The FAIRification of data and the potentialities of FAIR resources showed, in practice, at the Rome Bring Your Own Data workshop

Claudio Carta[1], Marco Roos[2], Annika Jacobsen[2], Rajaram Kaliyaperumal[2], Mark Thompson[2], Mark D. Wilkinson[3], Ronald Cornet[4], Andra Waagmeester[5], David van Enckevort[6], Mascha Jansen[7], Luana Licata[8], Allegra Via[9], and Domenica Taruscio[1]

[1] Istituto Superiore di Sanitá, Italy
`{claudio.carta,domenica.taruscio}@iss.it`
[2] Leiden University Medical Centre, The Netherlands
`m.roos@lumc.nl,annika.jacobsen.86@gmail.com,rajireturn@gmail.com,`
`mark9630@gmail.com`
[3] Centro de Biotecnologa y Genmica de Plantas UPM-INIA, Spain
`mark.wilkinson@upm.es`
[4] Academic Medical Center - University of Amsterdam, The Netherlands
`r.cornet@amc.uva.nl`
[5] Micelio,Belgium
`andra@micelio.be`
[6] University Medical Centre Groningen, The Netherlands
`david.van.enckevort@umcg.nl`
[7] Dutch Techcentre for Life Sciences, The Netherlands
`mascha.jansen@dtls.nl`
[8] University of Rome Tor Vergata, Italy
`luana.licata@gmail.com`
[9] National Research Council, Italy
`allegra.via@uniroma1.it`

Abstract. It is widely agreed that rare disease patient registries should be international and follow the guiding principles of Findable, Accessible, Interoperable, Reusable for humans and computers (FAIR). Furthermore, the procedures to collect and exchange data should be harmonised. Four years ago we organized our first Bring Your Own Data (BYOD) workshop in Rome with the aim to promote the establishment of FAIR rare disease registries in compliance with IRDiRC and EU recommendations[4, 1]. Since 2014, BYOD workshops have been organized every year by - and held at - the National Centre for Rare Diseases - Istituto Superiore di Sanitá (CNMR-ISS), Rome, Italy with the support of RD-Connect and ELIXIR, particularly the Dutch Techcentre for Life Science representative of ELIXIR-NL [1, 2]. The general roadmap of the BYOD workshop contains at least a preparatory phase, an execution phase, and a follow-up phase to foster the results of the workshop by surveying and phone conferences with participants[3]. Here, we describe the 4th edition of the BYOD to link Rare Disease Registries held in September 2017 at CNMR-ISS in Rome, which included twenty eight

participants selected on the basis of participants background, role with reference to registry activities, and involvement in European Reference Networks (ERNs). This edition has been done in collaboration and/or with the support of several organisations such as RD-Connect, ELIXIR, RD-Action, EURORDIS, EpiRare, ERNs, and ICORD This year the focus was on (i) FAIRification process, with partially prepared semantic models, (ii) FAIR data tools, and (iii) FAIR data management. A new element was a semantic model sketching exercise with the registry managers. At the beginning of the workshop, an overview of the FAIRification process and an introduction to four selected, scrambled, and anonymous sample datasets were provided. In the following sessions, the BYOD became highly interactive and the participants, split in four groups, went step by step through the process of FAIRification with break-out sessions alternated to plenary sessions in which participants presented the results of their group work. In the first step, each group had to discuss and draw a conceptual model of their assigned dataset, followed by a plenary session to discuss commonalities and complementarities. In the second step, the groups created an ontological model based on the conceptual model from the previous step which required ontology search. In the third step, the OpenRefine-FAIRifier tool and metadata editor were used to aid in the creation of the machine readable format of the selected data. Next, to show the potential of data linkage, a cross resource question was executed over the four FAIRified sample datasets. The final part of the workshop was focused on a discussion about FAIR data management. Based on our experience from the previous editions of the workshop, the novel elements are: (i) in the preparatory phase only four heterogeneous datasets among those received were selected, ii) a breakout and plenary session at the end of each FAIRification step were introduced, (iii) a time slot allocated for the self sketching the semantic model with pen and paper by the registry managers, (iv) a time slot was allocated to discuss FAIR data management and FAIR project planning. We conclude that the main objectives of the BYOD for registry managers were achieved. The participants were able to see the potentiality of a FAIR registry and how this allowed to get more rapid responses to cross-resource questions improving the use of available information on rare diseases and speed up research. Furthermore, participants had the chance to get more into the FAIRification process and understand the importance of data management planning when setting up a registry. Finally, it is important to stress that the friendly environment and the high interactivity among the different players, i.e., data managers, researchers, clinicians, patient representatives, and IT-trainers, strongly contributed to make the BYOD workshop a success.

Keywords: Rare Disease Registries, FAIR data principles, BYOD

tative of ELIXIR-NL and ELIXIR-IIB), RD-Connect, RD-Action, EURORDIS, ERNs, EpiRare, and ICORD.

References

1. Carta, C., Roos, M., Gainotti, S., Lopes, P., Thompson, M., Kaliyaperumal, R., van der Horst, E., Queralt Rosinach, N., Thompson, R., Taruscio, D.: Linked data approach: some results and considerations from the first RD-Connect Bring Your Own Data meeting in Rome. In: European Human Genetics Conference. p. 313. Nature Publishing Group, Glasgow (2015), `https://www.eshg.org/fileadmin/www.eshg.org/conferences/2015/downloads/ESHG2015_Abstracts_May_19.pdf`
2. Carta, C., Roos, M., Kaliyaperumal, R., Gainotti, S., Torreri, P., Taruscio, D.: RD-Connect workshop for patient registries in Rome: a Bring Your Own Data for beginners on data linkage and ontologies. In: European Human Genetics Conference. p. 344. Nature Publishing Group, Barcelona (2016), `https://www.eshg.org/fileadmin/www.eshg.org/conferences/2016/downloads/ESHG2016_Abstracts_final.pdf`
3. Jansen, M., Carta, C., Roos, M., Bonino da Silva Santos, L.O.: The organisation of Bring Your Own Data (BYOD) workshops to make life science data linkable at the source. In: Semantic Web Applications and Tools for the Life Sciences (2016), `http://www.swat4ls.org/wp-content/uploads/2016/10/paper-18.pdf`
4. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., Roche, D.G., Kruuk, L.E.B., Lanfear, R., Binning, S.A., Bechhofer, S., Benson, D.A., Berman, H., Henrick, K., Nakamura, H., Wenger, M., Crosas, M., White, H.C., Carrier, S., Thompson, A., Greenberg, J., Scherle, R., Lecarpentier, D., Martone, M.E., White, E., Sandve, G.K., Nekrutenko, A., Taylor, J., Hovig, E., Wolstencroft, K., Bauch, A., Sansone, S.A., González-Beltrán, A., Maguire, E., Sansone, S.A., Rocca-Serra, P., González-Beltrán, A., Harland, L., Groth, P., Berman, H.M., Bourne, P.E., Berman, H.M., Watenpaugh, K., Westbrook, J.D., Fitzgerald, P.M.D., Rose, P.W., Kinjo, A.R., Gutmanas, A., Starr, J., Musen, M.A.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, 160018 (mar 2016), `http://www.nature.com/articles/sdata201618`