

How AI-Generated Content Affects Students' Perceived Information Safety: An Exploratory Study

dr Paulina Motylińska

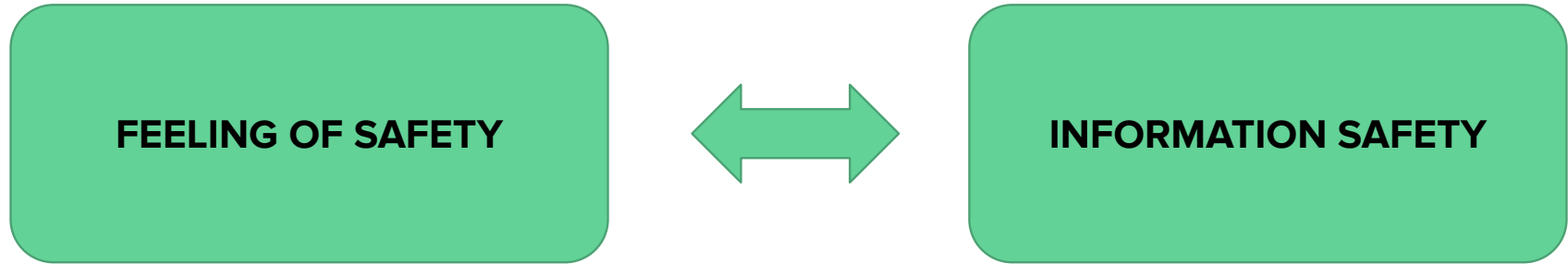
Institute of Security and Informatics, University of the National Education Commission, Krakow

mgr Anna Pieczka-Węgorkiewicz

Institute of History and Archival Studies, University of the National Education Commission, Krakow

Bratislava, 28.11.2025

Feeling of Information Safety



Feeling of Information Safety

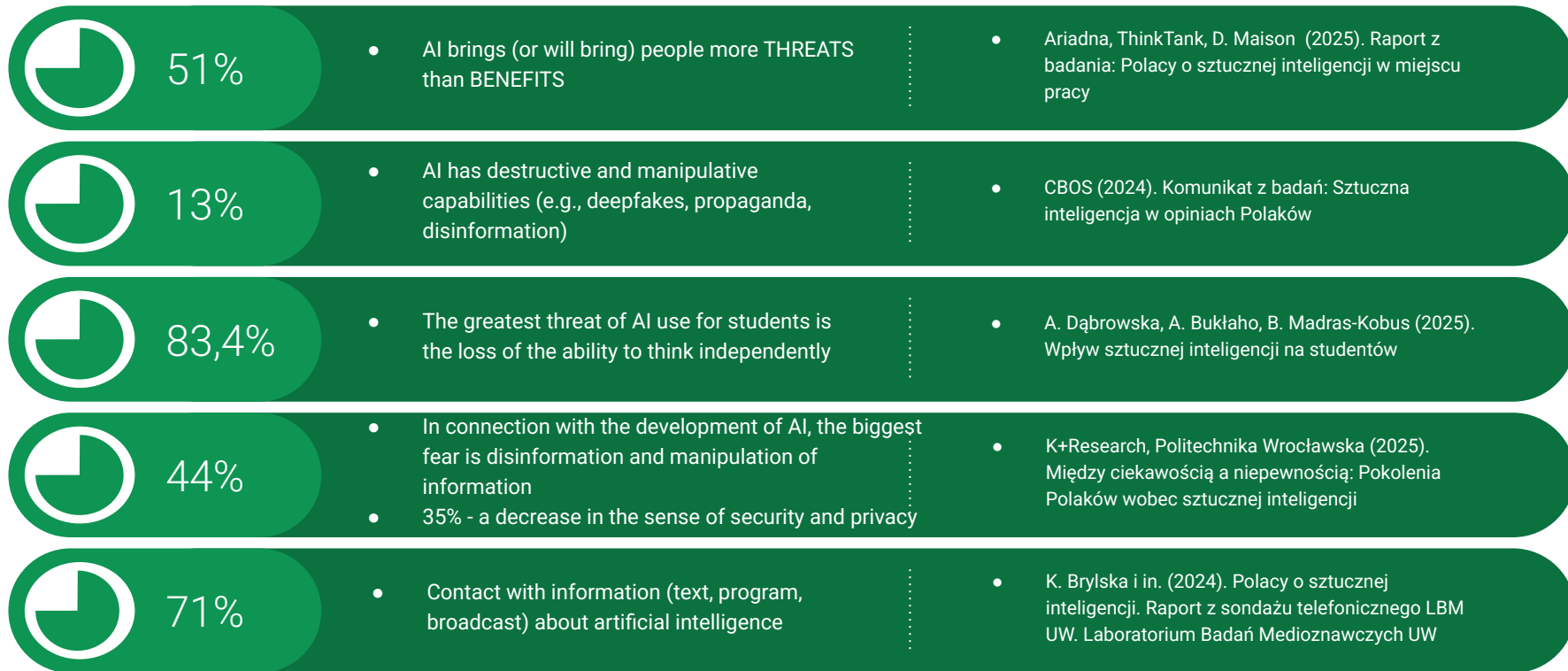
“A state in which an information user does not feel threatened by:

