

Goal-Oriented Modelling of Relations and Dependencies in Data Marketplaces

Arnab Chakrabarti^{1,2}, Christoph Quix^{1,2}, Sandra Geisler¹, Jaroslav Pullmann¹, Artur Khromov¹, Matthias Jarke^{1,2*}

¹ Fraunhofer Institute for Applied Information Technology FIT, Germany
² Databases and Information Systems, RWTH Aachen University, Germany
¹firstname.lastname@fit.fraunhofer.de ²lastname@dbis.rwth-aachen.de

Abstract. Data exchange between companies is becoming more important in a digitized economy. Business models are created for providing, enriching, or using big data in various domains. Although there are some success stories in this area, companies are still struggling to define goals and strategies for a successful participation in a data marketplace. Especially, companies from classical business domains such as automotive, mechanical engineering, or life sciences could benefit from a secure and trusted data exchange as it supports data-driven business processes. In this paper, we model goals and strategic relationships of actors of a data marketplace using i^* . The model can be used as a blueprint by a company to identify their strategy and to determine the objectives to establish a successful data-driven business. We also presented the model for our case study from the medical domain.

1 Introduction

In a data-driven economy, the availability of high quality data is crucial for many businesses. Although there are clear benefits of exchanging data for the economy and society in general, organizations and individuals are still careful in providing data to other entities as they are afraid of losing control over their data. This applies to industrial environments as well as to medical contexts, where especially patients are concerned about data sovereignty and privacy. A platform for a secure and trusted data exchange could address these concerns by providing a trustworthy data space in which users define the terms and the conditions of use for the data provided. The *Industrial Data Space*³ (IDS [4]) is an initiative in Germany to create such a platform, in which participants can exchange data securely and still keep the control over their data and maintaining their data sovereignty. For the development of business models for the participants of the IDS, it is important to analyse the goals and dependencies of the various stakeholders. The contribution of this paper is the application of the goal-oriented modelling approach of i^* [9] to the data exchange setting of the IDS. We provide a generic i^* model based on the general concepts of the IDS, as they have been

* Copyright 2018 for this paper by its authors. Copying permitted for private and academic purposes.

³ <http://www.industrialdataspace.de>

defined in the business architecture of IDS [6]. Goal-oriented modelling is a common approach during the creation of a business model. For example, a process for the integration of business models and goal models has been presented in [2]. Paja et al. in their work [5] analysed the applicability of various goal modelling techniques for strategic decision making and evaluated the usability of such techniques on a realistic case study. Business modelling with i^* has also been done in various approaches. The integration of technology-oriented models and business models is discussed in [3]. Reasoning about business models that are formalized in i^* is done in [7]. Samavi et al. address especially disruptive business models which might change frequently and thus, require a continuous evaluation of the goals, intentions, and roles. In the medical domain, i^* has been successfully used to model a wellness tracking system for patients [1]. Thus, it was beneficial in identifying the roles of users and the dependencies among them at a nascent stage of the development for the IDS infrastructure.

2 Modelling the Industrial Data Space using i^*

In this section we present the goal-oriented requirements analysis for the IDS using the i^* modelling approach. The business architecture of IDS as presented in [6] fails to provide an answer to the following questions:

- How are the stakeholders inside the IDS dependent on each other?
- What are the major goals that the stakeholders want to achieve?
- Which specific tasks do the actors need to perform to achieve these goals?
- What are the alternatives for a stakeholder to achieve a particular goal?

In order to answer these questions, we do an early requirements engineering in form of the i^* model. We identify the goals and the subgoals for the IDS and use them for modelling the Strategic Dependency (SD) diagram and the Strategic Rationale (SR) diagram (Fig. 1 and 2).

The following goals and the soft goals were identified for IDS and were used for modelling the SD diagram (Figure 1) and the SR diagrams (Figure 2).

