaessa)	
Päivämäärä 21.9.2023 **Sivumäärä** 17+15 **Kieli** englanti	

Tiivistelmä

Tiivistelmä on lyhyt kuvaus työn keskeisestä sisällöstä usein yhtenä kappaleena: mitä tutkittiin ja miten sekä mitkä olivat tärkeimmät tulokset. Suomenkielisen opinnäytteen tiivistelmä kirjoitetaan suomeksi ja englanniksi ja ruotsinkielisen vastaavasti ruotsiksi ja englanniksi. Suomen- tai ruotsinkielisten opiskelijoiden, joiden opinnäytteen kieli on englanti, tulee kirjoittaa tiivistelmänsä englanniksi ja koulusivistyskielellään. Muiden kuin koulusivistyskieleltään suomen- tai ruotsinkielisten tulee kirjoittaa tiivistelmänsä vain englanniksi. Opiskelija voi halutessaan lisätä opinnäytteeseensä toisen tai kolmannen tiivistelmän omalla äidinkielellään. Tämän opinnäytteen tiivistelmäteksti kirjoitetaan opinnäytteen luettavan osan lomakkeen lisäksi myös pdf-tiedoston metadataan. Kirjoita tähän metadataan kirjoitettavaa teksti. Metadatatekstissä ei saa olla erikoismerkkejä, rivinvaiho- tai kappaleenjakomerkkiä, joten näitä merkkeja ei saa käyttää tässä. Jos tiivistelmäsi ei sisällä erikoimerkkejä eikä kaipaa kappaleenjakoa, voit hyödynttää makroa abstracttext luodessasi lomakkeen tiivistelmä (katso kommentti tässä TeX-tiedostossa alla). Metadatatiivistelmatekstin on muuten oltava sama kuin lomakkeessa oleva teksti.

Avainsanat Vastus, resistanssi, lämpötila

A" Aalto-yliopisto
Sähkötekniikan
korkeakoulu

Författare	Gabriel Gomes Ziegler
Titel	Arbetets titel — Opinnäytteen mahdollinen alaotsikko
Utbildningsprogram	Electronik och electroteknik
Huvudämne	Sopiva pääaine
Övervakare	Prof. Pirjo Professori
Handledare	TkD Alan Advisor, DI Elsa Expert
Samarbetspartner	Company or institute name in Swedish (if relevant)

Datum 21.9.2023	**Sidantal** 17+15	**Språk** engelska

Sammandrag

Sammandraget är en kort beskrivning av arbetets centrala innehåll: vad undersöktes, hur undersöktes det och vilka var de viktigaste resultaten?

I lärdomsprov som skrivs på svenska skrivs sammandraget på svenska och engelska, på motsvarande sätt skrivs sammandraget på finska och engelska i lärdomsprov på finska. Finsk- eller svenskspråkiga studerande som skriver sitt lärdomsprov på engelska ska skriva sammandraget på engelska och på sitt skolutbildningsspråk. Studerande vars skolutbildningsspråk inte är svenska eller finska skriver sammandraget endast på engelska. Den studerande kan om hen så önskar lägga till ett andra eller tredje sammandrag på sitt eget modersmål. Sammandraget fungerar då ofta som mognadsprov och bör i så fall vara minst 300 ord långt. Information om mognadsprov på svenska finns på MyCourses:
https://mycourses.aalto.fi/course/view.php?id=26872.

Nyckelord Nyckelord på svenska, temperatur

Preface

I want to thank Professor Pirjo Professor and my instructors Dr Alan Advisor and Ms Elsa Expert for their guidance.

I also want to thank my partner for keeping me sane and alive.

Otaniemi, 9 February 2023

Eddie E. Engineer

Contents

Symbols and abbreviations

Symbols

$\mathbf{B}$	magnetic flux density
c	speed of light in vacuum $\approx 3 \times 10^8$ [m/s]
ω_D	Debye frequency
ω_latt	average phonon frequency of lattice
$\uparrow$	electron spin direction up
$\downarrow$	electron spin direction down

Operators

$\nabla \times \mathbf{A}$	curl of vectorin $\mathbf{A}$
$\dfrac{\mathrm{d}}{\mathrm{d}t}$	derivative with respect to variable t
$\dfrac{\partial}{\partial t}$	partial derivative with respect to variable t
$\sum_i$	sum over index i
$\mathbf{A} \cdot \mathbf{B}$	dot product of vectors $\mathbf{A}$ and $\mathbf{B}$

Abbreviations

AC	alternating current
APLAC	an object-oriented analog circuit simulator and design tool (originally Analysis Program for Linear Active Circuits)
BCS	Bardeen-Cooper-Schrieffer
DC	direct current
TEM	transverse eletromagnetic

1 Introduction

1.1 Background of the Field of Study

The field of data extraction from financial reports has evolved significantly with advancements in text processing and machine learning technologies. Historically, this task involved manual data entry or rule-based systems that were labor-intensive and prone to errors. The emergence of Large Language Models (LLMs), such as GPT (Generative Pre-trained Transformer) and BERT (Bidirectional Encoder Representations from Transformers), has revolutionized this domain. These models have the ability to understand and extract complex financial information from unstructured data, thereby increasing accuracy and efficiency. Recent studies have demonstrated the potential of LLMs in automating financial data extraction, highlighting improvements in processing time and data accuracy over traditional methods.

1.2 General Objective

This study aims to extend the current capabilities of data extraction systems by incorporating advanced Large Language Models (LLMs) and exploring novel methodologies in the field. The primary goals include: ellaborating a comprehensive comparison of methods for extracting information from financial reports, with a focus on non-standard reports,
    enhancing the precision and efficiency of data extraction from financial reports, developing a scalable system capable of processing large volumes of data, and comparing the effectiveness of various LLMs and extraction techniques. By achieving these goals, the study seeks to contribute to the broader understanding of automated data extraction and its application in financial analysis.

