Providing Recommendations of Services, Datasets and End-Users in the Data Market Austria (DMA)

Matthias Traub * – mtraub@know-center.at
Heimo Gursch * – hgursch@know-center.at
Dominik Kowald * - dkowald@know-center.at
Dieter Theiler * - dtheiler@know-center.at
Roman Kern ** – rkern@know-center.at
Elisabeth Lex ** – elisabeth.lex@tugraz.at

* Know-Center GmbH – Research Center for Data-Driven Business & Big Data Analytics, Inffeldgasse 13/6, 8010 Graz

** Institute of Interactive Systems and Data Science, Graz University of Technology, Inffeldgasse 13/6, 8010 Graz

Abstract

The Data Market Austria (DMA) [1] is a nationally funded lighthouse project to create a digital ecosystem i.e. a multi-sided market for shared datasets and data services. In its final expansion stage, DMA is intended to be a central hub for all actors participating in the Austrian data economy, regardless of their industry sector. Comparable initiatives target limit themselves to only a single specific domain, for example the Industrial Data Space in Germany [2].

Since the data-driven economy is a complex environment, participants in DMA are assigned diverse roles depending on the value they are able to offer: data-provider, service-provider, consulting-service-providers, and infrastructure-provider.

The first role is the data-provider, assigned to contributors that offer datasets. These datasets are not limited to any specific domain or structure, but all contributions can be offered on DMA. Through the use of innovative and intelligent cloud-based approaches, DMA will allow end users to process and analyse proprietary data without actually having to directly access or "see" the data. The next vital part in DMA are algorithms and services to process data, where there are also no restrictions on the type or complexity. Participants with the role of service-providers will define innovative new services utilizing multiple sources of data and analytics capabilities available in the DMA. Complementary to the mentioned roles are individuals and companies offering consulting services. These partners support the implementation of algorithms, inspection and judgment of datasets, expertise in finding matching pairs for datasets and algorithms, or any other technical or business-related consulting. Finally, to cater for the mentioned roles, another vital player in the DMA ecosystem are infrastructure providers, offering storage and computing power to host the datasets and allow for an execution environment for the algorithms.

An obstacle for the creation of a new data value chains [3] is the identification of all necessary partners. Since a data value chain may have multiple data sources, multiple algorithms and might be executed in a distributed manner even over different infrastructures, it is non-trivial coordination task to identify the best possible combination. In DMA, a personalized recommender system will support users by automatically suggesting meaningful combinations of datasets, algorithms, consulting, and infrastructure.

Recommendations in DMA are primarily based on metadata and interactions [4] of services, datasets and users, which create a context and enable matchmaking. In combination with the needs of each user, personalised recommendations for collaboration are generated. DMA will extend this matchmaking functionality to incorporate insights about user interaction with the DMA Portal and also about implicit and explicit feedback [5] by the user (explicit feedback is given directly by the user, e.g., by rating, while implicit feedback comes non-intrusively from interactions such as clicks). The recommendations will be generated using our scalable recommendation framework ScaR [6].

Currently the system is in the first stage of the prototype and a closed beta will be launched by the end of 2018. Therefore, it is necessary to setup a structured online A/B test in order to have initial results of which type of algorithm might suit the users needs. For this we are aiming to compare three models, most popular items as a baseline, context-based CF as well as content-based recommendations. In the end we propose to alter the recommendation strategies to be adaptive and incorporate the feedback into the process, which leads to a system capable of learning and adapting to current needs and situations.

In order to transfer our findings from DMA on the European level, we will integrate them into the recently approved H2020 flagship project AI4EU [7]. In the course of this project, we will contribute to the development of a platform that acts as a broker, developer and one-stop shop providing and showcasing services, expertise, algorithms, software frameworks, development tools, components, modules, data, computing resources and prototyping functions of AI in Europe.

In the course of this workshop, we would therefore demonstrate our planned evaluation setup with the used recommender algorithm as well as the underlying data. It would be very helpful for us to get feedback from the community in order to make correction and include it into our future work.

References:

[1] http://www.datamarket.at

[2] http://www.industrialdataspace.org/

[3] Curry, E. (2016). The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches. In New Horizons for a Data-Driven Economy.

[4] Damiani, E., Ceravolo, P., Frati, F., Bellandi, V., Maier, R., Seeber, I., & Waldhart, G. (2015). Applying recommender systems in collaboration environments. Computers in Human Behavior.

[5] Jawaheer, G., Weller, P., & Kostkova, P. (2014). Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback. ACM Transactions on Interactive Intelligent Systems (TiiS).

[6] http://scar.know-center.tugraz.at/

[7] https://ai4eu.org/