

**IEEE P7003 Working Group
Meeting Minutes
8 November 2018 /5:00 P.M. – 6:30 P.M. UTC
Teleconference**

1. Call to Order

The meeting was called to order at 17:02 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 November Agenda

Motion to approve the meeting agenda from 8th November 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 6 September meeting minutes

Motion to approve the minutes from September 6th, 2018. The minutes from the September 6th, 2018 meeting were approved without objection.

6. Announcements

- I. Proposed dates and places for P7003 Face-to-Face meetings at start of 2019
To get to know each other better and efficiently work through the draft outline and to answer the questions and comments that will have been listed by then. Offer from Ernst and Young that they could provide meeting space in London and New York and communicate between the two. Also from others in London, Berlin, and Mumbai. Suggested dates between 25th February and 1st March. No funding to support travel expenses, so multiple locations at the same time will be handy.
- II. Document needs to be put into IEEE-SA format, which has copyright notices and so on. This will happen in next week or so and needs a technical editor that will help to make sure new content keeps formatting etc.
Action: Anyone who wants to take on that role please get in touch with Ansgar and Christy. *Liz and Adam have expressed interest in the role*

7. Updated Outline Discussion

- I. Topic updates – Progress towards submission to main document and summary of desired feedback on submitted content
- It is important that the subsections add to the main document as soon as possible in order to start to identify overlap, contradictions, gaps etc. Only two sections have done this so far – legal and cultural aspects.
- Deadlines: submitting to working document, by end of this year/by next call. Ultimate deadline for standard is Q4 next year.
- Action: ALL subgroups to add their content to the main document and indicate what the current state of development is.**
- i. Taxonomy (not yet submitted) (Pascal)

A bit behind schedule. Did some pruning on existing document and made chart that visually represents the different types and stages of introduction of bias. Used some examples from the data justice lab in Cardiff. Content of document is pretty close to where they want to be, so should be in by next week or early week after that. As well as overlap, they need to go through all sections and pick up different ideas and terminology that other people have used – links to the section of key concepts and definitions which is planned.

Some annotations point to literature introducing a specific definition, which might not normally be included in the regular text, is it normal to have references in the document? How should examples for each part of the taxonomy be provided? References, appendix, or links? Suggested answer was that these should mostly go into appendix, and that citations for use cases can be useful. There is an exception for references to other standards which should be in line.
 - ii. Legal frameworks (submitted) (Marrousia)

International landscape and high level definitions. Human rights and big data ethics group are currently looking at it. Should be second iteration by next call. Might significantly change before next time so suggestion to look and get an overview but no need to comment on it as yet.
 - iii. Psychology (is anyone working on this?)

BC still interested but could not contribute to the call. If people want to do it, it really needs to be worked on. Suggestion to see what other sections are including to see if this is already being covered, eg what kind of overlap there is with cultural context? Currently not enough people to produce any kind of momentum on this section.
 - iv. Cultural aspects (submitted) (Arthur)

Text in the document is about how the section works rather than the final text. Considers that algorithms have their own culture and when this culture intersects with local culture (eg algorithms are designed in a particular cultural context) they operate a certain way, but what if they are being deployed in different cultures? What kind of bias do we see and how can we as designers ensure they are mitigated? How could designers work with local communities? There's a lot in this document, it could do with someone to guide which parts are most

important to the document as a technical standard, and which elements might be more useful as a supplementary document. Could do with one or two people to work on that – Arthur able to support. Currently it identifies problems and not always solutions; Kubar will add in some solutions into the document.

Mark U: Normative vs informative to help decide what goes in the document and what is an appendix; from chat:

Normative elements are defined in International Organization for Standardization Directives Part 2 as "elements that describe the scope of the document, and which set out provisions". Provisions include "requirements", "recommendations" and "statements". "Statements" include permissions, possibilities and capabilities. A "requirement" is an "expression in the content of a document conveying criteria to be fulfilled if compliance with the document is to be claimed and from which no deviation is permitted." It is not necessary to comply with recommendations and statements in order to comply with the standard; it is necessary to comply only with the requirements (that are denoted by the verbal form "shall"). There is much confusion between "normative" and "requirement", however the ISO terminology is supported by national standards bodies worldwide and is the legitimate description of these terms in the context of standards documents.

In standards terminology still used by some organisations, "normative" means "considered to be a prescriptive part of the standard". It characterises that part of the standard which describes what ought (see philosophy above) to be done within the application of that standard. It is implicit that application of that standard will result in a valuable outcome (ibid.). For example, many standards have an introduction, preface, or summary that is considered non-normative, as well as a main body that is considered normative. "Compliance" is defined as "complies with the normative sections of the standard"; an object that complies with the normative sections but not the non-normative sections of a standard is still considered to be in compliance.

Normative = prescriptive = how to comply

Informative = descriptive = help with conceptual understanding

Typically, normative is contrasted with informative (referring to the standard's descriptive, explanatory or positive content). Informative data is supplemental information such as additional guidance, supplemental recommendations, tutorials, commentary as well as background, history, development, and relationship with other elements. Informative data is not a requirement and doesn't compel compliance.

