#Standards4AI

‘Putting Science Into Standards’ workshop

Welcome!
We will start soon

AI for Media including Social Media, content moderation, recommender systems

9 June, 10:45-12:00

Panel discussion
AI for Media

Roundtable speakers

Symeon Papadopoulos

Centre for Research and Technology Hellas

Jochen Leidner

Coburg University

Rapporteur: Alexandra Balahur (JRC)

Audience interaction

slido.com

#Standards4AI

- Select the **Media** room on Slido
- Zoom chat - only technical questions to host
- Camera and audio OFF

Session structure

1. Brainstorming

Identify specific aspects which require standardization
Identify standardization committees or working groups and existing standards

3. Prioritisation

⏱ 15 minutes

Based on the feasibility and importance of standardization activities, identify priority needs. Copy and paste the sticky notes from previous steps.

2. Mapping - categorisation

⏱ 30 minutes

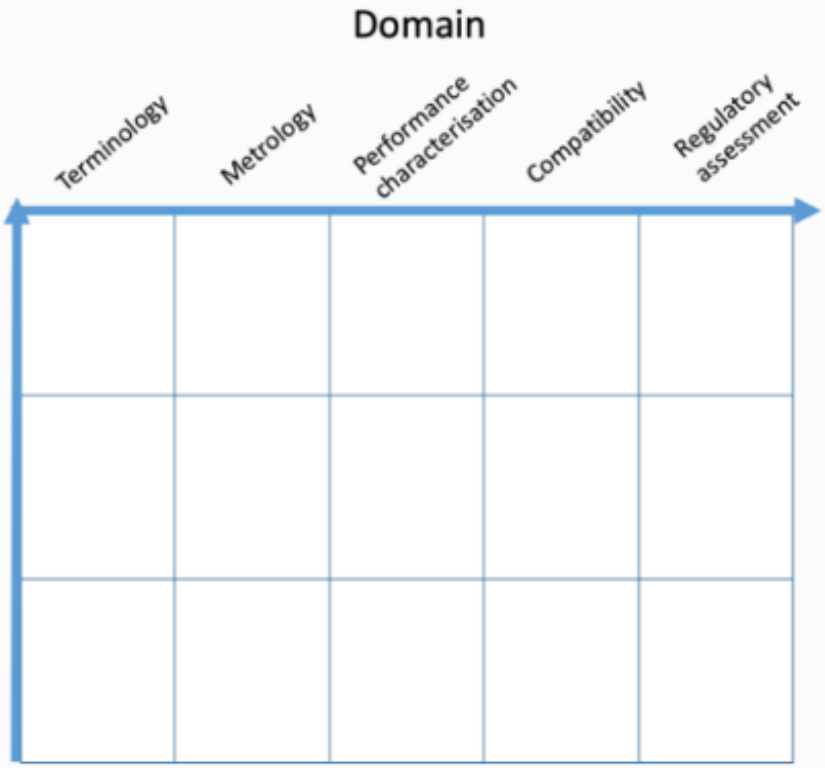
- Map standardisation needs for a) identifying and compiling data for eventual training of the AI system (first matrix) and b) data use within the AI system to be delivered (second matrix).
- Map required standards by considering the category of standards (x axis: terminology, metrology etc.) versus the innovation stage (y axis: technology, production, market)

Development pathway

AI system deployment & marketing
- *Regulatory assessment*
- *Users*
- *Transparency & specification*
- *Accountability / Responsibility*
- *Maintenance, post-market follow-up & bias monitoring*
- *Supply network*

AI system creation & production
- *Data sets & algorithms (incl. bias) / models*
- *Cybersecurity*
- *System design & integration*
- *Upscaling & evaluation*
- *Quality control*

Data creation
- *Compilation, preparation, bias testing,*
- *Analysis, processing, labelling,*
- *licensing & restrictions,*
- *Sharing & marketing*



#Standards4AI

'Putting Science Into Standards' workshop

Symeon Papadopoulos

CERTH/ITI

Professional background

- Computer science, large-scale multimedia analysis and mining with particular emphasis on applications in the Media and Social Media
- Media verification, online disinformation
  - Image forensics
  - Deepfake detection
  - Content moderation
  - Social network mining
- AI Bias and Fairness
  - Bias discovery in visual datasets
  - Bias mitigation

Challenges Faced & Solutions

- Multiple definitions of bias in AI
  - What is most suitable/appropriate is domain-/task-dependent
  - Individual vs group fairness (potentially conflicting)
  - Certain forms of bias are hard to capture and quantify
- Good understanding of bias often relies on access to datasets
  - Typically not possible to access “real-world” data
  - Concept drift makes the problem even harder
- Working in inter-disciplinary teams involving both AI and domain experts can lead to better understanding of issues
- Creating representative datasets that can be reused by the research community, e.g. for benchmarking) is important
  - Copyright and privacy issues!

Way Forward, Next Steps

- Benchmark datasets
- Auditing tools (e.g. dataset auditing)
- Guidelines towards compliance
- Short term goals:
  - Inter-disciplinary groups working towards foundations and resources
- Long term goals:
  - Tools that can be used by academia and industry to measure and mitigate bias

#Standards4AI

'Putting Science Into Standards' workshop

Thank you!

