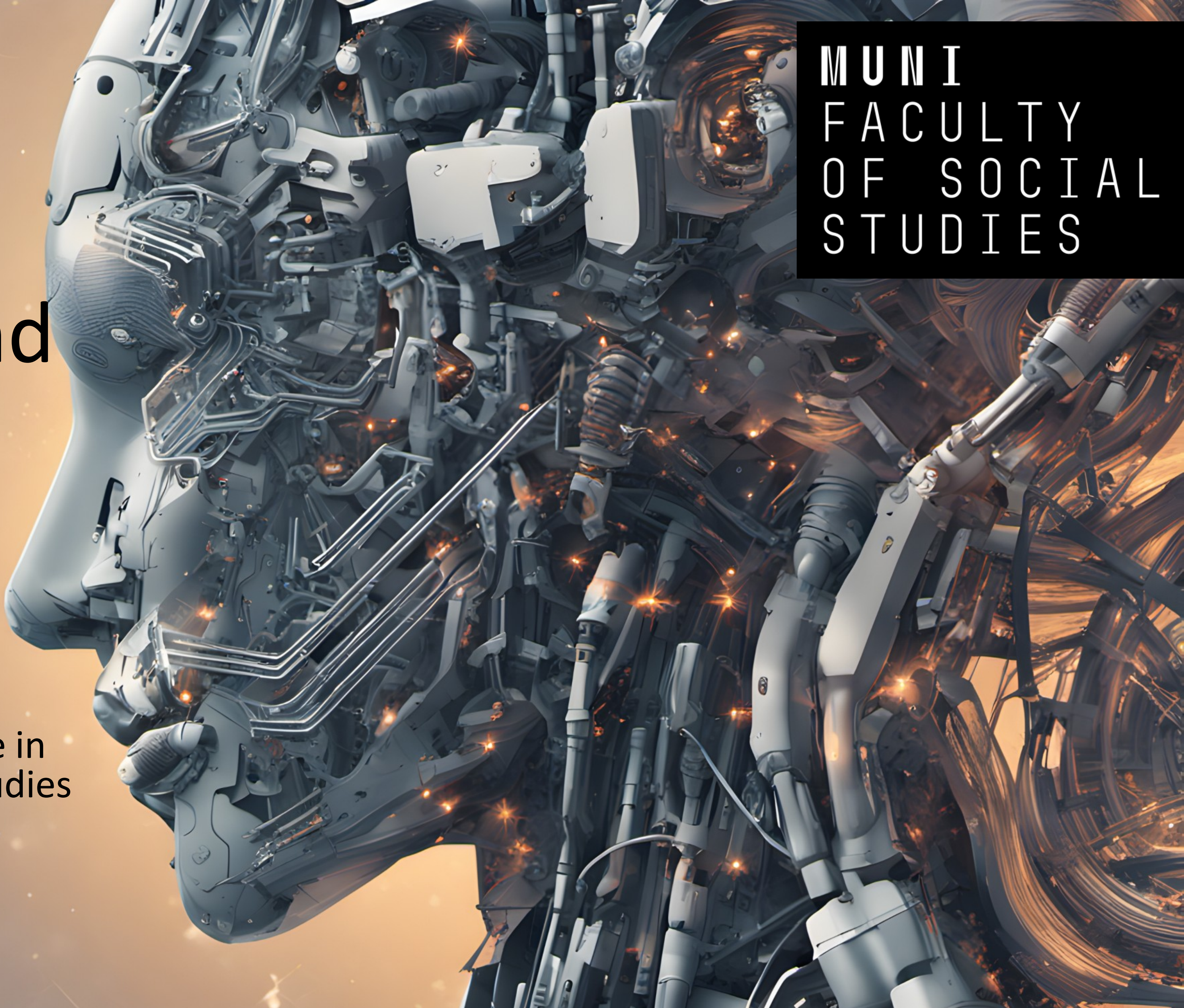
MUNI
FACULTY
OF SOCIAL
STUDIES

AI in Security: Applications and Ethics



15.10.2024

GLCb2028 Artificial Intelligence in Political Science and Security Studies

Jan **KLEINER**

jkleiner@mail.muni.cz

Literature

- Mahsur, N. (2019). AI in Military Enabling Applications. CSS Analyses in Security Policy, 251. https://doi.org/10.3929/ethz-b-000367663

- Pedron, S. M., & da Cruz, J. D. A. (2020). The future of wars: Artificial intelligence (ai) and lethal autonomous weapon systems (laws). International Journal of Security Studies, 2(1), 2.