1.3 Primary Research Question and Sub-Problems

The primary research question of this study focuses on: "How can Large Language Models (LLMs) be optimized for more accurate and efficient extraction of financial data from unstructured reports?" Sub-problems in this line of inquiry include: identifying

the most effective LLM architectures for financial data recognition, developing methodologies for context-aware data extraction, enhancing the system's ability to handle diverse report formats, and evaluating the impact of training data quality and volume on model performance. These sub-problems are essential for understanding the intricacies of applying LLMs to financial data extraction and for developing a comprehensive solution.

Scope and Constraints

The scope of this study is limited to the extraction of financial data from English-language reports, focusing on publicly available annual and quarterly financial statements. Key constraints include the variability in report formats, the complexity of financial terminology, and the inherent limitations of current LLM technologies in understanding domain-specific contexts. The study primarily revolves around the use of GPT and BERT models, considering their widespread adoption and state-of-the-art performance in text processing tasks. Main concepts involved include natural language processing (NLP), machine learning, data extraction, and financial analysis, with a particular emphasis on the adaptation and optimization of LLMs for specialized data extraction tasks.

2   Literature review

Ever since Portable Document Formats (PDFs) were created by Adobe in 1993, they have been used to store and share information. These document standard quickly became a way of companies reporting their financial information for the public as well as KPIs and other important information internally. This has led to a large amount of information being stored in PDFs, which has led to a need to extract information from these files. A series of professions have arised from this need, such as data entry, data extraction, and data analysis. The extraction of information from PDFs has been a manual process for most of the tasks until recent years, when Optical Character Recognition (OCR) and Natural Language Processing (NLP) technologies have been developed to automate processes involving processing PDFs.

Extracting information from a document, recently referred to "Document AI" is a complex problem that often involves cross-modal interactions where information is represented in both text and visual form. This is particularly true for financial reports, where information is often presented in tables, charts, and text. The problem is further complicated by the fact that financial reports are often not standardized, and the information is presented in diverse range of formats. methods:

vector search gpt4 multimodal vision Q&A with RAG

2.1   Structure of the thesis

The thesis is composed by a comprehensive comparison of methods for extracting information from financial reports, with a focus on non-standard reports. The thesis is structured as follows:

1. Introduction

2. Literature review

3. Research material and methods

4. Results / Findings

5. Discussion

6. Summary / Conclusions

7. References

3   Research material and methods

This part is the core of your work, where you explain the methodological choices you made, its limitations, how you pick your research material or subjects, the implementation of your study and the methods used. This section determines the methodological strengths and weaknesses of your thesis. Any earlier description of the method should limit itself to work done earlier by others. Here you tell your reader what you have done.

4 Results

Present the results of your study here and answer the research questions, asked earlier in the thesis (in the introduction, perhaps), this study strives to answer. The scientific value of your work is measured by the results you obtain along with the arguments you give to back the answers to your research questions.

Be critical of the significance of your results. You may critically scrutinise the results and your interpretation of the results here, or you may do so later in the chapter with the discussion of your work or in the conclusions part.

This part should discuss how reliable the data used in the study are. You may discuss the reliability of the conclusions drawn from the study either in this chapter or later in the discussions part. You may have the discussion in a chapter of its own, separate from the summary or conclusions.

5 Summary/Conclusions

This is where you tie up any loose ends. Tell your reader briefly and clearly what you have done, what you have discovered, and the value of your discovery in the context of similar work done earlier. Draw clear conclusions regarding the research problem, sub-problems or hypotheses. You also discuss future lines of study and new questions your study might have posed.

As the author of the thesis, you alone are responsible for ensuring that the layout, form and structure of your thesis adheres to the guidelines outlined by your school. This template aims to help you meet these requirements.

References

This is the list of references to the sources cited in appendix B. The list more or less follows the Vancouver style (IEEE). See appendix C for a detailed exposition on cross-referencing and bibliography styles. Follow the description there.

[1] Citation Guide: Making a bibliography, *Aalto University Learning Centre*. Online article. Available `https://libguides.aalto.fi/c.php?g=41067 4&p=2797572` (accessed on 14.7.2021)

[2] Bringhurst, R., *Horizontal Motion. The Elements of Typographic Style*, Point Roberts, WA: Hartley & Marks, 1992. p. 26, pp. 25–36. Also available online as version 3.0 at `https://smallpressblog.files.wordpress.com/201 7/11/bringhurstelementsselections1.pdf` (accessed on 7 May 2021).

[3] de Buen Unna, J., *Manual de diseño editorial*, 4. ed. corrigida y aumentada, Somonte-Cenero, Gijón: Ediciones Trea, 2014.

[4] Dyson, M. C., and Kipping, G. J., "The Effects of Line Length and Method of Movement on Patterns of Reading from Screen," *Visible Language,* vol. 2, no. 2, pp. 150–181, 1998.

[5] Shaikh, A. D., "The Effects of Line Length on Reading Online News," *Usability News*, vol. 7, no. 2, July 2005.

[6] Bailey, C., *The Basics of Typography*. [Online]. `https://www.webfx.com/bl og/web-design/the-basics-of-typography` (accessed on 14 July 2021).

[7] Wikipedia contributors, "Line length," *Wikipedia: The Free Encyclopedia*, Wikimedia Foundation, Inc., 22 July 2004. `https://en.wikipedia.org /w/index.php?title=Line_length&oldid=997524503` (accessed 7 May 2021).

[8] Wikipedia contributors, "Leading," *Wikipedia, The Free Encyclopedia*, 2021. `https://en.wikipedia.org/w/index.php?title=Leading&oldid=102 6690618` (accessed 14 July 2021).

[9] "Visual elements," *Aalto University Brand Library*. [Online]. `https://www.aa lto.fi/en/brand-library#/visual-elements/typography` (accessed 14 July 2021)

The reference list that follows are examples containing the required information about the cited sources, and it more or less abides to the Vancouver style (IEEE). See appendix C for guidelines for making your reference list correctly.

[10] Kauranen, I., Mustakallio, M. and Palmgren, V., *Tutkimusraportin kirjoittamisen opas opinnäytetyön tekijöille*, Espoo, Teknillinen korkeakoulu, 2006.

[11] Itkonen, M., *Typografian käsikirja*, 3rd edition, Helsinki, RPS-yhtiöt, 2007.

