HOW TO AVOID PLAGIARISM?

Tiina Niklander
AMICT 2011, Petrozavodsk
<table>
<thead>
<tr>
<th>Table of content</th>
</tr>
</thead>
<tbody>
<tr>
<td>➢ Plagiarism definition</td>
</tr>
<tr>
<td>➢ Copy detection tools</td>
</tr>
<tr>
<td>➢ Turnitin</td>
</tr>
<tr>
<td>➢ Tool overview</td>
</tr>
<tr>
<td>➢ Criticism</td>
</tr>
<tr>
<td>➢ Avoiding unintentional plagiarism</td>
</tr>
<tr>
<td>➢ Conclusion</td>
</tr>
</tbody>
</table>
Plagiarism definition

- "Plagiarism is the use of another’s thoughts, or work, without acknowledgement or permission. In plagiarism, one author takes another’s idea and presents it as his/her own.” (Smith & Wren 2010)

- “Plagiarism originated from the Latin word *plagiarius* meaning the theft of words as well as slaves” (Gu & Brooks, 2008)

- “The word plagiarism comes from Latin word *plagium* which means kidnapping” (Imran, 2010)


Types of plagiarism in research

- **Method**
  - how it is implemented

- **Cast**
  - the form of plagiarism

- **Purpose**
  - the motivation behind

N. Imran: Electronic Media, Creativity and Plagiarism. SIGCAS Computers and Society, Volume 40, No. 4, December 2010
Copy detection tools

- Automated tools to detect similarities in compared documents
- Statistical methods – most common
  - N-gram, Lancaster word pairs, word frequencies, compression metrics – fingerprints
  - Euclidean distance, cosine function, Jaccard measure, Dice measure
- Several tools available
  - De facto standard: turnitin.com
  - Others: Eve2, CopyCatchGold, WordCheck, Glatt, Moss, JPlag

Turnitin

- Web-based tool
- Own database of articles and web pages
- No ACM & IEEE? - at least claimed in 2008 by Kaner and Fiedler
- "Ignoring the commonly used words, the software looks for matches of strings of eight to ten words" (Jones, 2008)

Student submits her work, possibility for teacher submission
Can be integrated to course platforms like blackboard or moodle

Example course with two separate paper submission tasks and peer review of the first one
### Originality measure

- Traffic lights (blue, green, yellow, orange, red) + percentage
- How much of the submission matches with existing material

---

![Fig 2 from K. Jones: Practical Issues for Academics Using the Turnitin Plagiarism Detection Software.](image)

*International Conference on Computer Systems and Technologies - CompSysTech’08, ACM, 2008, IV.1-1 - IV.1-5*
Submitted text in the big picture, total originality value on top right, matched articles on the right with similarity percentage
Cut and paste plagiarism

Clearly fulfills the definitions of plagiarism: 

**direct copying and no citations.**

(colors show different sources)
Another limitation of CloneCloud is its inability to migrate native state to export unique native resources remotely. That's why it focuses on migration at execution points where no native state in the stack or the heap needs to be collected and migrated. CloneCloud neither virtualizes access to native resources that are not virtualized already or are not available on the cloud.

Despite of CloneCloud's limitations, Chun et al. noticed significant speed-ups when they tested their prototype with three different applications. Their prototype delivered up to 21.2x speed-up without programmer involvement.

B. Alfredo Giurgiu et al. [5] have also researched how to dynamically distribute applications between the cloud and mobile devices. Their approach has many similarities and differences with CloneCloud. They use Alfredo to distribute an application between a mobile phone and a server. It is based on OSGi, which has been used to decompose and loosely couple Java applications to software modules.

Giurgiu's et. al. approach does not yet do resource profiling automatically so it has to be done manually. They compose a...
Criticism about Turnitin

- Byrocratic decisions based on the originality percentage should not be done automatically using a boundary value for similarity
  - Asian students ’caught’ too easily
  - English-speaking students can rephrase better, but still copy idea
  - Student copying from single paper might remain under the limit
- Teaching is more important than punishment, because normally not intentional by the students

Avoiding unintentional plagiarism

- “Avoiding plagiarism does not need to be difficult or require an in-depth knowledge of copyright law.” (Smith&Wren 2010)

- Use proper referencing
- Paraphrase properly
- Summarize in own voice (with citation)
- Understand the source information (Imran 2010)


Explanations given by students

- Aspects and **behaviour** of students
  - bad **time** management,
  - personal **shortcomings when preparing assignments**,  
  - the elevated number of assignments to be handed in, etc.

- **Opportunities** conferred by ICT to locate, copy and paste information

- Aspects related to professors-lecturers and/or the characteristics of the **subject-course**
  - lecturers who show no **interest** in their work,
  - eminently **theoretical** subjects and assignments, etc.

---

Conclusion

- Common sense
- Do not steal ideas
- Give credit to the right persons and papers
- Tools help teachers in checking

"The very nature of research is to examine and build upon previous findings and to confirm, deny, or expand into new areas." (Smith & Wren 2010)
References