Position Paper: Approaches to Semantic Interoperability & Semantic Mapping

Michael Jacoby

Fraunhofer IOSB, Karlsruhe, Germany
michael.jacoby@iosb.fraunhofer.de

1 Motivation

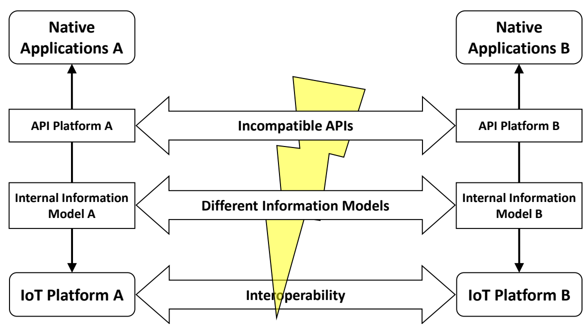

Fig. 1: Interoperability issues in current IoT landscape.

Semantic Interoperability is the key to "data exchange and service creation across large vertical applications" as seen as next step of evolution of the IoT [2]. As depicted in Figure 1, there are two major problems to achieving interoperability between IoT platforms. The first challenge are incompatible APIs and is referred to as syntactic interoperability. This can be addressed by using standardized protocols (of which there are many) or using a meta-level description of the interface, e.g. like Web of Things does. The second issue is that exchanging data is not enough to dynamically connect to other platforms and make use of the data. There is also the need to understand and make sense of the exchanged data. Therefore, this challenge is called semantic interoperability.

For now, standardization bodies rather seem to focus on achieving syntactic interoperability as the number of protocols and communication standards rapidly grows. Therefore, we decided to consider this issue as more or less solved and focus on addressing semantic interoperability, which is the logical next step in achieving a higher level of interoperability.

In this paper, we want to shortly introduce our theoretical analysis of the problem domain of semantic interoperability and present an approach how to address semantic interoperability using Semantic Mapping and SPAQRL Query Re-Writing technologies developed as part of a H2020 research project called symbIoTe (**Symb**iosis of smart objects across **IoT e**nvironments).

2 Approaches to Semantic Interoperability

The problem of semantic interoperability is defined by multiple systems (in that case IoT platforms) wanting to communicate and exchange information but using different information models and terms to describe their knowledge. We analysed the problem domain and found that its solution space can be thought of as a trade-off between two rather radical approaches like depicted in Figure 2. On the one side, there is the

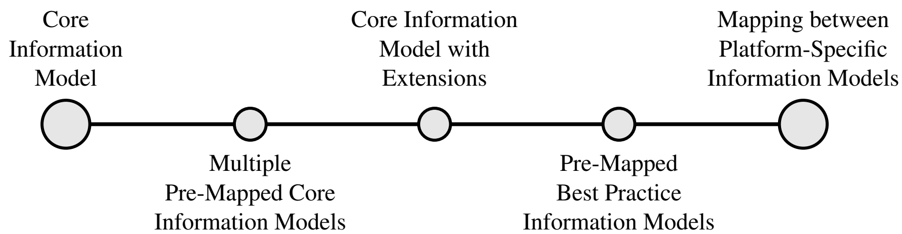

Fig. 2: Solution space for possible approaches to semantic interoperability.

approach we call *Core Information Model*. It describes what is done in almost every system today and what could be referred to by the term *interoperability by standardization*. The idea is that there is one agreed-upon and standardized model/ontology that everybody uses. This way, everybody is able to understand anybody else because they use the same terms. This solution is straightforward, easy to understand and does not induce any computational overhead. Therefore, it seems to be the perfect approach to achieve semantic interoperability. However, there are multiple reasons why this approach is not suitable for or even cannot be used in most real-life scenarios. Such a core information model would need to comprise any term that any platform may ever want to exchange. Assuming this would be possible, we definitely will not be able to foresee everything that will be needed by future, not yet existing platforms. Furthermore, there will be platforms dealing with the same domain but having a different view on it, therefore needing a (slightly) different model for this domain. Knowing this, we are sure that this approach is not a viable solution to a generic semantic interoperability approach, as it will always exclude some platforms. However, we think it should be used whenever possible, e.g. in scenarios with a clearly defined and limited domain.