[12] Koblitz, N., *A Course in Number Theory and Cryptography. Graduate Texts in Mathematics 114*, 2nd edition, New York, Springer, 1994.

[13] Bardeen, J., Cooper, L. N. ja Schrieffer, J. R., "Theory of Superconductivity," *Physical Review,* 1957, vol. 108, no. 5, pp. 1175–1204.

[14] Deschamps, G. A., "Electromagnetics and Differential Forms," *Proceedings of the IEEE,* 1981, vol. 69, no. 6, pp. 676–696.

[15] Sihvola, A. et al., "Interpretation of measurements of helix and bihelix superchiral structures," in: Jacob, A. F. and Reinert, J. (Eds.) *Bianisotropics '98 7th International Conference on Complex Media*, Braunschweig, 3.–6.6.1998, Braunscweig, Technische Universität Braunschweig, 1998, pp. 317–320.

[16] Lindblom-Ylänne, S. and Wager, M., "Tieteellisten opinnäytetöiden ohjaaminen," in: Lindblom-Ylänne, S. and Nevgi, A. (Eds.), *Yliopisto- ja korkeakouluopettajan käsikirja*, Helsinki, WSOY, 2004, pp. 314–325.

[17] Miinusmaa, H., *Neliskulmaisen reiän poraamisesta kolmikulmaisella poralla*, Master's thesis, Teknillinen korkeakoulu, Department of Mechanical Engineering, Espoo, 1977.

[18] Loh, N. C., *High-Resolution Micromachined Interferometric Accelerometer*, Master's thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1992.

[19] Lönnqvist, A., *Applications of hologram-based compact range: antenna radiation pattern, radar cross section, and absorber reflectivity measurements*, Doctoral thesis, Teknillinen korkeakoulu, Department of Electrical and Telecommunications Engineering, 2006.

[20] SFS 5342, *Kirjallisuusviitteiden laatiminen*, 2nd edition, Helsinki, Suomen standardisoimisliitto, 2004. 20 p.

[21] Palmgren, V., Planning Officer, Teknillinen korkeakoulu, library, Otaniementie 9, 02150 Espoo. Interview, 15.1.2007.

[22] Ribeiro, C. B., Ollila, E. and Koivunen, V., "Stochastic Maximum-Likelihood Method for MIMO Propagation Parameter Estimation," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 46–55. [Online]. Accessed 19.1.2007. Also available in print. DOI: 10.1109/TSP.2006.882057.

[23] Stieber, T., "GnuPG Hacks," *Linux Journal.* Online publication, 2006, November, no. 143. Accessed 19.1.2007. Also available in print. Available: `http://www.linuxjournal.com/article/8732.`

[24] Pohjois-Koivisto, T., "Voiko kone tulevaisuudessa arvata tahtosi?," *Apropos.* Online publication, February, no. 1, 2005. Accessed 19.1.2007. Available: `http://www.apropos.fi/1-2005/prima.php.`

[25] Adida, B., *Advances in Cryptographic Voting Systems*, PhD Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2006. Online document. Accessed 19.1.2007. Available: `http://crypto.csail.mit.edu/~cis/theses/adida-phd.pdf.`

[26] Kilpeläinen, P., *WWW-lähteisiin viittaaminen tutkielmatekstissä.* Online article. Updated 26.11.2001. Accessed: `http://www.cs.uku.fi/~kilpelai/wwwlahteet.html.`

A Contents of an appendix

Appendices are not essential in a thesis, and so you must plan the content of your thesis as if it does not contain an appendix. The appendix cannot be used as a dumping ground for text and ideas from an overgrown thesis.

An appendix is an independent entity, even though it complements the thesis. So, the appendix is not, say, just a list or image or table, but contains explanatory text as well that indicates the purpose of its content. It can contain code listings, like the one below for a simplified list of commands to create an appendix.

```
\clearpage
\appendix
\addcontentsline{toc}{section}Contents of an appendix}
\thispagestyle{empty}
\section*{Contents of an appendix}
...
text
```

Equation numbering in the appendix forms a separate, complete entity. Here are a couple of examples how equations in an appendix are numbered:

$$(x + a)^n = \sum_{k=0}^{n} \binom{a}{b} x^n a^{n-k}, \tag{A1}$$

$$\sin\alpha \pm \sin\beta = 2\sin\left(\frac{\alpha \pm \beta}{2}\right)\sin\left(\frac{\alpha \mp \beta}{2}\right). \tag{A2}$$

The appendix can contain figures that do not fit in to complement the text in the thesis. The numbering of figures is like that of equations: see figure **??**.

The numbering of tables is like that for equations and figures, as is evident from the caption of table A1.

Table A1: Caption for the table.

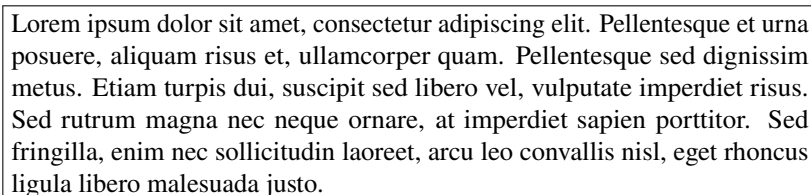|	
9.00–9.55	Safety instructions on the use of laboratories
9.55–10.00	Transfer to the laboratory

B   Page layout and typographical design

Layout choices

Designing a visually pleasing, balanced and easily readable document requires setting several typographic parameters, one of the most important ones being the line length. If too wide, the reader's eyes have trouble focusing on the text because the line length makes it difficult to gauge when the line begins and where it ends; too narrow, and the reader's eyes have to travel back too often, breaking their reading rhythm and often causing them to begin the next line before finishing the current one.

Traditional research on the effect of line length on readability of printed text have found that 45–75 characters per line (cpl) is an acceptable range, with 66 cpl being the ideal [2, 7]. This number includes letters, numbers and spaces. The line length in conventional books tends to be 30 times the type size, but anything between 20 and 40 times is considered acceptable. For example, for a 10-pt font, an acceptable line length is 300 points (30×10), or about 10,58 cm. The reader's experience as a reader also affects the preferred line length; an experienced reader can handle between 45 and 80 cpl for a comfortable reading experience, whereas a novice prefers a line length of between 34 and 600 cpl [3].