Below we are going to define the actors and describe the dependencies in detail as depicted in the models above. The identified actors within the IDS are: *Data Consumer*, *Data Provider*, *Broker*, *Data Owner*, *Data User*, *Clearing House*, *Identity Provider*, *App Store Provider*, *App Store*, *Vocabulary Manager*.

- *Data Usage Control/Data Sovereignty*(Soft-Goal): The data owner in the IDS may apply usage restrictions to its data which is being transmitted to the data consumer. Data Sovereignty can be further classified into sub-goals like handling permissions and ensuring control over data usage.
- *Secure Data Exchange*(Soft-Goal): The IDS aims at enabling secure data exchange between the data providers and data consumers, ranging from the source the data originates from (e.g., a sensor on an IoT device) to the actual point of use (e.g., an industrial smart service for real-time analysis).

- *Data Governance*(Soft-Goal): Data governance is supported (i) by establishing trustworthy relationships between data owners, data providers, and data consumers; (ii) by offering a decentralized architecture that is independent of any central authority; and (iii) by aiming at transparency and traceability of data exchange and data use.
- *Data Provenance*(Soft-Goal): The IDS should track the lineage of the data and provide data ownership; thus, ensuring data provenance.
- *Data Economy*(Goal): The IDS App Store provides data-driven services and publishes them as data apps (i.e., software components providing dedicated data-related service functionality). The IDS participants can request these data apps from an app store.
- *Data Exchange*(Goal): The main and the most significant goal of the IDS is to provide a data exchange platform between the data providers/data owners and the data consumers/data users, but this exchange takes place when the above mentioned goals are satisfied.

The SD Model: In the SD model, we map the dependencies between the actors that are identified within the IDS. For example in Fig. 1, we can see that the *Data Consumer* is dependent on the *Data Provider* having two types of dependencies: *Exchange Data*(goal-oriented dependency) and *Data*(resource-oriented dependency). As another interesting example, we can see the dependencies between the *Clearing House* and the *Data Consumer*. The Data Consumer is dependent on the Clearing House for *Data Governance* (soft-goal dependency), for *Logging Transaction* (task-dependency), and to receive *Confirmation Reports* (resource-dependency). Whereas the *Clearing House* is dependent on the *Data Consumer* for the successful achievement of *Confirm Reception*. Here it is important to note that by modelling such dependencies, we can identify cycles which indicate viable dependencies. For example, in the described scenario we can say that the Data Consumer dependencies are likely to be viable and in case a dependency fails, this failure can propagate along a chain of other dependencies within the model.

SR Models: In this section we describe the SR models for the two main actors within the IDS: the *Data Provider* and the *Data Consumer*. Fig. 2 shows the SR model for the data consumer, in which we can see that the main goal of the Data Consumer is to exchange data within the IDS. This has been defined as the goal *Data Exchange through IDS* in the SR model. To achieve this goal the Data Consumer must successfully connect to the IDS infrastructure (task ‘*Connect to IDS*’ in the SR model) being dependent on the *Service Provider*. The main goal of data exchange consists of many sub-goals which in turn can be further decomposed into several tasks. It should be noted that the subgoal *Data Economy* has an *OR* dependency on the main goal of data exchange because according to the IDS architecture a Data Consumer or a Data Provider may or may not fulfil the goal of Data Economy while exchanging data. The other sub-goals are mandatory for fulfilling the main goal for data exchange and thus, they have an *AND* dependency to the main goal. The decomposition is modelled in detail within the actor in the corresponding SR model. The model also provides

