Position Paper for

W3C Workshop on Data Models for Transportation
(https://www.w3.org/auto/events/data-ws-2019/)

Authors: Piero Bonatti, **Martin Kurze** (presenter),
acknowledgements to:Freddy De Meersmann, Ben Wittam Smith

Wordcount (from here): 987 ☺

Policies, Policy Language and Vocabulary to automatically handle privacy issues in future personal and professional transportation

1. Introduction/Motivation:

Transportation is undergoing a fundamental change today: From Mono-modality based on one vehicle owned by the person/company to "multi-modal systems of shared vehicles, orchestrated by online-transportation service providers". This change produces and requires enormous and still growing amounts of data and needs to exchange this data. Much of this data is potentially PII (personally identifiable information) and thus under special legal protection and at the same time of high value for (re)use outside the transportation field.

We propose to integrate means for tracking and controlling PII whenever they appear in the different stages and processes of transportation. The policy model, its components and applicability in GDPR-context and beyond (i.e. in transportation)

From this conceptual starting point, we propose and examine several real-world use cases, relevant to the respective industry and company with regard to applications in transportation. While the present position paper focusses on the needs and tools for end-to-end solutions, the respective industry partners describe related positions by other members of the SPECIAL project in separate papers.

We conclude with a brief summary and an outlook as well as an invitation for collaboration, on project or company level.

2. The SPECIAL approach: policy language, vocabulary and policy engine

The H2020 project SPECIAL[1] has already developed a rich framework for **consent management and automated compliance checking** that should be of interest for most of the participants to this workshop. The main components of the framework are:

- A **policy model** for data usage, specifically designed to address the requirements of the GDPR. This model covers both data subjects' **consent** and data controllers' data usage policies (**business policies**).
- **Vocabularies (ontologies)** for the main usage policy elements, such as purposes, personal data categories, recipients, storage duration, etc, as well as GDPR's legal bases and varieties of consent. A first version of the vocabularies is being released by the W3C Data Privacy Vocabularies and Controls Community Group, promoted by SPECIAL.

1 Funded by the European Union's Horizon 2020 research and innovation programme under grant agreement N. 731601.

- An encoding of the policy model and the vocabularies in the **standard OWL2**. The encoding provides the model with a **formal semantics** - sorely needed for interoperability and for guaranteeing the coherence of compliance checking, explanations, and policy validation.
- A **scalable engine** that can execute a compliance check every 0.5 ms without resorting to parallelism. The engine can process all expressions in a new profile of OWL2 called PL (Policy logic) that is general enough to accomodate a number of possible extensions to the current model, and follow the evolution of the vocabularies. SPECIAL developed also a **parallel big data architecture** for complianc checking for further speed-ups.
- A **transparency infrastructure** where (i) data controllers can log personal data processing operations and monitor their internal processes, (ii) data subjects can enquire how their personal data are processed, (iii) data protection officers can audit personal data processing activities.
- An innovative **dynamic consent** request method, aimed at reducing the privacy-related burden on users.

In other words, this framework constitutes a significant step towards a **standard representation of personal data and consent** and their processing, that already includes some personal data whose processing naturally arises in the **automotive** domain (e.g. locations).

3. Transportation related use cases and experiences from a Telecommunications perspective

Being a service provider for telecommunication and a data center operator at the same time, Deutsche Telekom can play (at least) two roles in the ecosystem of data models and related issues in tomorrow's world transportation:

1. As a trusted 3rd party, DT could ensure that (PII) data is transferred between partners in a secure and privacy aware way. Using the SPECIAL approach, DT could also apply user-defined policies for data sharing. E.g., DT could strip PII datasets down to a level compliant with the user's desire (formulated as policy). This of course requires advanced and interoperable vocabularies and ontologies.

2. As a service provider with a huge customer base and large data sets (much of it PII), telecommunications companies have the need/opportunity to monetize this data as other ("OTT") companies do anyway. These large data sets will become even larger and more valuable with upcoming technologies such as "5G".

5G "precise positioning" and other techniques will enable new applications and services in the transportation sector. Most of them affect and rely on PII. To keep the customer/user/citizen in control of his data (or the data that the vehicle he currently uses generates), several steps of data control need to be considered:

- Data collection (sensors in vehicles, wearables and the environment)
- Data transfer (via networks, e.g. 5G) to data center(s)
- Data analysis (by one or more entities with a limited set of purposes)
- Data sharing (with or without user consent)
- Data expiration
- Data accessibility and control for users (data subjects)

With the policy model of EU-project SPECIAL, this can be mapped to technology and business. Nevertheless, a number of questions already arises and we hope to provoke and collect not just answers but also more questions during the workshop. The key issue remains interoperability not just between players in the (PII) data space but also in the (upcoming new) transportation ecosystem. We should find compatible models and schemes and specify common standards.

4. Conclusions and invitation to collaborate

Transportation and telecommunication are just two industries dealing with data, some of it "personal/private" in nature. A few questions and open issues are intended for discussion during the workshop. The following list is just a starting point for a hopefully lively discussion and exchange of experience:

- How to get "user consent" to exchange PII?
- Do we need "one huge universal" ontology (for transportation and privacy) or several smaller ones?
- Would a "policy model" be applicable for transportation too?
- Transparency for the data subject and privacy protection might contradict the "open" use of PII (such as location information for advances traffic management)