On the other side, there is the approach we named *Mapping between Platform-Specific Information Models*. The idea behind this approach is that each platform uses its own information model. However, to achieve interoperability between these different information models, (semantic) mappings between them must be defined and somehow executed. The main advantage of this approach is that it works without every single platform agreeing on the same model. For the current IoT landscape where there are a vast number of heterogeneous platforms being deployed, this is the perfect fit. Unfortunately, many downsides need to be addressed and resolved. First, a mapping language is needed to express those mappings in. Further, some logic to make use of the mappings is needed to actually provide an interface that hides the fact that the systems are using different model. For each user it should appear as if all the platforms (there exists a mapping to) are using the same model as he. Furthermore, the mappings have to be defined by someone. This alone is already a quite challenging task as there is often not a single valid solution when models get more complex (actually, it is even harder to say what is a ”valid” solution). And finally yet importantly, this approach also adds a computational overhead to each query or data access as some kind of translation has to be done.

Besides those two quite radical approaches there may exist any level of trade-off between them. We identified three of them, which are explained in short hereafter. For details on this analysis, please refer to [4]. The approach called *Multiple Pre-Mapped Core Information Models* suggests having multiple core models that are statically aligned with each other. This reduces the complexity regarding mapping as there is neither a need for a specialized mapping language nor a generic execution framework but rather mapping logic can be hard-coded. This way, it is quite inflexible to changes and extensions in the future and does not solve most of the problems of the Core Information Model approach.

In the *Core Information Model with Extension* approach, the term core information model is used a bit differently than before. It now describes the concept of a minimal and abstract core vocabulary that is shared across all platforms defining a minimal common set of terms, e.g. the classes Sensor and Observation and the relation hasObservation between them (be aware, that already this assumption is quite disputable even it seems very basic and straightforward). ach platform can then use a custom extension of this core information model to add their own classes and relations. This way, we gain some minimalistic interoperability in terms of the core information model out-of-the-box (i.e. without mapping) but at the same time gain almost full flexibility in platform-specific information models.

The *Pre-Mapped Best Practice Information Models* approach is basically the same as the *Multiple Core Information Models* approach. The only difference is, that here the models are optional to use meaning a platform is not forced to use any of the provided models but can also use a completely different one.

3 symbIoTe: A real-world approach

symbIoTe [5] is an EU project and part of the European Union’s Horizon 2020 research and innovation programme. Its main objective is to provide an interoperability and mediation framework for collaboration and federation of vertical IoT platforms thus enabling creation of cross-domain applications using multiple heterogeneous IoT platforms in a unified way. One part of symbIoTe is a search functionality to find resources based on their meta data across registered platforms. The symbIoTe approach to semantic interoperability is shown in Figure 3 and follows the *Core Information Model with Extensions* approach. Besides the Core Information Model (CIM) it also defines the Meta Information Model (MIM) which is used to store meta information about platforms used for management purpose. To be symbIoTe-compliant, each platform must provide its data based on a Platform-Specific Information Model (PIM) that must be an extension of the CIM. Syntactic interoperability is achieved by implementing a common REST-based interface called *Interworking Interface*.

symbIoTe supports multiple kinds/levels of semantic interoperability. First, all platform are interoperable in terms of the CIM out-of-the-box. Furthermore, if they use the