Reading text from a screen poses challenges absent in paper: glare, flicker and scrolling. Research seems to indicate that longer lines are better for scanning through the text whereas shorter lines are preferred for accurate reading. One study says that reading speed at a certain level of comprehension seems to be better for longer lines (100 cpl) than for shorter lines (25 cpl) [4]. Another study [5] indicates that subjective preferences for longer or shorter line lengths appear to be contradictory. About 60% of the test subjects preferred the presented shortest (35 cpl) or longest (95 cpl) lines, and at the same time all of them disliked either the shortest or the longest lines. (See also [7].)

Another important typographical parameter is line leading or leading, which refers to the distance between adjacent lines of text. Double spacing is a practice from the era of typewriters particularly in academia to allow making handwritten comments in documents. Typewriters had a limited number of options for leading, and double spacing was chosen as a norm. Too much leading can cause continuity problems, since the reader's eyes must travel a greater distance between lines of text. The amount

> Lorem ipsum dolor sit amet, consectetur adipiscing elit. Pellentesque et urna posuere, aliquam risus et, ullamcorper quam. Pellentesque sed dignissim metus. Etiam turpis dui, suscipit sed libero vel, vulputate imperdiet risus. Sed rutrum magna nec neque ornare, at imperdiet sapien porttitor. Sed fringilla, enim nec sollicitudin laoreet, arcu leo convallis nisl, eget rhoncus ligula libero malesuada justo.

Figure B1: Text typeset in a 300-pt-wide box using a font size of 10 pt. The resulting average number of characters per line is about 60. The text is framed to make the box size evident.

of leading is a compromise between ease of reading, desired efficiency in the use of vertical space, weight and type (serif or sans serif) of the typeface used, and visual aesthetics. Naturally, whether a document is printed or published online affects the leading, but even the language of the text must be considered when deciding on the required leading. (See also [8].)

Aalto University's visual guideline for writing documents

Of Aalto University's visual guidelines [9], the one that applies to document writing is related to the use of fonts. This guideline specifies that the body of the text be in the serif font Sentinel and the sec-tion titles in boldface of the sans serif font Nimbus Sans. These fonts should be available on all Aalto computers. However, being a commercial product, Sentinel can be replaced with Georgia and Nimbus Sans with Arial, both of which come installed in all Windows machines. Thus, Georgia and Arial are the fonts used in the Word template. The font choices for this LaTeX template are a Times and Helvetica clone from the `newtxtext` package since it provides support for math fonts and its output is a PDF/A compliant pdf.

Layout and typographical specifications

Page layout in the thesis

The thesis is typeset on A4-sized paper. The text width is set to an average of about 75 cpl as a compromise between having to read the thesis on a computer screen and on paper, as discussed above. For the font size of 12 pt to be used in the body text, the text width works out to 14,2 cm. For the online version of the document, the text column is centred, implying that the left and right margins are both 3,4 cm. If you want to print the document and bind it, the binding margin must be 4,8 cm. The text height is set to 23 cm by setting the top margin to 3,7 cm and the bottom margin to 3 cm. The layout dimensions are summarised in table B1.

Table B1: Page layout dimensions.

Paper size	A4
Text width	14,2 cm
Top margin	3,7 cm
Bottom margin	3,0 cm
Online document	
Left margin	3,4 cm
Right margin	3,4 cm
Printed document	**(for binding)**
Left margin	4,8 cm
Right margin	2,0 cm

Sectioning and text body

The font for the body text is a 12-pt serif Times clone and a Helvetica bold clone for the section titles. Use at most three levels of hierarchy in your text: section, subsection and subsubsection. The lower section numbering must use the section number of the higher section. For example, section number 2.1.3 refers to section 2, subsection 1 and subsubsection 3.

C   Reference and in-text citation guidelines

With regard to standard practice across academic disciplines, it is important to use an in-text citation to indicate when your work borrows or refers to words or ideas from a specific source. A full reference of this source should be included in your 'References' / 'Bibliography' / 'Works Cited' section as well.

There are two major referencing styles: the Harvard style and the Vancouver style. The former was first introduced in a journal article by Professor of Zoology Edward Marks in 1881 during his tenure at Harvard University (Chernin, 1988, 1062). It has since become an umbrella term to refer to styles which utilise author-date (e.g., APA style) or author-page (e.g., MLA style) within parentheses (). It still remains in use to some degree in the natural sciences (e.g., American Chemical Society, 2006) and has become the predominant style within the social sciences (e.g., American Psychological Association, 2010), and the arts and humanities (e.g., Modern Languages Association, 2016; University of Chicago Press, 2017). In contrast, the latter Vancouver style has become a very common style within the fields of engineering, technology and science. The name is derived from the inaugural meeting in Vancouver, Canada in 1978 of a committee later referred to as the International Committee of Medical Journal Editors (ICMJE) (BMA, 2012). The style is typified by its use of numbers for in-text citations; as such, it is also known as the author-number system. One of the most common versions of the Vancouver Style is the IEEE Reference Guide (IEEE 2018). The numbers used for in-text citations correspond to a numbered reference list at the end of the text. Typically, the numbers refer to the order in which the referenced authors first appear in the text. Thereafter, they continue to be referred to by this number. A lesser utilised variation of this style can feature the numbers corresponding to an alphabetical list of authors.

There is a third system of notes and bibliography. This is one of the two styles in the Chicago Manual of Style; the other is an author-date style. Notes and bibliography is mostly used by the fields of literature and history. It is occasionally utilised in the field of arts, but it is considerably less common than the Harvard style. Among the 97 English doctoral dissertations available online in AaltoDoc as of 5 March 2020, only 14 utilised this method. As it is relatively rare within Aalto University, it is recommended to visit the Chicago Manual of Style online for the reference style guide for it.

