

Privacy in Proteomics Data

HUPO 2016 Bioinformatics Hub Mini-Pitch

Henning Hermjakob

European Bioinformatics Institute / Beijing Proteome Research Center

hhe@ebi.ac.uk

- In Genomics, management of personally identifiable data is a major topic
 - Data is often managed by resources like dbGaP or EGA, providing access only after (often lengthy) authorisation
- In Proteomics, the topic is only beginning to emerge:
 - Li S, Bandeira N, Wang X, Tang H. On the privacy risks of sharing clinical proteomics data. AMIA Jt Summits Transl Sci Proc. 2016 Aug 31;2016:122-31. eCollection 2016. PubMed PMID: 27595046
 - Parker GJ, ..., Leppert M. Demonstration of Protein-Based Human Identification Using the Hair Shaft Proteome. PLoS One. 2016 Sep 7;11(9):e0160653. doi: 10.1371/journal.pone.0160653. eCollection 2016. PubMed PMID: 27603779.
 - <https://twitter.com/search?q=nih%20proteomics%20policy&src=typd>
- Motivation: Keep data as open as possible
- Are there fundamental differences for proteomics?
- Are we just waiting how things develop, or do we want to develop a more proactive approach, for example in the context of HUPO?
- Example: Can we maintain more open access to data through privacy preserving interfaces?