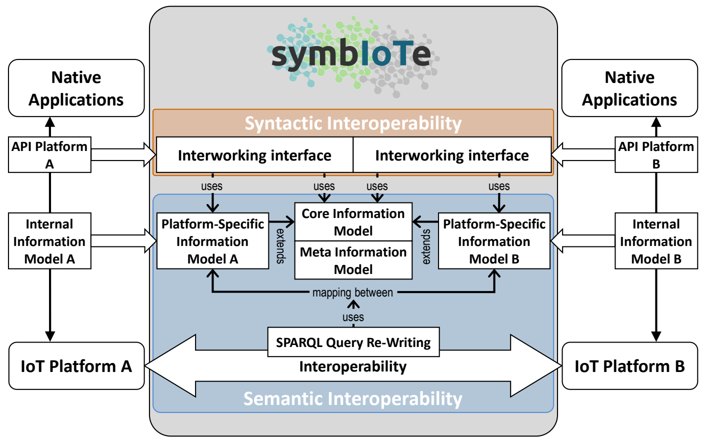

Fig. 3: High-level diagram showing symbIoTe's approaches to interoperability.

same terms to describe things they are also interoperable regarding these terms without any additional effort. If two platforms need a higher level of semantic interoperability, they need to define a mapping between their PIMs. These mappings can be used for two tasks. First, when querying symbIoTe for resources provided by other platforms. These queries are SPARQL queries and formulated against an information model. If the query is formulated against the CIM, then it can be answered across all platforms without further work. If it is formulated against a PIM, the query is executed against the data of all other platform returning results for all platforms that (by chance) use the same terms. Additionally, if there exist any outgoing mappings from the PIM, SPARQL Query Re-Writing is used to translate the query into the target PIM, execute it against any platform using the target PIM and translate the result back to PIM the query is written against. Once this is done for all mappings, the results are merged and returned. From a user perspective, it seems that all platforms he found are using the model he used within the query, hiding that they actually all may use completely different models.

4 Challenges

Realizing semantic interoperability through semantic mapping is quite a challenging task. This section is lists a few of the challenges the author is aware of to indicate where research is needed or problems may arise.

- A mapping language is needed to define mappings. The most promising one for symbIoTe purpose found so far is EDOAL (Expressive and Declarative Ontology Alignment Language) [3, 1].
- Definition of mappings is not easy, especially for non-expert users. Therefore, support for automatic pre-alignment of models (called ontology matching) and a visual mappings editor would be very helpful.
- SPARQL Query Re-Writing is a complex task. There should be a classification of ontology mismatches, an analysis how often they occur in real-world scenarios and how easily they can be resolved.

- How to find other platforms to define a mapping to and how to decide whether it is worth the effort to define a mapping to them? As without mapping no automated understanding of terms of other platform is possible, this needs to be done manually. To support this task, a system could provide any kind of suitable search functionality across the information models of other platforms, e.g. full-text search, phonetic search or NLP-based search taking into account multiple languages.

5 Conclusion

After syntactic interoperability can be treated as solved, achieving semantic interoperability is the next key challenge for the IoT. This paper shows, that a shared and standardized vocabulary between systems is desirable but nor realistic to achieve. Therefore, we need to look at alternative solutions of which semantic mapping seems very promising. However, achieving semantic interoperability through mapping is a challenging task. Nevertheless, the author is convinced that this is right (and probably only) way to establish semantic interoperability in a generic way within the IoT and that open challenges should be addressed by further research.

Acknowledgement

This work is supported by the H2020 symbIoTe project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688156. The authors would like to cordially thank the entire symbIoTe consortium for their valuable comments and discussions.

References

1. Correndo, G., Shadbolt, N.: Translating expressive ontology mappings into rewriting rules to implement query rewriting. In: International Semantic Web Conference, Ontology Matching workshop. CEUR-WS (2011)
2. European Commission: Reaping the full benefits of a digital single market. Comission Staff Working Document (2016)
3. Euzenat, J.: Edoal: Expressive and declarative ontology alignment language (2011)
4. Jacoby, M., Antonić, A., Kreiner, K., Łapacz, R., Pielorz, J.: Semantic interoperability as key to iot platform federation. In: International Workshop on Interoperability and Open-Source Solutions. pp. 3–19. Springer, Cham (2016)
5. Soursos, S., Žarko, I.P., Zwickl, P., Gojmerac, I., Bianchi, G., Carrozzo, G.: Towards the cross-domain interoperability of iot platforms. In: 2016 European Conference on Networks and Communications (2016)