FAIR GENOMES: Standardizing a meta-data schema for FAIRifying personal genome data workflows

Gurnoor Singh[1*], K. Joeri van der Velde[2*], Jeroen Beliën[4], Jasmin Böhmer[3], Daphne Stemkens[5], Lisenka Vissers[1], Jeroen van Reeuwijk[1], Saskia Hiltemann[7], Lennart F. Johansson[2], Nienke van der Stoep[6], Daoud Sie[4], Janneke Weiss[4], Geert Frederix[3], Marco Roos[6], Erik van Iperen[8], Terry Vrijenhoek[3], Folkert W. Asselbergs[3], Joris van Montfrans[3], Rolf Sijmons[2], Hanneke van Deutekom[3], Pieter Neerincx[2], Fernanda de Andrade[2], Anna Niehues[1], Hindrik H.D. Kerstens[10], Mark Thompson[6], Rajaram Kaliyaperumal[6], Annika Jacobsen[6], Katy Wolstencroft[6,14], Ies Nijman[3], Marcel Nelen[1], Ariaan Siezen[1], Koen ten Hove[1], Nine Knoers[2], Christian Gilissen[1], Hans Scheffer[1], Stefan Willems[3], Wendy van Zelst-Stams[1], Helger Ijntema[1], Kim Elsink[3], Bart de Koning[9], Bauke Ylstra[4], Erik Sistermans[4], Patrick Kemmeren[10], Henne Holstege[4], Christine Staiger[11], Bastiaan Tops[10], Susanne Rebers[12], David van Zessen[7], Valesca Retèl[12], Edwin Cuppen[13], Peter van Tintelen[3], David van Enckevort[2], Lieneke Steeghs[1], Salome Scholtens[2], Jeroen Laros[6], Leon Mei[6], Cor Oosterwijk[5], Andrew Stubbs[7], Peter A.C. 't Hoen[1], Mariëlle van Gijn[2], Morris Swertz[2]

GENOME DATA IS UNFAIR

The increase in personal genome data generated in diagnostics and research holds great promise for advancing personalized prevention and medicine. However, valuable genomic and associated clinical data is fragmented across many healthcare providers and research organizations, making it difficult to find and reuse. This prohibits us from exploiting the potential information contained within these genomes for health benefit.
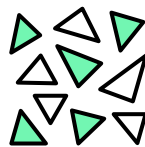
TOGETHER WE FAIRIFY

In multidisciplinary face-to-face and video conference meetings, we work towards defining a schema for genome data. This schema consists of common and optional data elements divided into five categories: general information, personal and clinical information, material information, technical information, and informed consent, enabling institutes to talk about genomes in 'the same language'.

BUILDING A SEMANTIC SCHEMA

By talking the same meta-data we increase **F**indability and **A**ccessibility of genomes, with additional technical details leading to better **I**nteroperability and **R**eusability of the data. The schema now consists of around 60 data elements with relationships and values mapped to existing ontologies such as SNOMED, DUO, HPO, UMLS and EDAM. It provides a strong basis for digital twin data in Dutch hospitals, development of personal genetic lockers, and active Dutch participation in the European '1+ Million Genomes' Initiative.
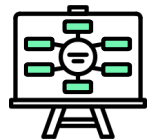
JOIN OUR INITIATIVE

The FAIR Genomes meetings are open to receive input from anyone to achieve the highest quality and usability of the resulting meta-data framework. Join FAIR Genomes at:
https://github.com/fairgenomes
or contact to join WP2/WP3:
Gurnoor.Singh@radboudumc.nl
K.J.van.der.Velde@umcg.nl
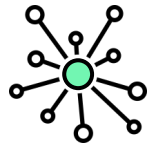
1 = Radboud University Medical Center, Nijmegen, The Netherlands
2 = University Medical Center Groningen, The Netherlands
3 = University Medical Center Utrecht, The Netherlands
4 = Amsterdam University Medical Centers, location VUmc, NL
5 = VSOP - Dutch Patient Alliance for Rare and Genetic Diseases
6 = Leiden University Medical Center, The Netherlands
7 = Erasmus Medical Center, Rotterdam, The Netherlands
8 = Durrer Center for Cardiovascular Research, Utrecht, The Netherlands
9 = Maastricht University Medical Center, The Netherlands
10 = Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands
11= Dutch Techcentre for Life Sciences, Utrecht, The Netherlands
12 = Netherlands Cancer Institute, Amsterdam, The Netherlands
13 = Hartwig Medical Foundation, Amsterdam, The Netherlands
14 = Leiden Institute for Advanced Computer Science, Leiden University, Leiden, NL

ZonMw