CCS/TMS Data Model - Scope and Approach for Collaboration and Specification

Contents

1 Purpose

This document describes the scope, approach, and general methodology of TCCS SD1. Since the overall data model is a central topic of this transversal subdomain, an emphasis is put on the mode of work and collaboration with the System Pillar Tasks/Domains and the Flagship Areas of the Innovation Pillar.

2 References

TCCS SD1 - Data Model_01_Introduction
TCCS SD1 - Data Model_02_Schema

3 Summary - Transversal CCS Sub Domain 1 (SD1)

SPT2TS-123978 - The Transversal CCS Sub Domain 1 (SD1) of the ERJU System Pillar started its activities in December 2022. As part of the ERJU System Pillar, standardisation focuses on traffic/train CS, traffic management, or use cases such as diagnostics. Subsystems and interfaces are standardised within the System Pillar to implement functionality such as Train Protection or Automated Operation, representing different use cases (white in Figure 2) for Domain Data (blue in Figure 2) and its underlying Engineering Data (green in Figure 2).**[**Content to be approved**]**

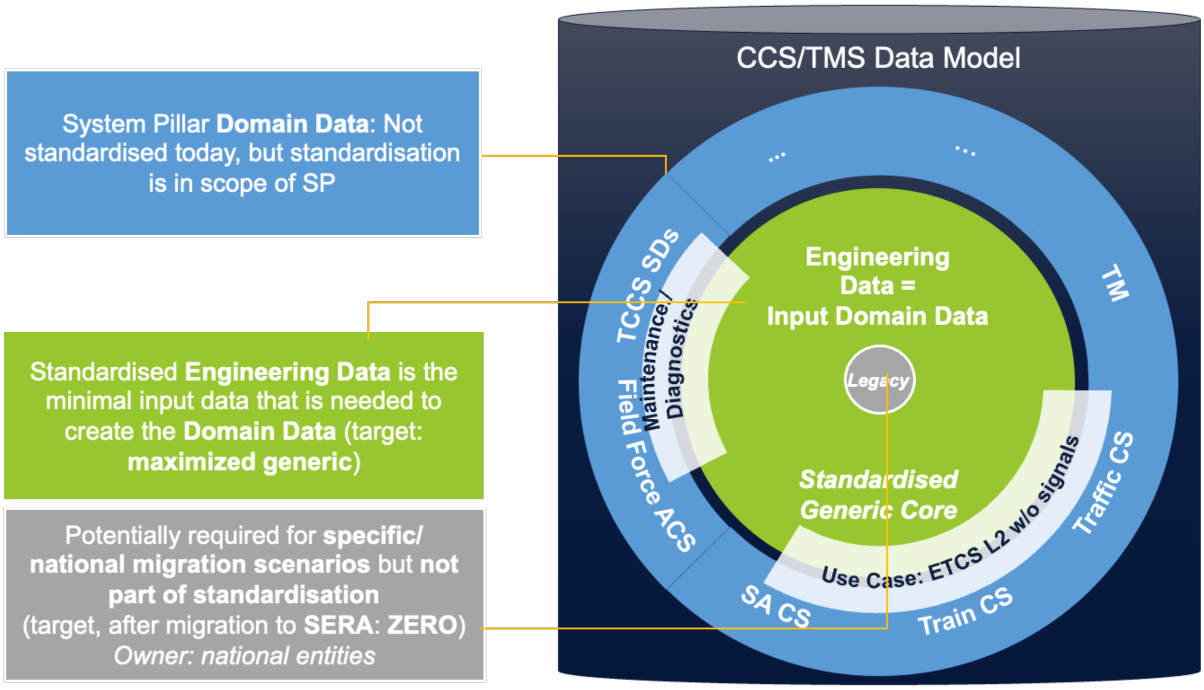

Figure 2: Scope of TCCS SD1 - harmonised Data Model for End-to-end process from Engineering Data to Domain Data