Presentation outline

- PolSci and AI generally and its origins.
- PolSci examples of communication research.
- AI and cybersecurity.
- LAWS.
- Migration.
- Wargames and theory (preparation for the incoming seminars).
  - In 2 weeks – we will add intelligence analysis tools and design a proper wargame.

PolSci and AI generally (Duffy & Tucker, 1995)

- Early applications of AI in research focused on constructing choice models in foreign-policy decision contexts.
- **Other applications:**
  - Production systems,
  - computational text analysis,
  - logic programming and computer learning,
  - conflict simulation and predicting outcomes in international conflicts via machine learning.
- **AI** + **computer vision** + **natural language processing** + **sentiment analysis** → set to **transform** society, the economy, and politics (Efthymiou-Egleton, Egleton & Sidiropoulos, 2020).
- AI can create **new ways of** (researchable) **communication** (alphabets, iconographics, languages etc.) (Mueller & Massaron, 2021).

Three examples of PolSci (communication) research

- 1. **Can AI communication tools increase legislative responsiveness and trust in democratic institutions**? (Kreps & Jakesh, 2023).
  - Recent.

- 2. **Artificial intelligence and European identity: the European Commission's struggle for reconciliation** (von Essen & Osseewarde, 2023).
  - Recent.

- 3. **Rise of the Machines? Examining the Influence of Social Bots on a Political Discussion Network** (Hagen et al., 2022).
  - Cited (30x – SCOPUS).

AI tools and responsiveness and trust in democratic institutions (Kreps & Jakesh, 2023)

- Legislative correspondence generated by AI with **human oversight** may be received favorably by constituents and increase trust and legislative responsiveness **compared to generic auto-responses**.

- Poorly performing AI may damage confidence in legislators.
- Still unclear specific impact of AI to political communication.

- Technologies like ChatGPT could **streamline democratic processes** rather than destabilize them → **BUT:** authors do not mention dis/mis/information or propaganda threats (cf. Hagen et al., 2022).

- **HITL** and **SITL concepts** (Rahwan, 2018).

EU´s approach to AI (von Essen & Osseewarde, 2023)

- The European Commission aims to develop European version of AI, but its communication efforts may not be sufficient to **generate trust** in AI among the European public.

- The EC frames European AI as **trustworthy** and **human-centric**, based on European values and historical success, but fails to connect its claims to specific European values

Social bots´ impact on political discussion network (Hagen et al., 2022)

- Social bots (automated accounts on social media), often utilize AI techniques to generate content, interact with users, spread information etc.

- Social bots can **significantly impact** political discussion networks by **creating the appearance of virtual communities**, attenuating the influence of traditional actors, and **amplifying** pro-Trump messaging.

- Bots are often utilized by actors with ideological positions reflective of a **small subset** of the public (e.g., the far-right).

- The potential for spreading misinformation, which **undermines democratic processes**.

AI and Cybersecurity (Bonfanti et al., 2021)

- AI as an **underdeveloped** field in social sciences (AI politics research years behind the cybersecurity politics one).
- Inter and **transdisciplinary** (decisions and research in one discipline transpires into other ones).
- Well suited for cyber **defense** and **offense** + **influence ops**.
- „…in what ways will AI **enhance the protection** of individuals, organizations, nations, and their cyber-dependent assets from **hostile threat actors**?
- How will it introduce **novel vulnerabilities** and enable additional typologies of actions?
- How will it induce cyber-security **stakeholders** to **adapt** to **changing** risk scenarios and opportunities?“ (p. 226).

LAWS (Sauer, 2021)

- Lethal autonomous weapons systems.
- **Autonomy** vs. **automation** – no consensus on delineation → e.g., functionalists: machine instead of human performing the task.
- „**kill chain**“ = finding, fixing, tracking, selecting, and engaging the target (+ assessing the aftereffects).
- Autonomy incl. critical functions is not new, but AI scales it up heavily.
- **Incentives** – no fear, emotions, fatigue, mercy, speed of (re)action etc.
- Technological, ethical, legal, strategic **criticism**.
  - E.g., „the accountability gap“ (p. 241) – someone has to be accountable for war actions.

