



The code-of-conduct process

Data management
in speech research



Motivation

Why write a “code of conduct”?

- No agreement on principles.
 - Every researcher, funding agency, university, company and political body seem to have their own interpretation.
- No agreement on concepts.
 - What is an acceptable level of anonymization?
 - Do we have to allow subjects to withdraw consent?
 - What do we mean by identifiability?
- No coherence = confusion
- Individual researchers at risk.
 - Possible legal action against researcher(s)
 - Lost opportunities (e.g. wrongly rejected applications and publications)
- Users and subjects at risk.
 - Possible ethical violations, identity theft, humiliation, violation of privacy
 - Bad user experience
- Organizations at risk.
 - Possible legal action against organizations
 - Potential media scandals, public outcry

What is a “code of conduct”?

“An association code of conduct is a set of rules which is commonly written for members of an association, which protects the community and informs the members of the association’s expectations. It is appropriate for even the smallest of associations to create a document containing important information on expectations for members. The document does not need to be complex or have elaborate policies.” (adapted from Wikipedia)

- Community standard codifies the-best-of-our-knowledge
 - If we don’t know, then nobody can be expected/required to know
 - Protect all parties (researchers, subjects, funders, employers)
 - Enable best possible research within ethical standards
- It is about *values*
 - What is important to us?



Objectives

Target groups and objectives

1. Researchers and research teams

What do I need to do when?

- Using data
- Collecting data
- Publishing data
- Publishing results
- Designing research plans
- Designing data management plans

2. Reviewers

What do we need to require from?

- Publications
- Data management plans
- Research plans

3. Administrative bodies

What are the responsibilities of?

- Conferences, editorial boards, funders, universities, companies, NGOs etc.

Design objectives

- Start with why, and stick to it
 - Why is this important?
 - Not just rules, but also the arguments and the shades of gray
 - Every part reflects back to motivation
- Useful and easy to use
 - Provides what is needed
 - Compact
 - Easy to read; Plain language, no legalese, not too technical etc.
- Only that which makes speech special
 - Generic data practices still apply
- Provides
 - Lists of potential threats and their potential remedies
 - Templates for typical use cases = if this then that
- Time scope
 - Continuous and never ending
 - Includes expiry date as well as mechanisms for review and update
 - Reflects current best-of-our-knowledge
 - Publish first draft (or recommendations) when done, not sooner

Draft of Process

Process, high level

- From general to specific & back-again
 - Define our values
 - Define objectives of document
 - Choose sub-areas (e.g. mindmap)
 - Sub-area specific questions
 - Generalize and unify where possible
 - Define terminology
 - Iterate
- Parallel tasks
 - Find relevant background material
 - Identify and keep relevant organizations in the-loop
- Involvement
 - Draft with working group
 - Comments from SIG
 - Request comments from wider audience (ISCA, IEEE, EU etc.)
 - Iterate until done
- For later (long-term goals)
 - Presentation at Interspeech for “general public”
 - Make it an ISCA “standard”
 - Propose for European Data Protection Supervisor (EDPS)

Process within working group

- Monthly working group meeting on Zoom
 - Anyone have a Jitsi server?
- Progress reports to SIG every 6 months
 - E.g. Seminar or webcafe
- Require progress, but no DL
 - Keep moving but no stress
- Drafting on Google-docs or similar
 - Open and private alternatives?
- Transparent and inclusive



Discussion