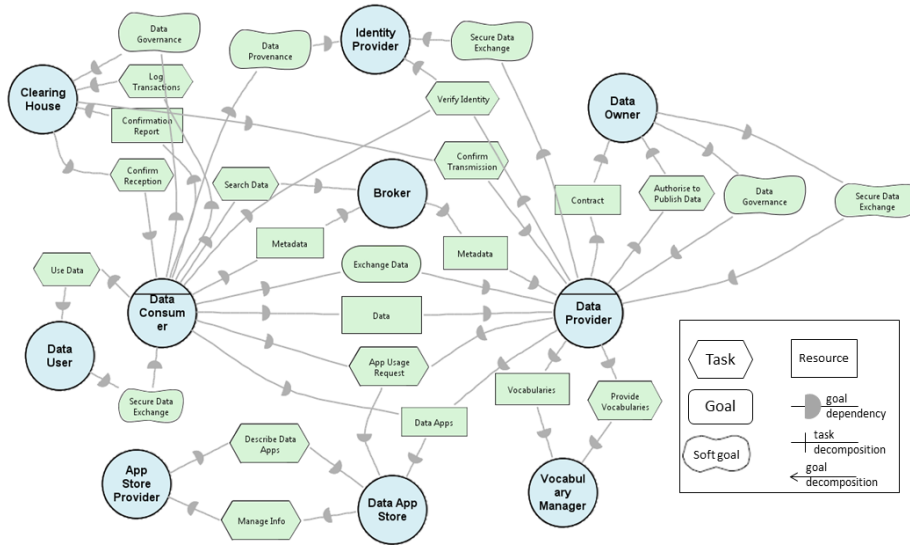


Fig. 1. Strategic Dependency (SD) Model for the IDS

the dependencies on how the internal goals and the tasks of the Data Consumer are related to the external actors of the IDS.

Due to lack of space we could not provide the model for Data Provider, however we would like to mention that the model is the mirror image of the SD model for the Data Consumer.

3 Use Case: Medical Data Space

The Medical Data Space (MedDS) ⁴ is a domain-specific verticalization of the IDS concept and designed to enable easy and secure exchange of data for medical domains. Crucial aspects in the design of the MedDS are data sovereignty, usage control, and data provenance to support transparency and security for data owners. Furthermore, easy implementation, efficient data transfer, and complex data analysis make the MedDS valuable for researchers and for companies, e.g., from pharmaceutical industry. The MedDS infrastructure is implemented according to the principles and along the architecture of the Industrial Data Space [6]

Modelling the P4 Medicine Scenario based on the IDS i* Model: Based on the i* modelling of the IDS in Section 2, we derived an i* model for the scenario. The scenario addresses the concept of P4 medicine for diabetics, where P4 stands for Preventive, Predictive, Personalized, and Participatory [8]. Two systems, a portal for diabetics lifestyle control involving patients and coaches (TeLiPro portal) and an electronic health record (EHR) system, exchange data.

An overview of the model can be found in Fig. 3.

⁴ <http://medicaldataspace.de/>

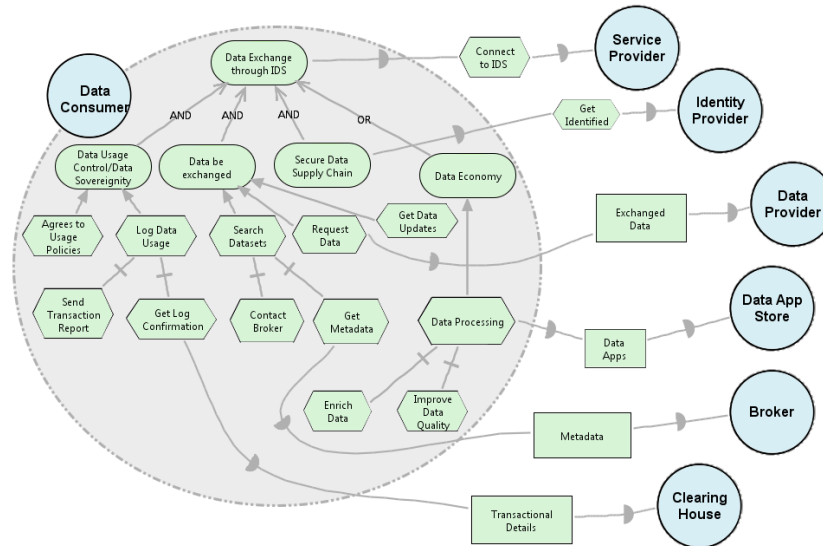


Fig. 2. Strategic Rationale (SR) Model for the Data Consumer

In the following we will describe the actors that we identified and model their interdependencies.

