Consent Framework Crosswalk (v1)

Establish a FAIR-aware patient consent framework

Corpas M, Kovalevskaya NV, McMurray A, Nielsen FGG (2018) A FAIR guide for data providers to maximise sharing of human genomic data. PLoS Comput Biol 14(3): e1005873. https://doi.org/10.1371/journal.pcbi.1005873

Consent frameworks dictate the extent to which human genomic data can be accessed and reused. Ensuring appropriate consent to collect genotype, phenotype, and any other type of human data is achieved will usually be the responsibility of the principal investigator (PI) overseeing the study. Data archives and repositories will be required to check that the consent forms of deposited datasets specify the goals of the immediate project. It is essential to explicitly describe in clear terms if the data is intended to be shared beyond the current scope of the project (i.e., general research use). If wider data sharing is intended, the consent form should set out potential risks and benefits to participants, as well as any data anonymisation procedures to be undertaken. Consent frameworks require special considerations from the data producer's point of view, given their extreme variability. To allow standardisation of consent frameworks, GA4GH has developed consent codes that facilitate the integration of distinct consent types across different legal systems [22].

The level of anonymisation that will be applied to the data should be clearly explained in consent forms, since different levels are possible. Participant consent requirements should be considered prior to data collection, alongside approval from an IRB. Different research questions may necessitate variable degrees of identity exposure by study participants. For example, the Personal Genomes Project (PGP) provides complete access to study participants' identities and phenotypic traits [23] under a Creative Commons Zero (CC0) license waiver [24]. This radically open consent framework is, however, a highly unusual one for clinical genomic data. NIH-funded studies require third-party researchers to submit a Data Access Request describing how they intend to use the data. A Data Use Certification Agreement is then produced, which must adhere to the NIH Genomic Data Sharing Policy's ethical principles governing data access and privacy safeguards [25]. In the UK, GeL consent forms are classified according to whether patients are affected with cancer or rare diseases, with the consent framework allowing access to summary statistics in a controlled environment to authorised users [26].

Patient data sharing consent frameworks vary country to country, funder to funder, and study to study. We thus suggest that, for interoperability purposes, data sharing consent frameworks adopt existing standards for digital consent formats and include, at a minimum:

1. Goals of the current research project and why data generation/sharing is being carried out.
2. Potential risks to the individual participant from the (ab)use of the data.
3. Confirmation that these issues have been discussed in person, with the individual and/or guardian involved in signing the form.
4. Contract of data access for the current research project and the extent to which the data custodian commits to make the data findable, accessible, interoperable, and reusable for future research projects.

Some consent forms (e.g., PGP or Genomes Unzipped [27]) may make the patient/donor's identity known. Others require the identity of research participants anonymised. The Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) database [28], a provider of anonymised human copy number variation (CNV) data and phenotypes (not datasets), offers a consent framework compliant with European Union guidelines for clinical sharing [29], allowing anonymous sharing of genomic and phenotypic data of patients. At all events, it is always advised that ethical and genetic counselling experts are consulted when choosing the appropriate consent form.

Establish a FAIR-aware social media data consent framework

Consent frameworks dictate the extent to which social media data can be accessed and reused. Making good faith efforts to ensure consent to collect and/or provide access to social media data is the responsibility of the institution overseeing the collection. Libraries, archives, and repositories must specify the goals of collecting effort. It is essential to explicitly describe in clear terms the scope of anticipated collection use (i.e., research on mass shootings, crisis response). The institution should make good faith efforts to share potential risks and benefits with individuals represented in the collection, as well as any planned data anonymization, cleaning, subsetting, and/or restructuring. Consent frameworks require special considerations from the social media data producer's point of view, given variability in data producer knowledge of risks and benefits that come with long term access to data they produce. To support standardization of consent frameworks, the Society of American Archivists, the Digital Library Federation, Open Repositories, the Association for Research Libraries, the Rare Books and Manuscripts Section of the Association of College and Research Libraries, and the International Federation of Library Associations and Institutions will work to develop consent codes that facilitate the integration of distinct consent types across different legal systems.

If pursued, the level of anonymization applied to social media data should be clearly explained, since different levels are possible. Consent requirements should be considered prior to data collection. To date, IRB generally grants exemptions to efforts that collect social media data. While an exemption may be granted, it is the responsibility of the institution overseeing collecting activity to address possible risks to individuals represented in social media data collections.

Consent frameworks vary country to country, institution to institution. We thus suggest that data sharing consent frameworks work to develop and/or adopt existing standards that include, at a minimum:

1. Purpose of the collection and why a particular institution is collecting.
2. Potential risks and benefits to individuals represented in the collection.
3. Confirmation that good faith efforts have been made to share potential risks and benefits with individuals represented in the collection, with an option for the individual and/or guardian to weigh in on consent within a reasonable period of time.
4. Contract of data access for the collection and the extent to which the data custodian commits to make the data findable, accessible, interoperable, and reusable.

It is advised that any ethical expertise or guidelines influencing an institution's approach to consent are made publicly accessible.

Thomas Padilla, 6.26.2018, CC-BY-SA