The two key considerations in adopting a particular style are that you agree with your supervisor on the particular style at the start of your thesis writing process and that you remain consistent in the use of the style throughout your thesis. Your supervisor can provide recommendations for you on the common style guides generally used in your field.

Below, you will find general guidelines to these two major styles with reference to the specific style guides in use. Corresponding examples are provided in the boxes below. Please note that the use of boxes here is only to visually separate the examples for clarity's sake and that boxes should not be used when quoting or paraphrasing in your own thesis.

Direct quotes

If you are using the exact wording from a source, you should place this in double quotation marks (" ") followed by the author, date, and page number in parentheses (if using the Harvard style) or a number in square brackets [if using the Vancouver style]. The citation should be on the same line and inside the punctuation. This citation should correspond to a full reference in the References / Bibliography / Works Cited section of your thesis. The punctuation used between the author, date, and page number may vary depending on which style guide you are using (compare, for example, The Chicago Manual of Style (University of Chicago Press, 2017), The Publication Manual of the American Psychological Association (APA, 2018), The MLA Handbook (MLA, 2016) and The New Oxford Style Manual (OUP, 2016)). However, it should be noted that the frequency of direct quote usage varies significantly between academic fields. While it is a fairly common practice to utilise quotes within art, design, architecture and business, it is rare in the science and engineering fields. Therefore, it is recommended that you ask your supervisor or check well-known journals in your field to gauge the frequency of quoting preferred by experts in your field.

Quoting in Harvard style

Example 1: Quote, information-prominent, American Sociological Association (2019) style

> Attending to these breakdowns not only result in an on-going re-constitution of relations between people and things but are also hotbeds for unleashing everyday "creativity, invention, imagination, and artfulness" (Jackson, 2014: 226).

Source: Durrani, M. 2018. Designers by any other name: exploring the sociomaterial practices of vernacular garment menders. *Design Research Society International Conference: Catalyst. DRS International Conference Series.* 4: 1731-1746. ISBN 978-1-912294-19-0 (electronic). DOI: 10.21606/dma.2018.495. © 2018 Design Research Society. This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. https://creativecommons.org/licenses/by-nc-sa/4.0/.

Example 2: Quote, author-prominent, American Psychological Association (2010) style

> Philosopher Mark Johnson (2007) argued that meanings emerge from "deeper explorations into the qualities, feelings, emotions, and bodily processes" (p. x).

Source: Aktaş, B. & Mäkelä, M. (2019). Negotiation between the maker and material: Observations on material interactions in felting studio. *International Journal of Design*, 13(2): 55–67. *© 2019 Aktaş & Mäkelä. Copyright for this article is retained by the authors, with first publication rights granted to the International Journal of Design. All journal content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License.*

Block quoting

Block quotes are used for longer quotes. Note the line indentation on each line of the block quote below, starting from "There was . . . ", and the absence of quotation marks. Each style guide recommends its own minimum text length before using block quoting. [Compare: AMA - four lines of text or more; APA - 40 words or more; or Chicago - 100 words or more.]

Example 3: Block quote, Harvard style, as recommended by journal

> When the Center for Bits and Atoms won the National Science Foundation Grant in 2003, MIT engineers began to look for local communities around the world they could help via digital fabrication: "Instead of bringing information technology to the masses, the fab labs bring information technology development to the masses," explained Gershenfeld, in the official press release (NSF 2004). Karlsen had a more colourful version:
>
> > There was an innovation competition launched by MIT globally to develop local projects. MIT sent some of its best teachers to Norway to find a suitable cooperation project. They found us through Telenor, who told them: 'There is this crazy guy lost in the fjord who devised sensors for his animals.' We enjoyed a great year of cooperation with MIT in 2001 and we were invited to Boston to present and develop this project.

Source: Kohtala, C & Bosqué, C. (2014). The Story of MIT-Fablab Norway: Community Embedding of Peer Production. Journal of Peer Production, 5 (8): 1–8. ISSN 2213-5316 (electronic). © 2014 public domain. http://peerproduction.net/issues/issue-5-shared-machine-shops/peer-reviewed-articles/the-story-of-mit-fablab-norway-community-embedding-of-peer-production/

Paraphrasing

Paraphrasing is a more prevalent citation practice in many academic fields. In engineering and science, paraphrasing tends to dominate with quotations used only sparingly. In other fields, such as art and design, the ratio of quoting to paraphrasing varies significantly. Again, it is best to ask your supervisor or review journals from your field to see how common the practice is amongst your academic peers.

The intent behind paraphrasing is that you write the ideas or arguments of a source in your own words. This citation practice can allow for better integration of ideas, argumentation, and flow within the text. A good rule of thumb for paraphrasing is to have more than 80% of the paraphrased text in your own words. If you only change a few words from the original, you can run the risk of plagiarism, even if you cite the source. Words matter. If you use the exact combination of words from another author, these must be in quotes.

Paraphrasing in Harvard style

Example 4: Paraphrase, author prominent, Chicago Manual of Style (2017) style

> von Hippel (1986) suggested a four-step process for working with lead users: first identifying important trends and key customer needs, then identifying lead users and understanding their needs and possible solutions and finally working with lead users in order to improve or generate product/service concepts.

Source: Hyysalo, S., Kohtala, C., Helminen, P., Mäkinen, S., Miettinen, V., & Muurinen, L. (2014). Collaborative futuring with and by makers. *CoDesign*, 10(3–4), 209–228. DOI: 10.1080/15710882.2014.983937. *© 2014 The Authors. This is an Open Access article. Non-commercial reuse, distribution, and reproduction in any medium, provided the original work is properly attributed, cited, and is not altered, transformed, or built upon in any way, is permitted.*

Example 5: Paraphrase, information-prominent, Harvard style as recommended by journal

> Obviously digital technologies will not destroy comics as we know them, but they may change their underlying decorum. In reality, these changes have continuously shaped the lives of the industry's amateurs and semi-professionals, who have to organize their time around a bricolage of fragmented schedules and poorly paid work (Woo 2015): from daily feeding a Patreon account while filling a scanlation request, to selling a print in Deviantart while reviewing the latest Doujinshi on a not-so-free-of-ads-blog are some of the patchwork tasks of the comics networked precariat in the age of semio-capitalism.

