IEEE P7003 Working Group
Meeting Minutes
6th December 2018 /1:00 P.M. – 2:30 P.M. UTC
Teleconference

1. **Call to Order**
   The meeting was called to order at 13:08 UTC
   A quorum has been established and noted.

2. **Roll call and Disclosure of Affiliation**
   The list of attendees present is attached.

3. **Approval of December Agenda**
   Motion to approve the meeting agenda from 6th December 2018. 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 8 November meeting minutes**
   Motion to approve the minutes from November 8th, 2018. The minutes from the November 8th, 2018 meeting were approved without objection.

6. **Announcements**
   I.   Migration of draft into SA template (appropriate front matter, copyright etc.)
   The proper version is now available on the iMeet platform, and will be kept in that form as a word document; it is not to be put into Google docs due to formatting issues. As sections get their material to a state that is sufficient for the main document, that content will be submitted as a separate document to Ansgar and Liz, then the Technical Editors (Adam and Liz) will help to shape it into the format that is required for the template. At that stage if it is near the complete final text, there will be a vote on whether to have it included. The version at iMeet is to be considered the proper draft version.
   Action: As sections get to the stage of wanting to move content, get in touch so we can do it properly.

   Adam has created a high level flow diagram to help to make sure we capture what kind of information each section provides to the other sections. It provides an overview of how to use the standard, depending on whether you have a built system

already or not, what prior information you should have to use each section, and what the outputs are.

Ansgar emailed to ask people to fill in their availability and location for face to face meeting. London and NY venue agreed for 28ᵗʰ Feb and 1ˢᵗ March. There is also a Mumbai option. We need an overview of how many people are looking to attend at each location. *Question in chat about Europe – currently only London but if you cannot do it you can put it on the form.* There will be a twice a day video call for those who can't make it.

ISO now has a group on trustworthiness, developing a standard around risk assessment focused primarily on machine learning.

7. Updated Outline Discussion

   I.   Update on Use Cases
Raymond has a medical use case, pretty well developed but might edit before Christmas.
Allison also intends to add to the use cases.

   II.   Topic updates – 5min summaries

      i. Key concepts
Needs to wait for the main document to decide on what definitions etc need to be provided – group at beginning of 2019

     ii. Taxonomy (Matthew)
Work done of graphical overview of taxonomy and on draft on full text. Not quite ready to officially put them out. Looking at GDPR for identification categories, and flagging potential ways certain categories could bring about bias capacity, along several axes, including time and longevity of the algorithm. Could be interesting to compare notes with some of the other groups. End of next week will have some things to share – draft of full text, overview (including stages of algorithms and bias capacity), and deep dive into some of the demographic considerations. Next call to be scheduled, probably around this time next Wednesday.

     iii. Legal frameworks
Email update to Ansgar - that section is mainly there, material for version 2.0 is being shaped by Lorna and her team at the Human Rights Research Group. Expecting final version pretty soon, then they will be asking for comments from others.

     iv. Psychology
No update.
Raymond asked a question about what we mean and from what perspective. Concept behind having the section was related to the idea that bias related to algorithmic decision making might not be directly in the algorithm, might be in how people interpret the outcome. Psychological aspects of people designing system that might produce bias. Raymond offered to write about confirmation bias/automation bias.
Adam highlighted also confirmation bias in people building the system, and people annotating or dealing with the data.

v.   Cultural aspects (Kuba via chat)
"The cultural aspects section has already worked on the significant portion of our contribution but we want to work on normalizing the text so that it can be included in the standard document. We will be meeting with couple of people eager to help us with that on December 19th and will follow up in the next meeting"

vi.   System design stages (Alejandro)
They had a call recently. Discussion on level of granularity, and now have more understanding of the key areas and the level at which this should be pitched. Content needs to be added from various documents and meetings and another meeting is to be held soon.
Action: If anyone else wants to be on the call, let Alejandro know so you can be included in the doodle for the next one.
Potential for another insurance use case
Action: Alejandro to circulate the video of the presentation and if it fits well we can add it as a use case.
Adam suggested building use cases in Jupiter notebooks or similar, but unsure whether appropriate for an IEEE standard. Might be good for testing out implementations of the standards: fictional use cases with worked examples in code and datasets. Could be for reference within the group. Alejandro suggests sending an example of what this would look like, to see if it could be implemented.

