

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

**1. Call to Order**

The meeting was called to order at 10:03 A.M.

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

The list of attendees present is attached.

**3. Approval of November Agenda**

***Motion to approve the meeting agenda from 17<sup>th</sup> November, 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 12 October meeting minutes**

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

**6. Updated Outline Discussion**

I. Update on Use Cases

Discussion of medical technology as an outlier or special case led to consideration of the development of autonomous weapons and military tech as another potential outlier.

Autonomous Weapons: This sector is likely to look at things differently and not use the standard, but it was suggested it could be a useful use case to stretch boundaries. If the standard does not cover this area well we should explicitly say it is not something we are handling. If it does, it could be used by those in AW to inform the space rather than be explicitly followed. It might be interesting to keep in initial considerations as the lack of transparency and other issues might be informative of other industries. However we need to be cautious because defining some criteria could be seen as justification for use of AW and this should be avoided. AW is a hot topic with the UN right now and recommendations could filter down to law enforcement.

Education: Maroussia brought up a case from Texas in May 2017, details to be found here: [www.courthousenews.com/wp-content/uploads/2017/05/houstonTeachers.pdf](http://www.courthousenews.com/wp-content/uploads/2017/05/houstonTeachers.pdf). Teachers have no way to know what's happening with the algorithm, why decisions are being made, criteria they are being judged on etc. which has been used to fire people. See also Weapons of Math Destruction by Cathy O'Neil (2016). It was also raised that student ratings of faculty show bias against women and minorities.

**Action: Maroussia to create new use case**

Ansgar suggested it may be worth sharing use cases with the transparency working group.

II. Stakeholder analysis (key actors: executive branch, legislative branch, regulators, companies, users, NGOs...?) – Pascal  
Pascal unable to join, update given by Matthew S. There is no existing SA taxonomy that fits, but the areas they are differentiating are:

- purpose/need for algorithm
- power to influence working of the algorithm
- awareness of downstream of effects
- impact received from the algorithm
- level of support

They are thinking about case studies and mapping out upstream stakeholder and downstream user impact, where stakeholders fit into the information chain, what are the active and passive effects on the way designed on the user.

They are working on a single use case at the moment to see the format and then this will be opened up to the other use cases.

**Action: Pascal/Mathew to talk us through the analysis of use case using SA in December meeting.**

III. Topic update: Bias Taxonomy (Matthew S)

The group has a call in the last week of November and will have some stakeholder mechanism candidates by then. Active google doc at <https://docs.google.com/document/d/1R2EsCWsZBdeb-cs-RJiPOw9COrjj3CqNMB1Utj2OLfc/edit> and they've had a meeting since the October monthly meeting. There are some comments to be addressed, and the bibliography/literature review is coming on well. Mapping to different areas (social science, law etc) is coming together.

*If anyone knows of a background or tender document around existing taxonomies of complex issues similar to bias in algorithms, or any literature, please let Matthew know. They are starting from scratch so examples would be useful, especially around a methodological approach. Next step is working on SA (point II) and continuing meetings between these monthly calls.*

IV. Topic update: Legal context (Maroussia)

Working on google docs to establish legislative framework in broad brush strokes, have a call with some legal scholars to advance research into LF.

Working to better articulate link between discrimination and procedural fairness. Need to contact transparency group to anchor the respective work better.

There was discussion about whether the overall goal was to identify bias or to counter bias. Positive bias may also trigger any safeguard as well as negative bias, is the standard robust enough to take this into account? Frank Pasquale suggested a need to look at existing conceptions of fairness to identify the preferred concepts. Matthew S asked where the neutral spot was in non-neutral technologies, and what we mean by unbiased. Ansgar pointed out the importance of criteria and justification of use, to see whether an algorithm system is acceptable in the particular society it is being used in. De-biasing: Martin Wattenberg, Fernanda Viégas, and Moritz Hardt (Google), *Attacking discrimination with smarter machine learning* (2016) Tolga Bolukbasi et al., *Man Is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings*, <https://arXiv.org/abs/1607.06520>  
Conceptions of fairness from a computer science point of view: <https://arxiv.org/abs/1609.07236>

Shrisha R asked if there are systems which are more likely to be biased, for example recommenders, and whether this has been looked at. Content moderation on social media was suggested, and Ansgar pointed out that algorithms derived from human behavior are likely to be the most at risk of being biased. Nozha B brought up the distinction between intentional and unintentional discrimination and it was suggested the impact was in unintentional discrimination.