Source: Manouach, I. (2019). Peanuts minus Schulz: Distributed Labor as a Compositional Practice. The Comics Grid: Journal of comics scholarship, 9(16), 1—21. https://doi.org/10.16995/cg.139 *© 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See* http://creativecommons.org/licenses/by/4.0/.

Paraphrasing in Vancouver style

Example 6: Paraphrase, information prominent, IEEE (2018) style

> When a laser beam is scattered by a dielectric microparticle, resulting in light refraction on entering and leaving the particle, a small amount of momentum is transferred from the photons to the matter. This change in momentum, known as the gradient force, results in the attraction of the particle to the high intensity part of the beam (usually the centre). Optical trapping of microscale particles via this mechanism was first reported in the 1970s [1] and duly led to the initial observation of a single beam optical trap in 1986 [2]. These preliminary experiments, and many of the methodologies that developed from them, utilized the gradient force

exerted by a single, tightly focused Gaussian laser beam to trap particles in solution through what has become known as the "optical tweezer" effect. Since these initial findings, optical technology has evolved significantly, and traps that facilitate three dimensional manipulation of particles are now readily available. While originally limited to the controlled manipulation of individual particles, multitrap setups involving either splitting [3,4] or time sharing [5,6] with a single laser beam are now also commonly utilized. As a more advanced form of the former, holographic optical tweezers that employ diffractive optical elements such as spatial light modulators now allow computer controlled, independent manipulation of multiple particles [7–9]. A number of multitrap devices have also been developed based on the application of laser beams with more complex phase and intensity profiles, as for example Bessel or higher order Laguerre Gaussian beams [10–12].

Source: "Chirality in Optical Trapping and Optical Binding" by David S. Bradshaw, Kayn A. Forbes, Jamie M. Leeder, and David L. Andrews in *Photonics* 2015, available under a Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/) at https://doi.org/10.3390/photonics2020483.

If you would like to emphasise the inventor, you could rephrase the first in-text citation as "was first reported by Ashkin in 1970 [1]" or "was first reported by Ashkin [1]" if the year is less important. (Ashkin is the sole author of the paper. It was published in 1970.)

Tips for paraphrasing

1. Identify the important points from the source text. Then, try to identify the relationship between these different parts. Is the relationship sequential, causal, contrasting, or conditional? Can these be replaced by synonyms? For example, the contrasting conjunction, but, can be replaced by however, although, nevertheless, yet or on the other hand. In many cases, these synonyms may require that you change the structure of the sentence, which in turn may help you to formulate the ideas in your own words.

2. Synonyms - A word such as give can be replaced with provide, supply or contribute. Refer to a thesaurus for more examples. You can find these online or at the Learning Centre. If you are unsure of how a new word might be used, check a dictionary or search the word in Google Scholar for examples of its use in context.

3. Common phrases - There are often multiple ways of expressing a given phrase in academic writing. For example, compare "Previous studies have not dealt with. . . " with "Researchers have not treated X in much detail" or "Most studies in the field of X have only focused on. . . " These examples were drawn from the University of Manchester's (2018) Academic Phrasebank

(http://www.phrasebank.manchester.ac.uk/). This corpus of common phrases includes hundreds of examples, categorised according to their functions.

4. Additions or deletions - Can you add a missing item? Can you leave something out?

5. Change the structure of the sentence - There are many ways to change the structure. Can you change the sentence from passive voice to active voice, or vice versa? Can you rewrite the sentence with "It is" or "There is"? Can you switch to the inanimate agent, e.g. "This thesis discovered"? Refer to the Aalto University Language Centre site (http://sana.aalto.fi/awe/) in the "Cohesion" section for more ways to change the structure of a sentence.

Example 7: Before and After Paraphrase

Before Paraphrase

"Significant progress has been made in the use of smart textiles in wearable technology, especially in the sport and well-being sector. However, the medical sector still lacks commercial and viable solutions" (Ilen et al., 2019, p. 2).

After Paraphrase

While some industries have taken advantage of smart textiles in wearable technology, Ilen et al. (2019, p. 2) contend that the medical industry has yet to produce any viable products.

Source: Ilen, E., Groth, G., Ahola, M., & Niinimäki, K. (2019). *Empathy in a Technology Driven Design Process: Designing for Users without a Voice of their Own.* Paper presented at 8th biannual Nordic Design Research Conference: Nordes 2019: Who Cares?, Espoo, Finland. Open Access paper.

Avoiding the pitfalls of the Finnish paragraph-length paraphrase (FPP)

In Finland, there exists a method of citation that is not permitted or recognised by many international style guides, including IEEE Reference Guide (IEEE 2018); Information and documentation — Guidelines for bibliographic references and citations to information resources (ISO 690:2010(E)) (ISO 2010); New Oxford Style Manual (OUP 2012); Scientific Style and Format: The CSE Manual for Authors, Editors, and Publishers (CSE 2012); The Chicago Manual of Style (University of Chicago Press 2017); The Publication Manual of the APA (APA 2018). Due to this and the reasons discussed below, it is not recommended that this method be used when writing in English. Interestingly, the Finnish Standards Association (Suomen Standardisoimisliitto SFS ry) which emulates the ISO standard, does not recognise this either (FSA 2010). In the Finnish style of citation, writers can write an entire paragraph where every sentence comes from a single source. Then, at the bottom of the paragraph, they provide a citation in brackets outside the final punctuation (see example below). Generally, the Finnish method suggests that by putting the citation after the period, the citation then refers to all the sentences preceding it in the paragraph.