Migration (Everuss, 2021)

- New fields like digital migration studies.
- Digitization of borders historically led by USA and EU.
- **Biometrics** → „…actionable inferences about personality, intent, emotional state, social conformity, sexual orientation, and many other… attributes“ (Crampton, 2019: 55).

AI and Wargames (Knack and Powell, 2023)

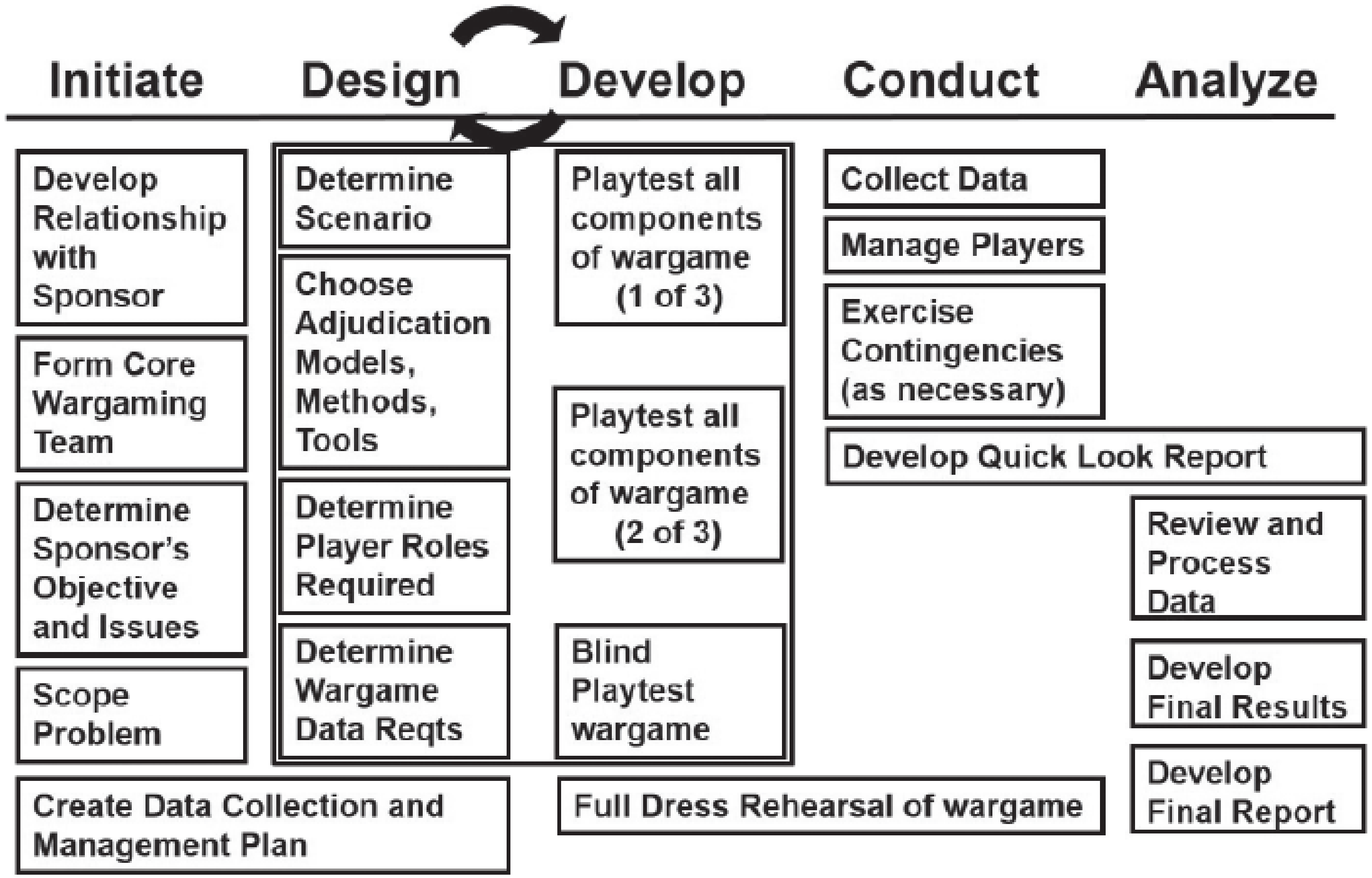
- Red Teaming in general (political/security/other simulations, table-tops -> identification of gaps in a strategy, SWOT analyses, policy analyses etc.).
    - **Narrow (safe) usage**: Repetitive tasks within sims and wargames (background info creation, automatic translation/transcription, textual data analysis, visuals etc.).
    - **High-risk usage**: Red team, game manager etc.
- Low cost/questionable reliability.
- Better on **tactical**/**operational** level than on the **strategic** one.