<https://en.wikipedia.org/wiki/Normative>

- v. System design stages (not yet submitted) (Chris C via chat)
I'm afraid the people working on system design stages have been swamped the last couple of months, and have made little progress. We have an outline and portions drafted, but still a bit of conflict resolution to be done before it makes sense to bring into the overall draft.

Alejandro: sent an email discussing overlap from other sections with the document. Suggested it would be good to have a call with relevant people from other sections once moved to the main document.

Ansgar: That's what these calls are for, so it would be best to put what exists into the main document for discussion. Can schedule into next agenda or a separate call.

Alejandro: happy to follow up via email and schedule a section on the next months call. For people who are interested all of the content is in the system design document.

- vi. Person categorization (not yet submitted)
No update.
- vii. Representativeness of data (not yet submitted)
No update.
- viii. Outcomes evaluation (not yet submitted) (Adam)
Content has now been moved into the main document. Would now like time to look at the rest of the document and see how it crosses over with this section. When looking at it, they tried to write it as a standard, not there yet, but closer to that point so need to see how it fits together.
Allison: flagging up overlap would be useful in case the group misses them.
Ansgar suggested colour coding of overlaps – each section title will be highlighted in a particular colour. If anyone notes an overlap in another section, use the relevant colour to highlight it.
- ix. Evaluation of processing (not yet submitted)
No update
- x. Resilience against manipulation (not yet submitted)
No update
- xi. Documenting & transparency (not yet submitted)
No update

Action: ALL to look through other sections and make comments/questions. If anyone identifies areas of overlap please highlight them in the relevant section colour (see point viii above)

- II. Update on Use Cases
Allison has quite a few use cases to provide. Will add titles in so we can see what we have.

8. Conferences and Whitepapers

- I. Who submitted to NIPS and/or AIES?

Quite a lot of people were working on conference papers in the last months. Deadlines for FAT*, various NIPS workshops, and AIES.

Adam, Allison and Maroussia – NIPS paper relating to regulatory frameworks related to data privacy in context of emerging standards. Accepted and working on final version.

FAT* submission on impact of using algorithmic decision making and big data and IoT more generally in insurance cases, building on insurance use cases. 24% acceptance rate, failed. Modified version submitted to AIES.

AIES on the development of user friendly tools for being able to do machine learning if you don't actually know anything about machine learning. What the implications of those are regarding ethical concerns of the use of ML systems. If you don't understand the principles you might start introducing bias etc. by using particular datasets.

9. Any Other Business

Mark: Working on a definition of transparency for NIST. The meaning of having an algorithm being transparent is highly context dependent, has anyone worked on this before and point in useful directions to help? Example use case: if you're in a clinical decisions support setting and the algorithm says you should prescribe drug x and you want an explanation. That transparency might be useful to an oncologist, but not the patient or radiologist. Can we operationalize this in a systematic way to be used in a standard? Please email dark@computer.org if you can help.

Marroussia responded that we've been trying to tackle this for a long time – link between explanation that is accessible to the user however expert or lay, as a prerequisite to articulate bias. Proposed Sandra Wachter's work might help, offering user-centric explanations and considering the slightest change a user could make for a different outcome:

<https://jolt.law.harvard.edu/assets/articlePDFs/v31/Counterfactual-Explanations-without-Opening-the-Black-Box-Sandra-Wachter-et-al.pdf>

Also ongoing technical research in deep learning applications for eg recognizing breast cancer, system highlights main criteria for reaching the decision. Can't talk about bias and equality if we can't have an explanation. Useful to crosslink with working groups that are already looking at this – P7001 is the transparency standards WG.

Other links shared in chat during discussions:

<https://kbros.co/2PMS6fkmy> "Codes of Ethics and the Ethics of Code"

<https://www.adlnet.gov/projects/ps4tla> "Privacy Support for the Total Learning Architecture"

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2920883 "Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System"

10. Future Meetings

Thursday December 6th, 13:00UTC

11. Adjourn

The meeting was adjourned at 18:10 UTC

Attendees:

Last Name	First Name	Employer/Affiliation	Voting
Antosca	Albert R		
Chaudhuri	Abhik	TATA Consultancy Services	X
Clifton	Chris	Purdue University	X
Costley	Jennifer		
Courtney	Patrick	tec-connection	X
Djeffal	Christian		X
Dowthwaite	Liz	University of Nottingham	X
Farley	Tim	Insitu, Inc.	
Gardner	Allison	Keele University	X
Gautam	Sumit	LG Electronics	
Gwagwa	E Arthur	Strathmore University	X
Hatada	Yohko	EMLS_RI	X
Jurgens	Pascal	U of Mainz, Germany	X
Koene (Chair)	Ansgar	University of Nottingham	X
Lévesque	Maroussia	Independent	X
Mandal	Sukanya	Self	X
Nadel	Larry	NIST	X
Rovatsos	Michael	University of Edinburgh	X
Saucedo	Alejandro	Self	
Smith	Adam L	Piccadilly Group	X
Stender	Matthew	Self	X
Underwood	Mark	Synchrony Financia/Krypton Brothers	X
Vidal	David	Idx	
Weger	Gerlinde	Independent	X
Bahn	Christy	IEEE-SA (staff)	