Example 8: Paraphrase, Finnish style, not recommended in English

> Additive manufacturing was originally developed to guide product design by providing a way to create prototypes directly from digital designs. This method called rapid prototyping (RP), as the name implies, consumes less time and resources than most preceding techniques. For instance, the manufacturing of an injection mold for prototyping purposes would be extremely expensive. However, the part can be created with additive manufacturing for a fraction of the cost. Moreover, rapid prototyping is cost and time effective when it can substitute handcrafting, CNC manufacturing, or silicon molding. The downside when compared to these methods is often poor surface quality and inferior dimensional accuracy. However, RP enables fast iterative testing of products with a low threshold of prototypes failing expectations. This makes it a superior tool in product development and explains why prototyping has been the leading application of AM. (Wohlers 2013)

Source: Anonymous. (2015). Bachelor Thesis. Adapted from Anonymous, Bachelor thesis, School of Engineering, Aalto University, Espoo Finland, 2015. This work is licensed under an Attribution-NonCommercialNoDerivatives 4.0 International license.

This style has several inherent problems. To begin with, it significantly limits the writer from utilising what the noted academic writing scholar Ken Hyland (2005) has called stance and engagement within the actual paragraph. This is where the writer adds their own authorial voice to the text by interacting with the source material and with the readership in order to situate themselves in their community of academic peers. This interaction is achieved in the context of stance and engagement by the use of attitudinal markers, boosting, hedging, self-mentioning, directives, personal asides, questions, reader pronouns, and shared knowledge (Hyland 2005, 177). Within this long form Finnish paraphrase, the stance and engagement can logically be attributed only to the original cited source, not the current writer summarising the ideas of the source, for there is no meaningful way to discern ownership of ideas beyond the citation attribution. In example 8 above, it is possible to identify at least 10 separate statements made in the paragraph. Some of these are representative of stance, including claims (e.g., 'The downside when compared to these methods is. . . ' and 'This makes it a superior tool. . . and explains why prototyping has been the leading application in AM') and logical connections ('Moreover', and 'However'). Moreover, there are clear attempts at engagement with asides (e.g., 'as the names implies') and creating interpersonal solidarity with experts in the field through boosting (e.g., 'extremely' and 'superior'). While some of these may have been the thesis writer's own arguments, it would, for the reader, appear that these have all been paraphrased from the source (Wohlers 2013). From a larger perspective, this Finnish paraphrase style tends to promote a less critical style of writing where writers tend to summarise the work of others and avoid critically discussing individual ideas from sources as they are introduced.

Similarly, this Finnish paraphrase style can lead fledgling writers quite easily into practices that constitute plagiarism. For instance, when students write their bachelor's thesis, it can often take the form of a literature review. When the perceived goal is to

summarise others' ideas, it can readily lead to an overuse of this FPP style. According to the Chicago Manual of Style, this could constitute plagiarism: '(u)se that is not fair will not be excused by paraphrasing. Traditional copyright doctrine treats extensive paraphrase as merely disguised copying' (University of Chicago Press 2017, 212). It is similar to the problem of overuse of block quotes in the arts, humanities, and social sciences.

Worse yet, the logic of the FPP style has led some students to the practice of summarising two or more sources into a single paragraph with the citations provided only at the end of the paragraph. This would certainly constitute plagiarism as there is no meaningful way for the reader to discern which idea or argument belongs to which author. Unfortunately, this logical leap does not appear to be a rare occurrence. In a recent survey of 41 theses written in English at Aalto University between 2008 and 2018, 85% (35 samples) used the FPP style, while the multiple citation problem occurred in 41% (17 samples) (Forget and Paloposki 2019).

While it is generally better to introduce a citation and immediately contextualise it (i.e., integrate it into your argumentation; explain its significance; agree/disagree with it; or contrast it), it is possible to continue discussing a single source without using the same citation at the end of every sentence. For instance, the Publication Manual of the APA (APA 2018) recommends maintaining a clear progression of the topic at the start of the next sentence.

Example 9: APA-recommended method of paraphrasing same source across consecutive sentences

> Chen and Liu (2004) studied the effect of aggregate size distributions and the volume fraction of aggregate on the fracture parameters of concretes with strength 50–89 MPa under three-point bending test. For this purpose three various maximum aggregate sizes of 10, 15 and 20 mm were employed.

Source: Rashad, A. and Seleem, H. (2014). A Study of High Strength Concrete with Moderate Cement Content Incorporating Limestone Powder. *Building Re-search Journal*, 61(1): 43–58. DOI https://doi.org/10.2478/brj-2014-0004. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 International license.

In example 9, it is clear the Rashad and Seleem are still referring to Chen and Liu's article in the second sentence by starting with 'For this purpose... aggregate sizes...' Although one would suffice, both noun phrases are clearly connected by topical progression to the narrative of the first sentence.

References

References in the Harvard Style

The reference list below contains examples of a scholarly article, book, chapter in a book with editors, conference publication, doctoral dissertation, interview, master's thesis, motion picture, painting, photograph, standard and webpage. These have been

formatted in the American Psychological Association (APA) style, one variation of the Harvard Style. Please note that the APA recommends including the digital object identifier (DOI) for sources found online. When the DOI is missing, they recommend including the home page URL. Although the APA guidelines do not require a date of retrieval, it is good practice to note the date retrieved as online sources can change over time.

American Medical Association. (2007). *AMA Manual of Style: A Guide for Authors and Editors* (10th ed.). New York, USA: Oxford University Press.

American Psychological Association. (2018). *Publication Manual of the American Psychological Association* (6th ed.). Washington, USA: American Psychological Association.

American Sociological Association. (2019). *American Sociological Association Style Guide* (6th ed.). Washington, USA: American Sociological Association.

British Medical Association. (2012). Reference Styles [Webpage]. Updated 28 February 2019. Retrieved 01 March 2020 from `https://www.bma.org.uk/library/library-guide/reference-styles`

Bruce, E. & Hamp-Lyons, L. (2013). Looking for the academic voice: Assessing undergraduate writing. In J. Wrigglesworth (Ed.), *EAP within the higher education garden: Cross-pollination between disciplines, departments and research*. Proceedings of the BALEAP Conference, Portsmouth, UK, 2011. Reading, UK: Garnet Education.

Chernin, E. (1988). The "Harvard System": a mystery dispelled, *BMJ*, 297(6655): 1062–1063.

Council of Science Editors. (2012). *Scientific Style and Format: The CSE Manual for Authors, Editors, and Publishers*. Chicago, USA: University of Chicago Press.