- (a) **exposure to low-quality information**, or
- (b) **the loss** of all or part of their stored **information resources**.

Instead, they experience **a sense of calm, safety, and satisfaction with their current level of information security and safety**, as well as **confidence in possessing the resources** (e.g., knowledge and skills for assessing information quality) necessary to take appropriate action in a crisis situation.”

FoIS as a research category -> e.g. Pieczka, Anna; Motylińska, Paulina (2021). Poczucie bezpieczeństwa informacyjnego. W: Horyzonty informacji 2, red. P. Korycińska. Kraków, Uniwersytet Jagielloński : Biblioteka Jagiellońska, s. 30-46.

AI and Users in Poland



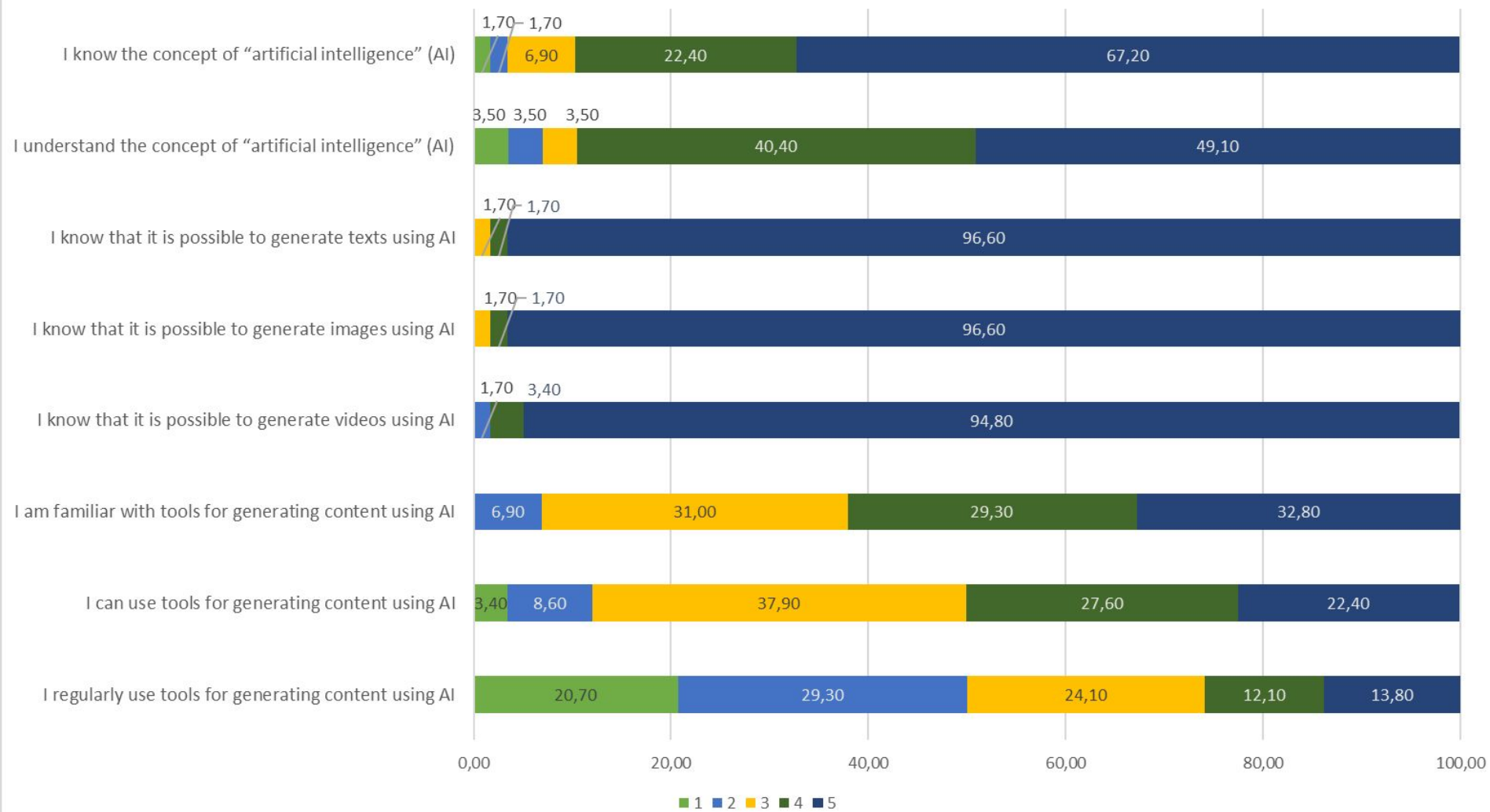
Our Study – Aim and Methodology

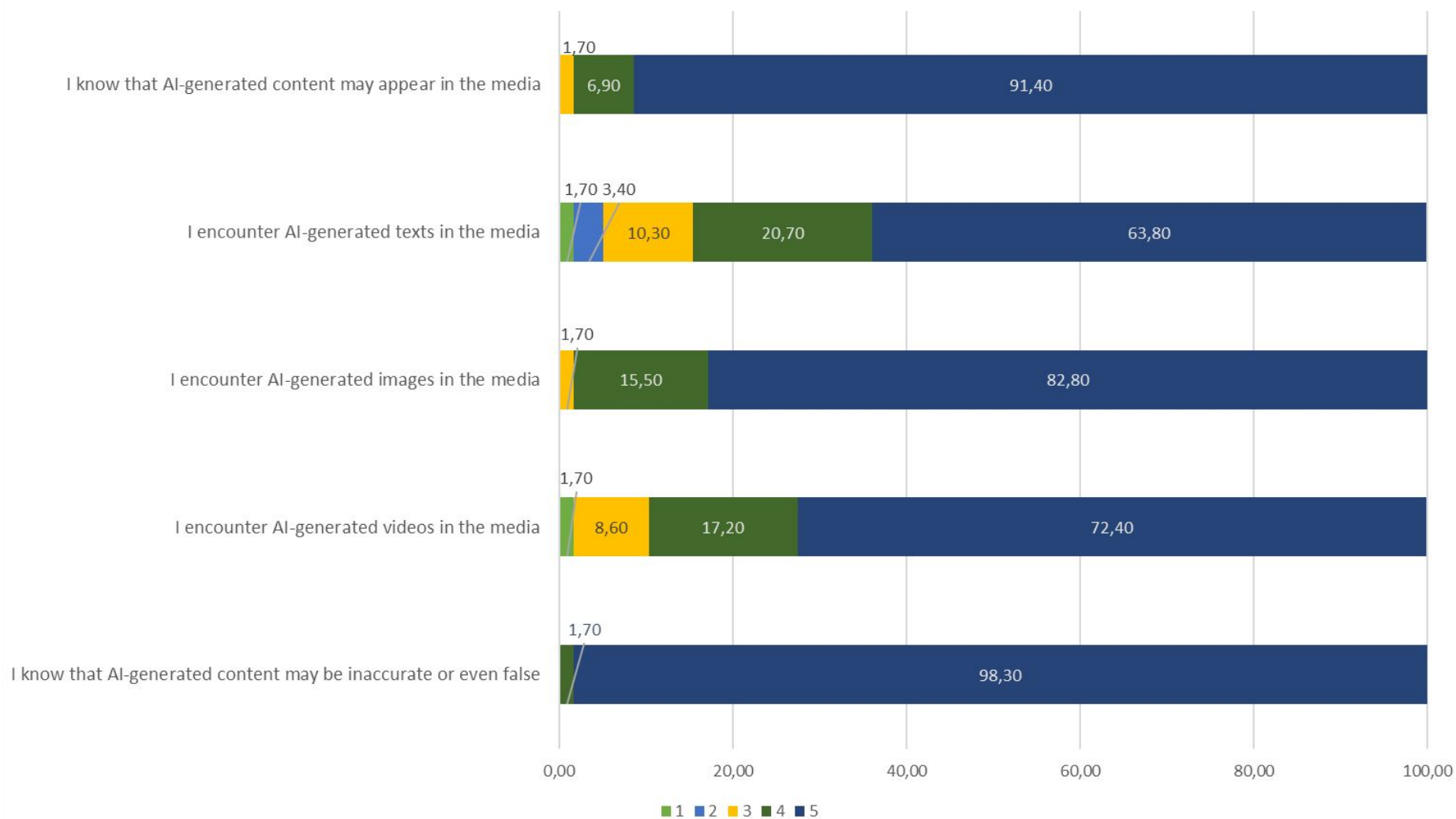
- **Aim:** determine the degree and scope to which awareness of AI-generated content in the information space affects users' sense of information security
- **Broad research perspective:** 4 thematic areas
- **Respondents:** 58 MA students (“Archiving, Office Management and Digitisation” and “National Security”)
- **Methodology:** questionnaire (self-assessment) and experiment (detecting synthetic content), declarations vs. actual feelings
- **Timeline:** October–November 2025

