

**IEEE P7003 Working Group
Meeting Minutes
15 February 2017 /10:00 A.M. – 11:30 A.M. EST
Teleconference**

1. Call to Order

The meeting was called to order at 10:00 A.M.

2. Roll call and Disclosure of [Affiliation](#)

The list of attendees present is attached.

3. Approval of February Agenda

Motion to approve the meeting agenda from 15th February, 2017. The agenda was approved as submitted without objection.

4. [IEEE Patent Policy](#) (Call for Patents)

The call for patents was raised; no one raised any concerns or any comments for consideration.

5. Approval of 12th January meeting minutes

Motion to approve the minutes from January 12th, 2017. The minutes from the January 12th, 2017 meeting were approved without objection.

6. Updated Outline Discussion

I. Update on Use Cases

No new use cases were raised.

Michael A and David sent a paper suggestion to the group on justified bias, and sensitive characteristics within the context of medical use, which could be turned into a use case. Medicine is a great opportunity for refining the lines of justifiable and unjustifiable bias, so it will also be included in one of the sections of the standard.

There were several questions about what appropriate bias is and whether any work had been done in this area. For example, if it does not disadvantage anyone, is it rightly called bias at all? This led to discussion of how to clarify what justification means in terms of this standard (statistical bias versus moral bias)

II. Topic updates (development plans) – 5min summaries

If anyone didn't receive an invite to the expected doc, let Ansgar know

i. Key concepts

No update

ii. Taxonomy (Pascal)

Next call in about 2 weeks. If anyone wants to join, let Pascal or Ansgar know. They have collected a large list of potential sources and papers to work on, should start bringing taxonomy together in useable form in 1-2 months

iii. Person categorization (Joseph)

Looking at various ways people can be categorized and creating a list.

Looking for more documentation.

iv. Legal frameworks (Maroussia)

Started mapping out in the docs, the framework, what equality means and how it plays out. Need more people to be involved, especially with knowledge of European frameworks and other countries. Shared examples of bias etc. in chat.

- Predictive policing reproduces bias against racial minorities by deploying patrols according to historical arrest data

- Dynamic pricing tools charge higher fees for college admission test preparation in Asian-dominated areas

- Facial recognition tools trained on Caucasian faces label dark-skinned individuals as gorillas and fail to recognize Asian features, pointing to a lack of diversity in the training data as well as in the engineers creating these systems

- Online ads for high paying job ignoring women - because more coveted demographics (unintentional bias)

- Word2Vec framework mapping semantic proximity between words. Trained on Google News articles, Word2Vec associates father with doctor and mother with nurse

- Female digital assistants reinforcing gendered stereotypes

Also discussion of avenues for undoing bias. There was some debate over changing which features are used or allowed, compared to being transparent about what is being used (hidden proxies eg postcode not race)

Problem of universal framework led to suggestion of creating metaprinciples which may play out differently across jurisdictions

v. Psychology (BC)

Putting together rough outline of approach and how to research types of categorization in information processing – cognitive and motivations.

Correlations between x and y (see doc)

vi. System design (Chris C)

Haven't got together yet – hopefully something this week.

Action: Ansgar to resend link to doc.

vii. Representativeness of data

No update

viii. Outcomes evaluation (Michael R)

Arranging a call to discuss asap, next week hopefully

ix. Evaluation of processing (Michael R)

As viii

x. Resilience against manipulation (Michael R)

As viii

xi. Documenting (and culture) (Maya)

They had a call this morning focusing on the culture section. It is important that this is not a consideration tagged on at the end, need to have a critical

approach to culture throughout the document. Including different concepts of culture: geography, law, engineering vs social science etc.
 Maya asked that the culture group also get access to the other google docs.
 Led to discussion of working in the main google doc versus all the small ones, and would a shared folder be better? General feeling that lots of docs is good but people should have access to them.

7. Any other Business

There are still some subsections that need people to work on.

The best way to organize a meeting with a subgroup is either through google docs, or contact Ansgar for emails.

Had Eastern time zone meeting with one participant (Lyria). Another one is to be scheduled in March aimed at west coast of US.

Ansgar submitted short paper to Fairware2018 (Ansgar, Maya, Liz)

KDD workshop: Chris C suggested getting involved in another workshop proposal already in submission (contributing paper etc.)

Amsterdam Privacy Conference: 15 march deadline.

AAAI fall symposium series at Stanford suggested as good venue

We've been asked to comment on a white paper from the IEEE Internet policy group on Algorithmic Decision making. *If you are interested, please email Ansgar.*

8. Future Meetings

Wednesday, March 14th, 3:00pm-4:30pm (UTC) /10:00am-11:30am (EST)

Tuesday, April 10th, 3:00pm-4:30pm (UTC) /10:00am-11:30am (EST)

Monday, May 7th, 3:00pm-4:30pm (UTC) /10:00am-11:30am (EST)

Friday, June 8th, 3:00pm-4:30pm (UTC)/10:00am-11:30am (EST)

9. Adjourn

Meeting was adjourned at 11:09 AM

Attendees:

Last Name	First Name	Employer/Affiliation	Voting
Abramoff	Michael	University of Iowa and IDx	
Biermann	BC	Heavy Projects & USC Annenberg Innovation Lab	x
Clifton	Chris	Purdue University	x
Coelho	Joseph	Marquette University, Milwaukee	
Dechesne	Francien	Leiden University	x
Dowthwaite	Liz	University of Nottingham	x
Duffy	Stewart	RadcliffesLeBasseur	x
Ganesh	Maya	Leuphana University	x
Hatada	Yohko	EMLS_RI	x
Horner	John	Self	
Jurgens	Pascal	U of Mainz, Germany	x
Koene (Chair)	Ansgar	University of Nottingham	x

Lévesque	Maroussia	Independent	x
Ng	Vivian	University of Essex	
Nyrup	Rune	Leverhulme Centre for the Future of Intelligence, University of Cambridge	
Piwowar	Kuba	University of Social Psychology and Humanities, Warsaw	
Rao	Shrisha	IIIT Bangalore	
Rovatsos	Michael	University of Edinburgh	x
Ruggeri	Salvatore	University of Pisa	
Srivathsan	KM	TATA Consultancy Services	
Stender	Matthew	Self	x
Vidal	David	IDx	
Wolf	Christine	IBM Research	x
Bahn	Christy	IEEE-SA (staff)	