- **Patient / Coach dependencies:** The actor *Patient* is dependent on the actor *Coach*, because she wants to get coaching from him (goal *Coaching*). On the other hand the actor **Coach** depends on the actor *Patient* as a *Data Owner* (role) as the Coach needs information (resource *Patient Info*) from the patient to fulfill the goal *Coaching*.
- **Patient / TeLiPro Portal dependencies:** The *Patient* actor playing the role of a *Data Owner* is dependent on the actor *TeLiPro Portal* trying to achieve the soft goal of *Improving her Health Status*. Also *Data Sovereignty* is a soft goal in this dependency as the patient needs to keep control over her data and the TeLiPro Portal needs to implement means to ensure this. The *TeLiPro Portal* is also dependent on the *Patient* actor, as it needs the patient as its customer to fulfill the goal of *Doing Business*.
- **TeLiPro Portal / Coach dependencies:** The *TeLiPro Portal* actor is dependent on the *Coach* actor, because it enables individual support to patients. The actor *Coach* is also dependent on the *TeLiPro Portal* actor as he can view and add data to fulfill his goal of *Data Management* regarding the coaching of the patient.
- **TeLiPro Portal / EHR Provider dependencies:** Both actors play the role of a *Data Consumer* and the role of a *Data Provider* as they mutually exchange data. The TeLiPro Portal provides as *Vital Parameters* as resource to the EHR Portal, while the EHR Portal provides the resource *Medication Plan* of the patient as data to the TeLiPro Portal and each side consumes the corresponding data. Hence, they are mutually dependent on each other to fulfill their services.

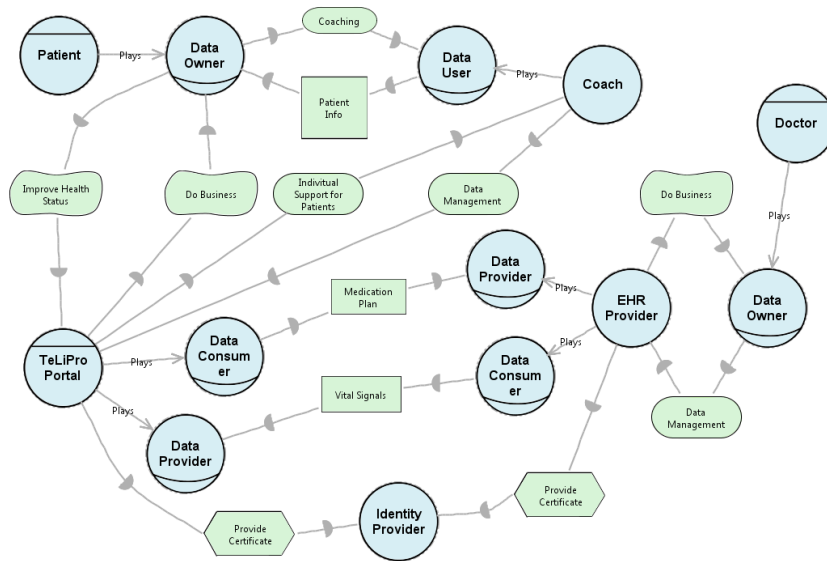


Fig. 3. The i^* Model for the P4 Medicine Scenario

- **Doctor / EHR Provider dependencies:** The *EHR Provider* actor is dependent on the *Doctor* actor as a customer to fulfil its soft goal of *Doing Business*. The actor *Doctor* is also dependent on the actor *EHR Provider* and plays the role of a *Data Owner* who fulfils the goal of managing his patients’ data with the support of the actor *EHR Provider*.
- **TeLiPro Portal Portal & EHR Provider / Identity Provider dependencies:** The actor *TeLiPro Portal* is dependent on the *Identity Provider* to get a certificate from the provider for secure data exchange. This is fulfilled by the task *Provide Certificate* by the *Identity Provider*. The same is true for the *EHR Provider* who also needs a certificate from the *Identity Provider*.