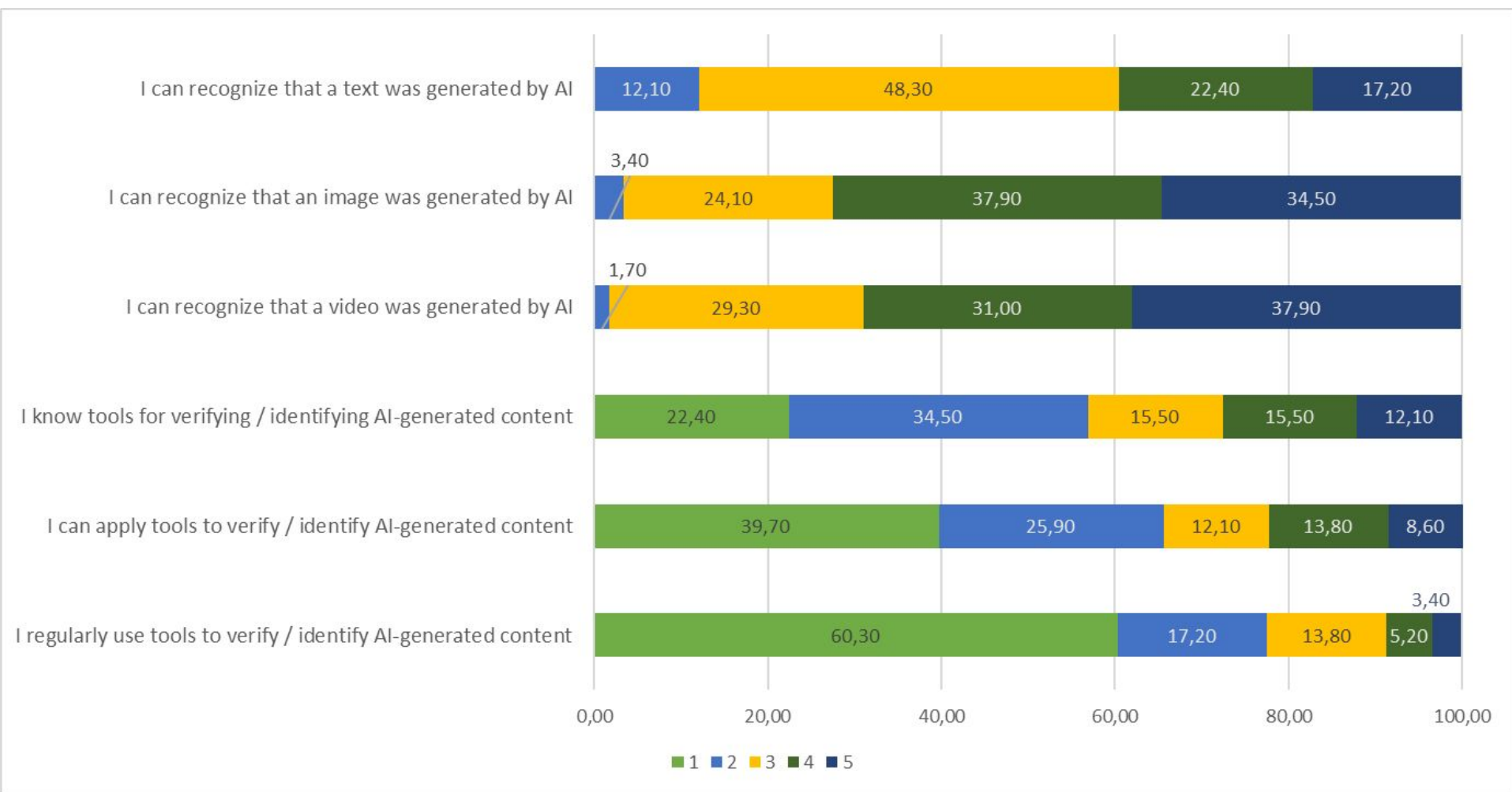
Research Questions

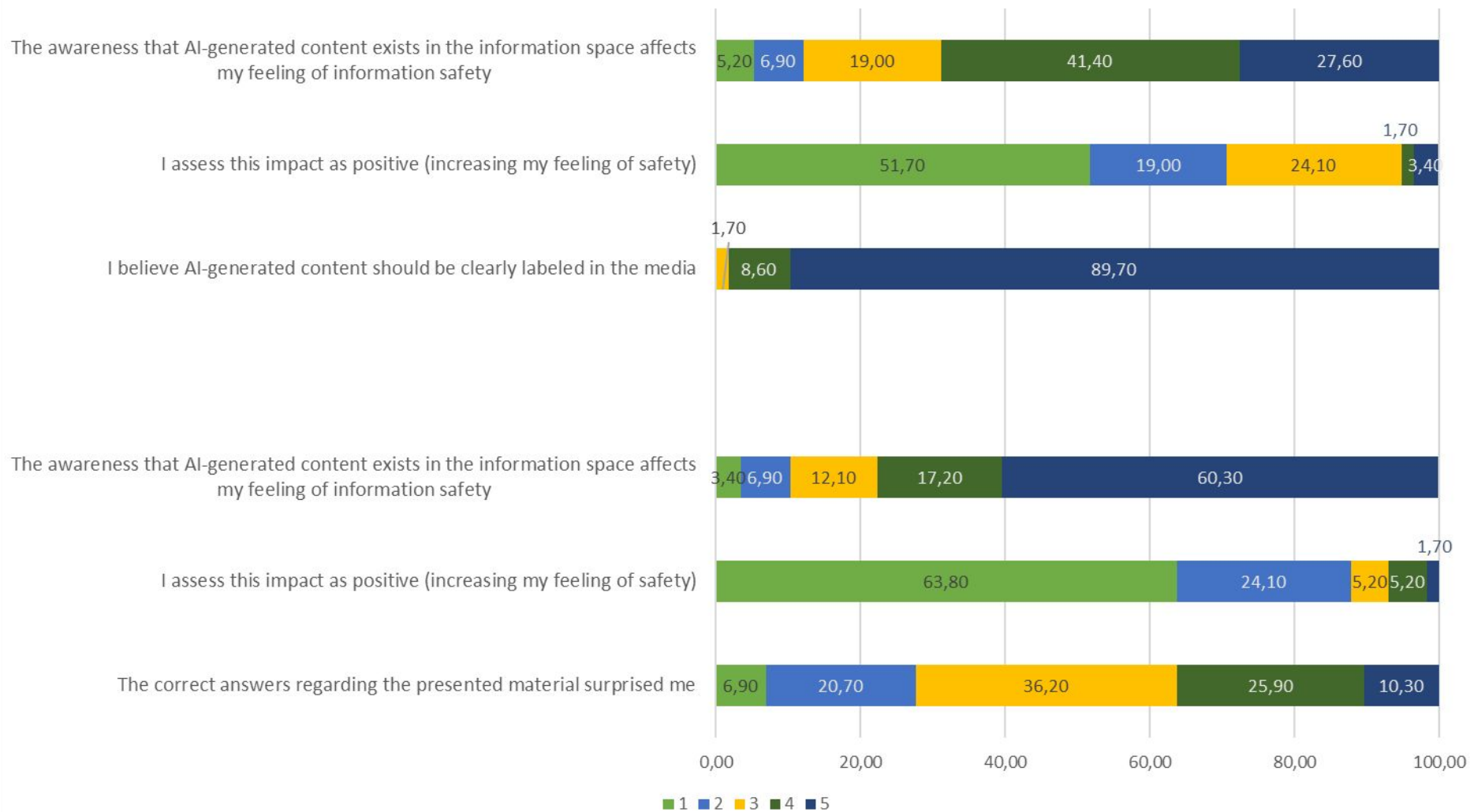
- Q1.** How do students evaluate their awareness and understanding of AI-related issues?
- Q2.** To what extent do they encounter AI-generated content in their information environment?
- Q3.** How competent do they feel in identifying AI-generated content?
- Q4.** To what extent does awareness of AI-generated content affect their feeling of information safety?