Wargame theory – introduction I (Appleget et. al, 2020)

- Usually a sponsor – sets goals and timeframe.
- Sole purpose is to collect **analytic data** to answer sponsor´s (research) questions – data determine wargame´s success → **well tought-out data collection plan is needed!**
- Roadmap = data collection and management plan (DCMP).
- Not just for combat/conflict scenarios, but for **Analysis of alternatives (AoA)** – e.g., M1A2 Abrams and its replacement options.
- + pedagogic, research tool.

Wargame theory – introduction II (Appleget et. al, 2020)

- **Course of action wargaming**.
- **BOGGSAT** = "bunch of guys and gals sitting around a table„.
- Vs.
- **Seminar wargames** - designed around the DCMP (Decision-Centric Methodology Process) and have a structured approach.
- **Quantitative/qualitative/hybrid** models.
- Strong role of **probability** and chance (dice rolls) + conditioned probability (e.g., missile interception of Iron Dome AA system – informed by statistics).


Initiate

Design

Develop

Conduct

Analyze

Develop Relationship with Sponsor

Form Core Wargaming Team

Determine Sponsor's Objective and Issues

Scope Problem

Create Data Collection and Management Plan

Determine Scenario

Choose Adjudication Models, Methods, Tools

Determine Player Roles Required

Determine Wargame Data Reqts

Playtest all components of wargame (1 of 3)

Playtest all components of wargame (2 of 3)

Blind Playtest wargame

Full Dress Rehearsal of wargame

Collect Data

Manage Players

Exercise Contingencies (as necessary)

Develop Quick Look Report

Review and Process Data

Develop Final Results

Develop Final Report

Five Phases of Wargame Construction

Source: Appleget et al. (2020, p. 73).

Let´s do some BOGGSAT wargame!

What are your areas of research interest?