Symeon Papadopoulos / papadop@iti.gr / @sympap

#Standards4AI

'Putting Science Into Standards' workshop

Jochen Leidner

Coburg University

Professional background

- Professor of Explainable and Responsible Artificial Intelligence (Coburg, Germany) and Visiting Professor, Data Analytics (Sheffield, UK)
- Former roles include Software Engineer, SAP and Director of Research (R&D), Reuters London
  - built many industry AI systems (2008-2021) in the domains news/journalism, legal, risk/compliance/counterterrorism, finance/trading/investment, pharma
  - Many systems used in production (by traders, banks, lawyers, US supreme judges, Reuters journalists)
- Education: M.A. (computational linguistics, English language & literature, computer science, FAU Erlangen), M.Phil. (computer speech, text and Internet technology, University of Cambridge), Ph.D. (computer science, University of Edinburgh)
- Teaching: Machine Learning, AI & Ethics, Bias/Fake News
- Research on Natural Language Processing (NLP), Information Retrieval (IR, search) and Machine Learning (ML): question answering, summarisation, topic modelling, automatic risk identification, terminology mining, entity resolution/geolocation of text, professional vertical ad-hoc retrieval engines using learning to rank etc.
- Public policy: Input to UK *House of Lords* AI Select Committee on AI Regulation

Challenges Faced & Solutions

- Important questions to consider:
  - 1. What makes good data 'good data'? quality dimensions
  - 2. What is standardization? - What it can and cannot do.
  - 3. What makes a data ecosystem (e.g. the Web, commercial data) work?
    - incentives (e.g. present Web: attention = $$$)
  - 4. Who should be the gatekeeper of truth?
    - = power, so avoid single point of failure (e.g. not a big tech platform owner)
- Has been working on a data science methodology, ***Data-to-Value*** (e.g. Leidner, 2022, *Proc. NLDB*) that made ethical considerations an integral part of project work.

Way Forward, Next Steps

- Recent trend: large corporations are teaming up to pay for lobbying to prevent privacy protection laws
- GDPR was a way forward (European is leading regarding privacy protection) but implementation is still spotty & lacks "teeth"
- Data quality or ethics audits lack experienced talent to conduct them & are expensive, time intensive
  - more use of open source can help (= transparency in development processes)
- No solution for the fact that all efforts are jurisdiction bound

#Standards4AI

'Putting Science Into Standards' workshop

Thank you!

#Standards4AI

'Putting Science Into Standards' workshop

Manuel Gomez Rodriguez

MPI

Professional background

- Manuel Gomez Rodriguez is a faculty member at the Max Planck Institute for Software Systems, where he leads the Human-Centric Machine Learning group. He holds a PhD in Electrical Engineering from Stanford University.

- Manuel develops machine learning algorithms to enhance the functioning of social, information and networked systems.

- Manuel has received several recognitions for his research, including an ERC Starting grant and an outstanding paper award at NeurIPS, the flagship conference in AI.

Challenges Faced & Solutions

- We have shown that one of the root causes of data bias and lack of inclusiveness in AI for Media is the feedback loop between algorithmic & human decisions. And found a solution!

- However, there are many other root causes for data bias and lack of inclusiveness and trustworthiness of AI. Different root causes require different solutions.

- Difficult to find the root cause in specific cases due to the opacity of media companies regarding the data and optimization objectives used in their AI models.

Way Forward, Next Steps

- User data standardization in (social) media may help users change sites without *loosing* their data. However, difficult to incentivize social media companies to adopt data standards.

- Standardization of privacy/data policies used in (social) media sites (e.g., ensure private ML/advertising/etc.).

- In the long term, as more content is created algorithmically (e.g., images, voice, text), watermarking generated content is needed. It would help to have a standard for watermarking.

#Standards4AI

'Putting Science Into Standards' workshop

Thank you!

Programme – what's next?

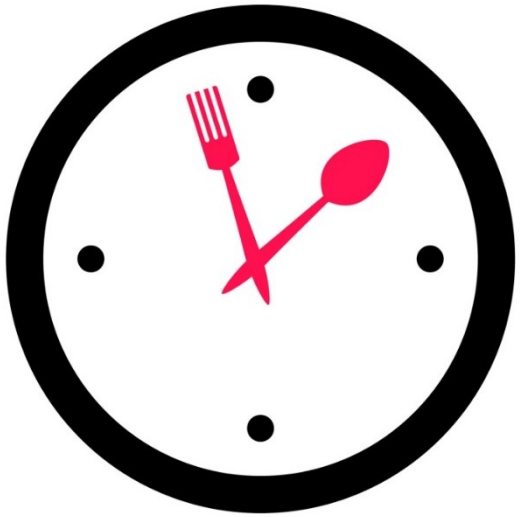
Main Plenary room

12:00 - 13:30	*Lunch break*
13:30 - 14:15	**Flash Summaries of parallel sessions**
14:15 - 15:45	**Panel discussion on ways forward**
15:45 - 16:00	**Concluding remarks**

Please check your confirmation email for the links to access main plenary room

The link will also be published on Slido and Zoom chat

Let’s take a break!



LUNCH BREAK

See you in the plenary room at 13:30!