RESULTS









Qualitative Results

Awareness that AI-generated content exists in the information space evokes the following feelings (selected responses):

01

Need to verify information (21)

- "AI-generated content creates more work that requires even greater focus. This comes from the need to verify encountered content more thoroughly and more deeply."
- "I know there is a lot of fraud related to this, but for now I try to verify it"

02

Feeling misled / manipulated (14)

- "When I see something unusual in a photo or video, I immediately doubt its authenticity and feel like someone is trying to 'deceive' me."
- "It is harmful and dangerous because it makes it easy to manipulate people"

03

Anxiety (10)

- "I feel uneasy because I don't know whether what I see is real and whether I can trust it"
- "It makes me feel uncertain and anxious (especially videos), because I know that many people — including myself sometimes — are inclined to believe in unreal or even real-looking but false events or situations."

04

Aversion to AI-generated content (2)

- "Situations in which I even have to stop and think about it make me uncomfortable. AI content is not needed in life. What for?!"
- "I avoid all AI-generated content."

Qualitative Results — AFTER reviewing materials

Awareness that AI-generated content exists in the information space evokes the following feelings (selected responses):

01

Difficulty distinguishing AI-generated content from human-created content (18)

- "I feel that you can't recognize AI-generated text with the naked eye."
- "It's becoming harder and harder to distinguish real content from AI, and the awareness that you can never be 100% sure what is what is very frustrating."

02

Distrust / lack of trust (12)

- "I'm still distrustful, but now I assume by default that something was generated by AI. It's my first instinct."
- "I'm not sure whether the information I receive is true, which makes me distrustful and creates the need to find confirmation in other sources."

03

Concern / fear / anxiety (10)

- "I'm worried about what will happen in the coming years."
- "The prospect of fraud against, for example, elderly people using AI terrifies me."
- "It's frightening that people create such things usually for their own benefit and use them for various purposes."

04

Need for labelling AI-generated content (3)

- "In my opinion, AI-generated information should be labelled."
- "AI-generated materials should be labelled."