4 Conclusion

The definition of appropriate business models and their continuous evaluation is a challenging task in a data exchange setting. The Industrial Data Space aims at providing an open architecture for a data exchange platform. We have applied the i^* modelling approach to this scenario to capture the intentions and goals of the stakeholders in the IDS. We designed strategy dependency and strategy rationale models at the general IDS level. These models have been validated in our use case implementation for the Medical Data Space, in which we instantiated the generic models with concrete roles and goals of the use case.

The reasoning about goals and dependencies is still at an early stage. We verified that the SR model for the use case contains dependency cycles to avoid the situation that an actor does not have a motivation to contribute to the data

exchange setting. This analysis is important to identify viable business models; the lack of a cycle indicates an incompleteness of the business model. Yet, a deeper analysis of the dependencies and the possibilities to reason about these models has to be performed. For our future work we intend to do an in-depth analysis of the design trade-offs and modelling alternatives as well as validate our model on running use cases from different vertical industries. As we have a strong community in the IDS in which companies and researchers work collaboratively on such issues, we are confident that the models presented in this paper are a first step for a more structured analysis of business models.

Acknowledgements: We thank our project partners Tim Wilking, Christian Mader, and Steffen Lohmann for their contributions on parts of the IDS architecture and the Medical Data Space. This work has been funded by the German Federal Ministry of Education and Research (BMBF) (projects InDaSpace and InDaSpacePlus, grant no. 01IS15054 and 01IS17031).

References

1. Y. An, P. W. Dalrymple, M. Rogers, P. Gerrity, J. Horkoff, E. Yu. Collaborative Social Modeling for Designing a Patient Wellness Tracking System in a Nurse-managed Health Care Center. In *Proc. 4th Intl. Conf. on Design Science Research in Information Systems and Technology (DESRIST)*, pp. 2:1–2:14. New York, NY, USA, 2009.
2. S. Jacobs, R. Holten. Goal driven business modelling: supporting decision making within information systems development. In *Proceedings of the Conference on Organizational Computing Systems, COOCS 1995, Milpitas, California, USA, August 13-16, 1995*, pp. 96–105. ACM, 1995.
3. E. Morales, X. Franch, A. Martínez, H. Estrada. Considering Technology Representation in Service-Oriented Business Models. In *Workshop Proceedings of the 35th Annual IEEE International Computer Software and Applications Conference, COMPSAC Workshops 2011, Munich, Germany, 18-22 July 2011*, pp. 482–487. IEEE Computer Society, 2011.
4. B. Otto, et al. Reference Architecture Model for the Industrial Data Space. *Technical report*, Fraunhofer-Gesellschaft, 2017.
5. E. Paja, A. Maté, C. Woo, J. Mylopoulos. Can Goal Reasoning Techniques Be Used for Strategic Decision-Making? In *Conceptual Modeling*, pp. 530–543. Springer International Publishing, 2016.
6. C. Quix, A. Chakrabarti, S. Kleff, J. Pullmann. Business Process Modelling for a Data Exchange Platform. In *Proc. Forum at the 29th International Conference on Advanced Information Systems Engineering (CAiSE), CEUR Workshop Proceedings*, vol. 1848, pp. 153–160. CEUR-WS.org, Essen, Germany, 2017.
7. R. Samavi, E. S. K. Yu, T. Topaloglou. Strategic reasoning about business models: a conceptual modeling approach. *Inf. Syst. E-Business Management*, **7**(2):171–198, 2009.
8. A. D. Weston, L. Hood. Systems biology, proteomics, and the future of health care: toward predictive, preventative, and personalized medicine. *Journal of proteome research*, **3**(2):179–196, 2004.
9. E. S. Yu. *Social Modeling and i**, pp. 99–121. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.