

WORKSHOP: INTERPRETING TEXT-MATCHING SOFTWARE SIMILARITY REPORTS

Dita Henek Dlabolová¹, Tomáš Foltýnek¹

¹*Mendel University in Brno, Czechia*

This is a follow-up workshop to one of the most successful educational activities of the European Network for Academic Integrity – the workshop “Where is the borderline of plagiarism” (ENAI, n.d.), which has been presented at numerous events with great success. Participants of the workshop often asked for a continuation, to deal with selected topics more in detail.

Interpreting text-matching software similarity reports is a challenging task requiring expertise and experience. Percentages presented by the system usually do not convey much useful information. Each report has to be carefully evaluated by humans. Their task is to distinguish false positives caused by random matches from possible seeds of disguised plagiarism, identify translation or paraphrase plagiarism, and spot other oddities that may indicate that some parts of the submitted document were taken from elsewhere. Although lots of universities use various text-matching software tools, only some of them organize sufficient training for their staff (Foltýnek and Glendinning, 2015). This workshop aims to fill this gap. The workshop is not specific to any text-matching software, it focuses on general features of disguised plagiarism and on how to discover it.

There are three learning outcomes of the workshop:

1. Understand the pros and cons of different institutional approaches to plagiarism cases
2. Learn more about the advantages and pitfalls of text-matching software tools
3. Be able to objectively judge a text-matching similarity report

The core activity of the workshop is a judgement of two cases of suspicion of plagiarism. The task of the participants will be to decide whether the given case is plagiarism or not by interpretation of a report from a text-matching software. Both cases

are prepared artificially but they are based on real student assignments. None of them can be decided at first sight, so their aim is to simulate real cases which a teacher might need to deal with when deciding about plagiarism. The reports do not come from any particular text-matching software. They were also created artificially based on features which are common in reports of the wide-spread text-matching tools, so that the activity is independent of software available at institutions of workshop participants.

Further, there will be a space for discussion and sharing experience.

The workshop outline is as follows:

- Definition of plagiarism
- Institutional processes dealing with plagiarism
 - Sharing examples of good practice
 - Pros and cons of particular scenarios
- Results of the project Testing of Support Tools for Plagiarism Detection (Foltýnek et al., 2020)
 - What text-matching tools are capable of?
 - What are their advantages and pitfalls?
- The core of the workshop: Interpreting the reports: Judgement of the cases
 - A pair of cases with the same percentage. Which of them is plagiarism and which is a false positive?
- Prevention of plagiarism
 - Addressing all sides of the Fraud triangle
 - Preventative measures substantiated by scientific literature

During the workshop, participants will be encouraged to share their personal experience not only with specific cases, but mainly with institutional processes, policies, and preventative measures. The discussion will be interlaced with the results of several international research projects.

REFERENCES

- FOLTÝNEK, T., and GLENDINNING, I. (2015). Impact of policies for plagiarism in higher education across Europe: Results of the project. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63(1), 207–216.
- FOLTÝNEK, T., DLABOLOVÁ, D., ANOHINA-NAUMECA, A., RAZI, S., KRAVJAR, J., KAMZOLA, L., ... WEBER-WULFF, D. (2020). Testing of support tools for plagiarism detection. *International Journal of Educational Technology in Higher Education*, 17(1), 46. <https://doi.org/10.1186/s41239-020-00192-4>
- ENAI (n.d.) *Where is the Borderline of Plagiarism?* Available from <http://www.academicintegrity.eu/wp/materials/where-is-the-borderline-of-plagiarism/>