Qualitative Results

Differences before and after reviewing the materials:

01	Change in the most frequently occurring respondent answers	• BEFORE → NEED TO VERIFY INFORMATION and AWARENESS OF THE EXISTENCE OF FALSE CONTENT (dominant feelings: distrust, fatigue, cognitive effort required) • AFTER → DIFFICULTY DISTINGUISHING AI vs. HUMAN CONTENT (feelings: greater awareness of perceptual difficulties, slightly less emphasis on verification)
02	New observations	• Need for regulation and transparency in labelling AI-generated content • RADICAL REACTIONS – desire to ban AI, reluctance to seek out and verify information “everything may be fake”
03	Changes in the level of detail in respondents’ statements	• BEFORE – respondents declared AWARENESS of the existence of false content, but AFTER → they pointed to specific concerns / fears / distrust (e.g., concerns about fraud targeting older people)
04	Unchanged reactions	• some respondents indicated no change in their sense of safety and/or in their perception of AI-generated content (16) after reviewing the materials

Discussion

- concerns related to the consequences of using AI: manipulation, fraud
 - (consistent with Gackowki et al.)
- high confidence in knowledge of AI technologies and ability to recognise synthetic content
 - (consistent with Nowina-Konopka 2023; consistent with Motylińska & Pieczka-Węgorkiewicz after ECIL)
- a high level of respondent confidence in their own ability to recognise AI-generated content
 - (consistent with Motylińska & Pieczka-Węgorkiewicz after ECIL)
- uncertainty about authorship, difficulty in making a clear assessment, risk of manipulation and disinformation; sense of being “deceived”
 - (consistent with Nowina-Konopka — “disorientation”)

Conclusion

- Awareness of AI-generated content in the information space **does affect** users' feeling of information safety
- The effect is negative — it **decreases the sense of security**
- **Recurring themes:**
 - fear, worry, unease,
 - risk of deception / manipulation
 - need for information verification
- Participation in the experiment **intensified or polarised** respondents' reactions (post-experiment responses were less neutral)

Further Research Directions

01

Real skills in detecting AI-generated content

- Comparison of self-assessed skills and actual recognition of AI-generated content

02

Feeling of information safety

- Before and after contact with AI-generated content
- across different user groups

03

Developing competencies in detecting AI-generated content

- analysis of strategies for improving these competencies (desk research)
- empirical studies on the effectiveness of these strategies (e.g., feedback training)

Bibliography

Gackowski, Tomasz et al. (2025). Między lękiem przed inteligentnymi maszynami a zaufaniem do nowych technologii: postrzeganie sztucznej inteligencji w badaniach opinii publicznej. W: Praktyki medialne w zdigitalizowanym świecie. Kraków, Instytut Dziennikarstwa, Mediów i Komunikacji Społecznej UJ, s. 47-69.

Motylińska, Paulina; Pieczka-Węgorkiewicz, Anna (w druku). Self-Assessment of the Polish Students' Artificial Intelligence (AI) Literacy in the Context of AI-Generated Content Detection. CCSI, Springer.

Nowina-Konopka, Maria (2023). "Każdy zna się na AI": przegląd badań polskiej opinii publicznej na temat sztucznej inteligencji. Zeszyty Prasoznawcze, nr 4 (256), s. 127-142.

Skrzypiec, Aleksandra (2025). Korzystanie z generatywnej sztucznej inteligencji w codziennym życiu na przykładzie chatbota GPT. W: Praktyki medialne w zdigitalizowanym świecie. Kraków, Instytut Dziennikarstwa, Mediów i Komunikacji Społecznej UJ, s. 313-335.

Wang, D., Guo, J., Zheng, K. (2025). AI or Human? An Analysis of University Students' Awareness of Reference Services Agent. In: Oliver, G., Frings-Hessami, V., Du, J.T., Tezuka, T. (eds.) Sustainability and Empowerment in the Context of Digital Libraries. ICADL 2024. Lecture Notes in Computer Science, vol. 15494, pp. 230-240.

Thank you for your attention

paulina.motylinska@uken.krakow.pl

anna.pieczka-wegorkiewicz@uken.krakow.pl