Re: discrimination in Machine learning, activity in the chat pointed out that a lot of ML/deep learning algorithms are fundamentally unexplainable, for example we cannot say how neural networks reach certain answers <https://research.google.com/bigpicture/attacking-discrimination-in-ml/>

Ansgar suggested a team might work on mapping out algorithm uses that are more or less likely to need an analysis of bias. This is to be included in the taxonomy of bias work. *Anyone who wants to work on it please email Ansgar.*

V. Topic update: Psychology & Bias  
No members available to update

## **7. Progress with improving Working Group membership diversity**

Regional – Emails have been sent to various lists, including ICT for Development reaching out to non-US no-Europe parts of the world. We have some new members, who are thanked for joining. Ansgar has contacted Digital Asia Hub but no response yet, and some emails have been forwarded to groups in Brazil.  
*Please email any connections you have to Ansgar.*

## **8. Identifying and filling gaps in WG expertise**

Legal context is improving. There is a call next week with two legal scholars, and thanks to new members who have joined us. For other areas that might be useful, it was suggested that theoretical computer science would be useful.

**Action: Francien to suggest some Theoretical CS people who might join the group.**

In the chat it was suggested that function should be used for regular feedback and comments, instead of relying on the shaky audio connections that some struggle with. The (audio) floor can then be given to anyone with a longer update or discussion. A lot of other Working Groups do this and then the chat is posted in iMeet. *Any comments?*

## 9. Future Meetings

Friday, December 8<sup>th</sup>, 10:00am-11:30am (EDT)

Friday, January 12<sup>th</sup>, 10:00am-11:30am (EDT)

Thursday, February 15<sup>th</sup>, 10:00am-11:30am (EDT)

Wednesday, March 14<sup>th</sup>, 10:00am-11:30am (EDT)

Tuesday, April 10<sup>th</sup>, 10:00am-11:30am (EDT)

Monday, May 7<sup>th</sup>, 10:00am-11:30am (EDT)

Friday, June 8<sup>th</sup>, 10:00am-11:30am (EDT)

## 10. Adjourn

Meeting was adjourned at 11:35 AM

### Attendees:

| Last Name     | First Name | Employer/Affiliation     | Voting |
|---------------|------------|--------------------------|--------|
| Biermann      | BC         | Heavy Projects           | x      |
| Boujemaa      | Nozha      | Inria                    | x      |
| Chire         | Josimar    | ICMC                     | x      |
| Clifton       | Chris      | Purdue University        | x      |
| Coates        | Daphne     | Warwick Business School  |        |
| Courtney      | Patrick    | tec-connection           | x      |
| Dechesne      | Francien   | Leiden University        |        |
| Dowthwaite    | Liz        | University of Nottingham | x      |
| Egawa         | Takashi    | NEC Corporation          | x      |
| Ganesh        | Maya       | Leuphana University      |        |
| Hatada        | Yohko      | EMLS_RI                  | x      |
| Hense         | Peter      | Spirit Legal             |        |
| Jain          | Aman       | Samsung                  |        |
| Koene (Chair) | Ansgar     | University of Nottingham | x      |
| Levesque      | Maroussia  | Independent              | x      |
| Pasquale      | Frank      | University of Maryland   |        |
| Platcheck     | Ivsen      |                          |        |
| Rao           | Shrisha    | IIIT Bangalore           |        |

|          |         |                         |   |
|----------|---------|-------------------------|---|
| Rovatsos | Michael | University of Edinburgh | x |
| Sobey    | Charles |                         |   |
| Stender  | Matthew | Self                    | x |
| Bahn     | Christy | IEEE-SA (staff)         |   |