Finnish Standards Association. (2010). SFS 5989, Guidelines for bibliographic references and citations to information sources. Helsinki, Finland: Finnish Standards Association.

Forget, M. and Paloposki, T. (2019). *When academic writing cultures collide: Plagiarism requirements in the English Thesis Seminar at Aalto University*. Paper presented at 3rd International Seminar English as a Medium of Instruction (EMI): embracing pluricultural education, Valencia, Spain.

Halliday, M. (2013). *Halliday's Introduction to Functional Grammar* (4th ed.). London, UK: Routledge.

Heo, M. and Lee, M. (Producers), & Joon-Ho, B. (Director). (2019). *Parasite* [Motion Picture]. South Korea: CJ Entertainment.

Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse, *Discourse Studies*, 7(2), 173–192. `doi.org/10.1177/1461445605050365`

Ilen, E., Groth, G., Ahola, M., & Niinimäki, K. (2019). *Empathy in a Technology Driven Design Process: Designing for Users without a Voice of their Own*. Paper

presented at 8th biannual Nordic Design Research Conference: Nordes 2019: Who Cares?, Espoo, Finland.

International Organization for Standardization. (2010). *Information and documentation – Guidelines for bibliographic references and citations to information resources* (ISO 690:2010(E)) (3rd ed.). Retrieved 01 March 2020 from `https://www.iso.org/standard/43320.html`

Lu, Y. (2018). *Experience goals in designing professional tools: evoking meaningful experiences at work* (Doctoral Dissertation, Aalto University, Espoo, Finland). Retrieved 1 March 2020 from `https://aaltodoc.aalto.fi/handle/123456789/34084` Modern Language Association. (2016). MLA Handbook. New York, USA: Modern Language Association.

Nixon, R. (1977, May 4). Interview by D. Frost (Video recording). David Paradine Productions Ltd., Hertfordshire, U.K.

Oxford University Press. (2012). *New Oxford Style Manual*. Oxford, UK: Oxford University Press.

Patrias, K. (2007). *Citing medicine: the NLM style guide for authors, editors, and publishers* (2nd ed.). Bethesda, USA: National Library of Medicine (US). Retrieved 1 March 2020 from: `http://www.nlm.nih.gov/citingmedicine`

Sutherland-Smith, W. (2019). Is student plagiarism still a serious problem in universities today? In D. Pecorari & P. Shaw (Eds.), *Student Plagiarism in Higher Education* (pp. 47–61). Oxford, UK: Routledge.

Tutal, E. (2015). *Participatory design of visual product identity concepts* (Master's thesis, Aalto University, Espoo, Finland). Retrieved 1 March 2020 from `https://aaltodoc.aalto.fi/`

Unknown. (1904). *Alexander Graham Bell* [Photograph]. Library of Congress Prints and Photographs Division, Washington, U.S.A.

University of Chicago Press. (2017). *The Chicago Manual of Style* (17th ed.). Chicago, USA: University of Chicago Press.

University of Manchester. (2018). Academic Phrasebank [Webpage]. Retrieved 1 March 2020 from `http://www.phrasebank.manchester.ac.uk/`

Van Gogh, V. (1889). *Sunflowers* [Oil on canvas]. Van Gogh Museum, Vincent Van Gogh Foundation, Amsterdam, The Netherlands.

References in the Vancouver (IEEE) style

The reference list below contains examples of scholarly articles [1, 2], a book [3], a chapter in a book with editors [4], a conference publication [5], a master's thesis [6], a doctoral dissertation [7], a standard [8] and a webpage [9]. Note that the IEEE style allows DOIs, but does not require them. For online references, you should specify the 'accessed date' if it is a webpage that might change. Scholarly articles and theses should not change, and thus the access date is not important for such references. Today,

some periodicals use article numbers [2] instead of ordinary page numbers [1]. Please see [10] for more details.

Referring to interviews or works of art seems very uncommon using the IEEE style and the guide [10] does not mention such references at all.

[1] J. B. Pendry, "Negative refraction makes a perfect lens," *Phys. Rev. Lett.*, vol. 85, no. 18, pp. 3966–3969, Oct. 2000, doi: 10.1103/PhysRevLett.85.3966.

[2] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, "Equivalence of restricted Boltzmann machines and tensor network states," *Phys. Rev. B*, vol. 97, no. 8, 2018, Art. no. 085104, doi: 10.1103/PhysRevB.97.085104.

[3] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Weinheim, Germany: Wiley-VCH, 2004.

[4] V. Yannopapas, A. G. Vanakaras, and D. J. Photinos, "Electrodynamic theory of three-dimensional metamaterials of hierarchically organized nanoparticles," in *Amorphous Nanophotonics*, C. Rockstuhl and T. Scharf, Eds., Berlin Heidelberg, Germany: Springer, 2013, pp. 119–141.

[5] T. Joachims, "Optimizing search engines using clickthrough data," in *Proc. 8th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Edmonton, Canada, Jul. 23–26, 2002, pp. 133–142.

[6] J. Martela, "Lifecycle of Mobile Phones," M.Sc. thesis, Dept. Materials Science and Engineering, Aalto University, Espoo, Finland, 2019. [Online]. Available: http://urn.fi/URN:NBN:fi:aalto-201908254898

[7] R.J. Garbacz, "A generalized expansion for radiated and scattered fields," Ph.D. dissertation, ElectroScience Lab., Ohio State Univ., USA, 1968. [Online]. Available: http://rave.ohiolink.edu/etdc/view?acc_num=osu1302723653

[8] *Simple Mail Transfer Protocol*, RFC 5321, J. Klensin, Oct. 2008, [Online]. Available: https://tools.ietf.org/html/rfc5321

[9] B. Casselman, "Jacob Bernoulli's zoo," AMS feature column, http://www.ams.org/publicoutreach/feature-column/fc-2018-02 (accessed Feb. 6, 2018).

[10] IEEE, "IEEE Reference Guide," 2018. [Online]. Available: https://ieeeauthorcenter.ieee.org/wp-content/uploads/IEEE-Reference-Guide.pdf