PLAGIARISM

Decision Making & Dealing with Grey Zones across Academic Fields

Presented by Michael Seadle, Thorsten Beck & Melanie Rügenhagen

25 June 2020
Presenters

Prof. Dr. Michael Seadle
Director
Principal Investigator

Dr. Thorsten Beck
Researcher

Melanie Rügenhagen
Researcher
Doctoral Candidate
Introductions

Background image by Gerd Altmann from Pixabay
Table of Contents

Introductions

Part 1: Introduction to Plagiarism  // Melanie Rügenhagen
Part 2: Image Plagiarism  // Thorsten Beck
Part 3: Detection and Judgement  // Michael Seadle

Summary
Part 1: Introduction to Plagiarism

// Melanie Rügenhagen
What is plagiarism?

“Plagiarism is a problem that comes from a lack of creativity—copying the expression of others [...].”
(Seadle 2017, p. 43)

“Plagiarism is hard to define in any precise way. At the simplest level it is copying texts from other authors and using them as if they were their own.”
(Seadle 2017, p. 16)
Part 1: Introduction to Plagiarism
Part 1: Introduction to Plagiarism

“Making a video using footage from others’ videos or using copyrighted music as part of the soundtrack.”

“Performing another person’s copyrighted music (i.e., playing a cover).”

“Composing a piece of music that borrows heavily from another composition.”
Part 1: Introduction to Plagiarism

What else could be plagiarized?

“Ideas”

“Data”

“Computer programming”

“Products of any other creative endeavour”

Source: [University of Guelph, n.d., b]
Part 1: Introduction to Plagiarism

Plagiarism is a hard accusation that may or may not be true.
What to look for

matching words

determining the intellectual similarity between two texts

identifying whether words or phrases are very common in some context or even across contexts

discipline-based customs of writing
What is considered to be plagiarism?

The VroniPlag Wiki team distinguishes between four major categories of plagiarism:

- “complete plagiarism”
- “cover-up”
- “pawn sacrifice”
- “translation plagiarism”
What do the VroniPlag Wiki categories mean?

- Complete plagiarism: “the text in question is completely or almost unaltered compared to a text found in a source that was created before the text in question without making this obvious”
  (this quotation is translated from the German original: “Text, der sich vollständig oder nahezu unverändert in einer anderen (früher datierten) Quelle findet, ohne dass dies kenntlich gemacht ist.”)

- A cover-up is an “instance where text deviates from text in an earlier source to the extent that the text uses different words to the same effect or partly the same words without making this obvious”
  (this quotation again is translated from the German original: “Text, der sich mit Abweichungen sinngemäß und/oder teilweise wörtlich in einer anderen (früher datierten) Quelle findet, ohne dass dies kenntlich gemacht ist.”)
What do the VroniPlag Wiki categories mean?

- The very brief description of a pawn sacrifice is “insufficient reference”. For example, this would include giving a perfectly correct reference for a quoted passage, but then continuing the quotation or paraphrasing without making that explicit.
  (this is translated from the German original on the VroniPlag Wiki website: see Wikia, Inc. 2019)

- Translation plagiarism includes “translating a text literally or using words to the same effect” without giving the proper reference. The VroniPlag Wiki team acknowledge on their website that there are often no exact translations, which means translation involves interpretation.
  (this is translated from the German original on the VroniPlag Wiki website: see Wikia, Inc. 2019)
What is considered to be plagiarism?

- “Not Crediting a Source”
- “Paraphrasing too Closely, even if you do Credit the Source”
- “Using statistics from someone else's work without crediting the source”
- “Using images, photographs, maps or other illustrative devices without citing sources”
- “Buying or copying another person's paper or assignment or hiring an essay-writing service to write your paper”
- “Sharing code for a computer program with another student or taking code from the Web or another source without citing that source”
- “Downloading material from the Internet without proper citation”
- “Padding a Bibliography or Reference List, to suggest that you have done research when you haven't”
- “Making up statistics or other important facts, and citing a false source”

Source: [University of Guelph, n.d., a]
What is originality?
What is originality?

“The ability to think independently and creatively”

(Lexico.com, n.d.)
Part 1: Introduction to Plagiarism

What is not plagiarism?
Part 2: Image Plagiarism  // Thorsten Beck
Part 2: Image Plagiarism

How is Image Plagiarism defined?

“Plagiarism is a form of academic dishonesty and may include, but is not limited to, the following: (...) Presenting text, digital work (e.g. computer code or programs), music, **video recordings or images copied with only minor changes from sources** such as the internet, books, journals or any other media, **without due acknowledgement**; (...)“

(Collery, 2020, emphasis added by Thorsten Beck)
Part 2: Image Plagiarism

Terminology & Types

- Duplication (Self/Other)
- Reuse/Recycling
- Copy-Paste
- Copy-Move

Example: Hank Willis Thomas ./. Graeme Williams

Hank Willis Thomas (2018): Greyscale reproduction of Graeme Williams’ original photograph (1990): “Police watch an ANC rally while children taunt them”.

Illustration, TB (2020) -
How to Detect Plagiarized Images?

- Visual Inspection
- Reverse Image Search
- TinEye
- Other
Part 2: Image Plagiarism

How to Avoid Image Plagiarism?