vii.   Person categorization
No update

viii.   Representativeness of data (Adam)
No meetings but there have been edits in the document. Planning to look closer in January.

ix.   Outcomes evaluation (Adam)
Meeting since last WG meeting with a small number of attendees. Progress around metrics area but incomplete. Hoping that as the overall structure of the standard is evolved it will reinvigorate the focus.

x.   Evaluation of processing (Michael)
Nothing has happened recently. People need pushing to add content.
Action: Ansgar to follow up on people in this group

xi.   Resilience against manipulation (Michael)
As above.

xii.   Documenting & transparency
Haven't heard from the group for a while so we need an update.
Action: Someone from this group to please send an update as a matter of urgency.

Flow diagram: There has been quite a lot of confusion about the overlap between different areas of the standard. This provides a top level understanding of the context needed for all projects, requirements analyses for projects, and then more process sections. You may already have a built algorithm (whether by you or someone else) or you may be looking at

creating one, so where you start may change but the two main circles are iterative. Need inputs and outputs from each validated step.
Action: Everyone to examine document and see if they agree, need clarification, or want to add anything.

8. **Conferences and Whitepapers**
AIES conference submissions – insurance use cases not accepted. Any others?
IEEE AI and testing conference – Adam, Allison, and Marrousia submitted

9. **Any Other Business**
Discussion about Cathy O'Neil's ORCAA service (http://www.oneilrisk.com/) led by Raymond Bond
Service to rubber stamp algorithms and affirm that they are fair. It's not transparent how this is done, and one can't tell how they are audited. Is there a series of checks that we can think about using for our purpose?
Do we provide some kind of guidelines around which types of auditing certification services would be considered sufficient as a way of testing your data within the concept of being compliant with the standard? Should we look at some minimum requirements of any auditing tool that you use, or is this beyond the scope of the standard?
Formal accreditation? Becomes very complicated to give an official recommendation. We don't want to be in charge of accrediting people and testing whether they did actually comply.
Maybe those sites should say that they are compliant with P7003, that they are using the standard in their evaluations. If they are compliant then that implies some trust is worthy for that system.
The standard should have some general guidelines. A tool should not just use a simple accuracy metric, for example.
IEEE-SA is starting a separate initiative around certification, including algorithmic bias.

10. **Future Meetings**
Thursday, 11th January, 3:00pm – 4:30pm (UTC)
Thursday, 7th February, 3:00pm – 4:30pm (UTC)
28th February and 1st March P7003 face-to-face meeting TBC
Thursday, 4th April, 1:00pm – 2:30pm (UTC)
Thursday, 2nd May, 3:00pm – 4:30pm (UTC)
Thursday, 6th June, 5:00pm – 6:30pm (UTC)

11. **Adjourn**
The meeting was adjourned at 14:16 UTC

Attendees:

Last Name	First Name	Employer/Affiliation	Voting
Bond	Raymond		

Chaudhuri	Abhik	TATA Consultancy Services	**X**
Courtney	Patrick	tec-connection	**X**
Dechesne	Francien	Leiden University	**X**
Djeffal	Christian		**X**
Dowthwaite	Liz	University of Nottingham	**X**
Egawa	Takashi	NEC Corporation	
Hailey	Victoria	VHG	
Hatada	Yohko	EMLS_RI	**X**
Jessen	Hans	EY	
Koene (Chair)	Ansgar	University of Nottingham	**X**
Mandal	Sukanya	Self	**X**
McIntosh	Suzanne	New York University	
Nadel	Larry	NIST	**X**
Piwowar	Kuba	University of Social Psychology and Humanities, Warsaw	
Rovatsos	Michael	University of Edinburgh	**X**
Saucedo	Alejandro	The Institute for Ethical AI & Machine Learning	
Smith	Adam	Piccadilly Group	**X**
Stender	Matthew	Self	**X**
Bahn	Christy	IEEE-SA (staff)	