References I

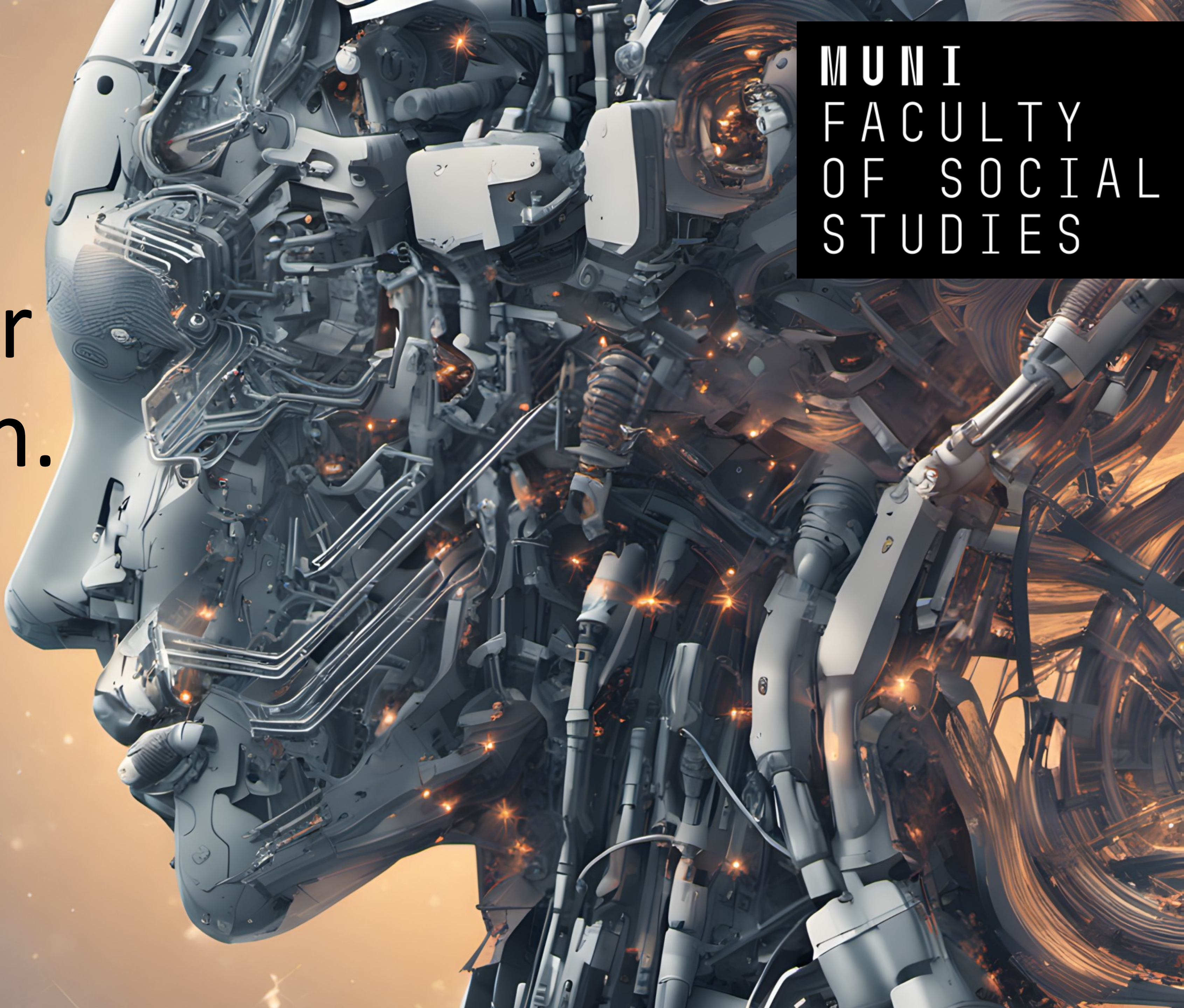
- Appleget, J. Burks, R. & Cameron, F. (2020). The Craft of Wargaming – A Detailed Planning Guide for Defense Planners and Analysts. Annapolis: Naval Institute Press. ISBN 9781682473771
- Bonfanti, M. E., Cavelty M. D., Wenger, A. (2021). Artificial intelligence and cyber-security. In: Elliott, A. (Ed.). (2021). The Routledge Social Science Handbook of AI (1st ed.). Routledge. https://doi.org/10.4324/9780429198533.
- Duffy, G., & Tucker, S. A. (1995). Political Science: Artificial Intelligence Applications. Social Science Computer Review, 13(1), 1–20. https://doi.org/10.1177/089443939501300101.
- Efthymiou-Egleton, I. P., Egleton, T. W. E., & Sidiropoulos, S. (2020). Artificial Intelligence (AI) in Politics: Should Political AI be Controlled?. International Journal of Innovative Science and Research Technology, 5(2).
- Everuss, L. AI, Smart Borders and Migration. In: Elliott, A. (Ed.). (2021). The Routledge Social Science Handbook of AI (1st ed.). Routledge. https://doi.org/10.4324/9780429198533.

References II

- Hagen, L., Neely, S., Keller, T. E., Scharf, R., & Vasquez, F. E. (2022). Rise of the Machines? Examining the Influence of Social Bots on a Political Discussion Network. Social Science Computer Review, 40(2), 264–287. https://doi.org/10.1177/0894439320908190
- Knack, A. & Powell, R. (2023). Artificial Intelligence in Wargaming An evidence-based assessment of AI applications. The Alan Turing Institute.
- Kreps, S., & Jakesch, M. (2023). Can AI communication tools increase legislative responsiveness and trust in democratic institutions? Government Information Quarterly, 40(3), 101829. https://doi.org/10.1016/j.giq.2023.101829
- Mueller, J. P. & Massaron, L. (2021). Artificial Intelligence for Dummies. Hoboken: John Wiley and Sons, Inc.
- Sauer, F. (2021). Lethal autonomous weapons systems In: Elliott, A. (Ed.). (2021). The Routledge Social Science Handbook of AI (1st ed.). Routledge. https://doi.org/10.4324/9780429198533.
- von Essen, L., & Ossewaarde, M. (2023). Artificial intelligence and European identity: the European Commission’s struggle for reconciliation. European Politics and Society, 1–28. https://doi.org/10.1080/23745118.2023.2244385

MUNI
FACULTY
OF SOCIAL
STUDIES

Thank you for your attention.



Questions?

Jan KLEINER

jkleiner@mail.muni.cz