- Correct citation
- Correct attribution
- Manage copyrights
Part 3: Detection and Judgement

Approaches to detection

**iThenticate - Automated detection** ([http://www.ithenticate.com/](http://www.ithenticate.com/))
“ [...] the most trusted plagiarism checker by the world’s top researchers, publishers, and scholars.” (Turnitin, LLC., n.d.)
- Matching words, interpretation is manual labor

**VroniPlag - Detection by crowdsourcing** ([https://vroniplag.wikia.org/de/wiki/Home](https://vroniplag.wikia.org/de/wiki/Home))
Website to detect plagiarism in German academic publications and theses
- Percentages can appear misleading, does not consider context
### Plagiarism greyscale metrics for texts with a reference

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Label</th>
<th>Valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing quotation marks</td>
<td>A correct reference is there, but the quotation marks are missing</td>
<td>MQ</td>
<td>Bad practice, but no deception</td>
</tr>
<tr>
<td>Incomplete reference</td>
<td>A partial reference is there, but quotation marks are missing</td>
<td>IR</td>
<td>Bad practice, probably no deception</td>
</tr>
<tr>
<td>Poor reference</td>
<td>An inaccurate reference is there, with or without quotation marks</td>
<td>PR</td>
<td>Bad practice, potential deception</td>
</tr>
</tbody>
</table>

---

2 This table is an extract from Seadle, 2017, p.39, with minor changes made for this exhibition.
**HEADT Centre Greyscale**

### Plagiarism greyscale metrics for texts without a reference:

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Label</th>
<th>Valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Significant matches</td>
<td>Multiple identical sentences in a paragraph</td>
<td>C1</td>
<td>Probable plagiarism</td>
</tr>
<tr>
<td>Some matches</td>
<td>$\geq 5$ contiguous words within a sentence of $\geq 5$ words</td>
<td>C2</td>
<td>Potential plagiarism</td>
</tr>
<tr>
<td>Near copying 1</td>
<td>Multiple ($\geq 3$) exact phrases ($\geq 3$ words) overlap in contiguous sentences in a paragraph</td>
<td>C3</td>
<td>Potential plagiarism</td>
</tr>
<tr>
<td>Near copying 2</td>
<td>Fewer than 5 contiguous words within a sentence of more than 5 words</td>
<td>C4</td>
<td>Potential plagiarism</td>
</tr>
<tr>
<td>Similarity 1</td>
<td>Several ($&lt;3$) exact phrases ($\geq 3$ words) overlap in contiguous sentences in a paragraph or more than 9 words in sequence</td>
<td>S1</td>
<td>Questionable practice</td>
</tr>
<tr>
<td>Similarity 2</td>
<td>Many ($\geq 5$) of the exact same words (excluding function words) in contiguous sentences in the same paragraph</td>
<td>S2</td>
<td>Questionable practice</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>----</td>
<td>----------------------</td>
</tr>
<tr>
<td>Topic overlap</td>
<td>Topic overlap with facts and standard phrases</td>
<td>S3</td>
<td>Probably not plagiarism</td>
</tr>
<tr>
<td>Standard words or phrases</td>
<td>Widely used words or phrases in this context</td>
<td>SW</td>
<td>Not plagiarism</td>
</tr>
<tr>
<td>General knowledge/facts</td>
<td>Information so broadly known that formulations cannot be considered unique</td>
<td>GK</td>
<td>Not plagiarism</td>
</tr>
<tr>
<td>Implied quotation</td>
<td>Something such as an indentation or colon implies that the passage was quoted, but no reference is supplied</td>
<td>IQ</td>
<td>Matching without an intent to deceive</td>
</tr>
</tbody>
</table>

1 This table is an extract from Seadle, 2017, S. 38, with minor changes made for this exhibition.
Examples - Case 1

Dissertation by Karl-Theodor Freiherr zu Guttenberg

1218 Plagiatsfragmente aus 135 Quellen
auf 371 von 393 Seiten (94.4%)
in 10421 plagiierten Zeilen (63.8%)

Stand: 03.04.2011 11:55 Uhr
Examples - Case 2

In a German dissertation by an author whose identity should be protected, one of the accusations was of a passage that ran in my translation:

*Mecklenburg-Vorpommern is a state in the north east of Germany.*

It was certainly true that others had used the same words, but since the sentence is a statement of fact, it is unreasonable to say it has sufficient creativity to count as plagiarism. Facts are facts.
Potential questions for discussion

How does one detect the difference between bad practice and deception in plagiarism?

What are the best forms of measurement for determining when the amount of word and phrase copying becomes serious enough to be called plagiarism?

To what degree is plagiarism an economic issue (theft of intellectual property) and to what degree an ethical one?
References


Slide 2: Greyscale changes to original photos by Thorsten Beck (2020)

Slide 3: Background image by Gerd Altmann from Pixabay

Slide 7: Photo on the left by Vova Krasilnikov from Pexels

Slide 7: Photo on the right by Donald Tong from Pexels

Slide 8: Screenshot of https://headt.eu/Satire with menu open to copy image, taken on 29/5/2020

Slide 10: Photo by K. Mitch Hodge on Unsplash


Slides 16-18: Images and background images by Melanie Rügenhagen (2020)


Thank you.

Ask your questions on:

Researcher Academy Mendeley group
Follow us on Twitter: @ResearcherAcad