SPT2TS-123981 - TCCS SD1 will standardise the data exchange objects of functional interfaces to provide the resulting data from engineering activities (green in Figure 2) in a generic and validated form. This data can be automatically compiled into the so-called Domain Data (blue in Figure 2) for publishing to all the consuming systems of the System Pillar (SP) architecture. The data shall be validated to fulfil a harmonised set of engineering/validation rules imposed by and coordinated with relevant domains. The data model is not derived from existing legacy processes and artefacts but instead is driven by functional system data needs. The model shall enable digital solutions best adapted to provide data of the required quality in the most efficient way. This functional-driven approach maximises the potential for a generic engineering core with harmonised rules, data formats, processes, and tools.**[**Content to be approved**]**

SPT2TS-123982 - SD1 doesn't invent new models but picks and recombines fragments from existing standards or data models to fit the functional and non-functional needs that are elicited within subdomain SD1 and by the other domains representing the data users. These requirements and analyses are sources or rationales for design decisions. The result of SD1 is a merged data model based on technical specifications, supporting the direct usage by the systems and the whole life cycle of data within the System Pillar architecture (incl. safe data publishing, activation or deactivation of data versions). Therefore, the data model must comply with restrictions of "on-the-wire" data transmission (not only static file-based loading) and support all relevant schemata such as binary/protobuf, json, xml/xsd,..). With respect to expected safety-related functions using the data, the development of processes and interfaces must be according to EN 5012x.**[**Content to be approved**]**

SPT2TS-123983 - The resulting SPT2TS-2040 - CCS/TMS Data Model is controlled by the System Pillar TCCS SD1 to enable governance and stability. The latter shall support backward compatibility and address the aspect of securing investments into toolchain developments or expensive modifications to processes. The migration to the "Single European Railway Area" (SERA) shall be supported by specific extensions to the interface model.**[**Content to be approved**]**

SPT2TS-123980 - Regarding the link between SD1 and the existing ontology approach from ERA vocabulary, it is referred to the detailed descriptions of this document (6.6- Linking to other Data Models with Ontology).**[**Content to be approved**]**

4 Introduction

4.1 High level-process and terms

Railway Systems require engineering and asset configuration data to implement the specific application within a defined infrastructure area.

The following high-level data terms are introduced for further explanation (both are covered by the same SPT2TS-2040 - CCS/TMS Data Model):

Engineering Data

The Engineering Data as part of the SPT2TS-2040 - CCS/TMS Data Model contains all the base data (i.e., track topology, track geometry, track asset configuration) for compiling the next version(s) of use case-specific Domain Data. The standardized Engineering Data comprehensively covers the data needs for the SERA in the scope of the System Pillar (i.e. radio-based ETCS only). Specific data needed for migration is out of the scope of SPT2TS-2040 - CCS/TMS Data Model.

Domain Data

Domain Data as part of the SPT2TS-2040 - CCS/TMS Data Model is a use-case-specific representation of information following the specific needs of affected systems and their functionalities implementing the use case. Domain data are a tailored and potentially transformed subset of Engineering Data.

SPT2TS-2032 - The overall process can be summarised on a very high level by the following activities:

- Collecting inventory or newly acquired input data for the engineering phase
- Preparation and validation of the next release of Engineering Data as the single source of truth for the next steps.
- Transformation (Compilation) of this Engineering Data into the data structures as required by the consuming systems. This resulting data is referred to as Domain Data.
- Publishing this Domain Data to the consumers, I.e. systems that need the information, following a publishing principle (e.g. point-to-point, publish/subscribe,..) and a standardised interface specification.
  Depending on the kind of information, consuming system, and used publishing principle, the loading and activation of this Domain Data within the systems are required (potentially incl. a response to the source). This is performed to ensure that the data is always synchronised and ready to be applied/used for specific applications.
- Application of the Domain Data by the subsystem. This includes internal system functionality but also interfacing messages to other systems.

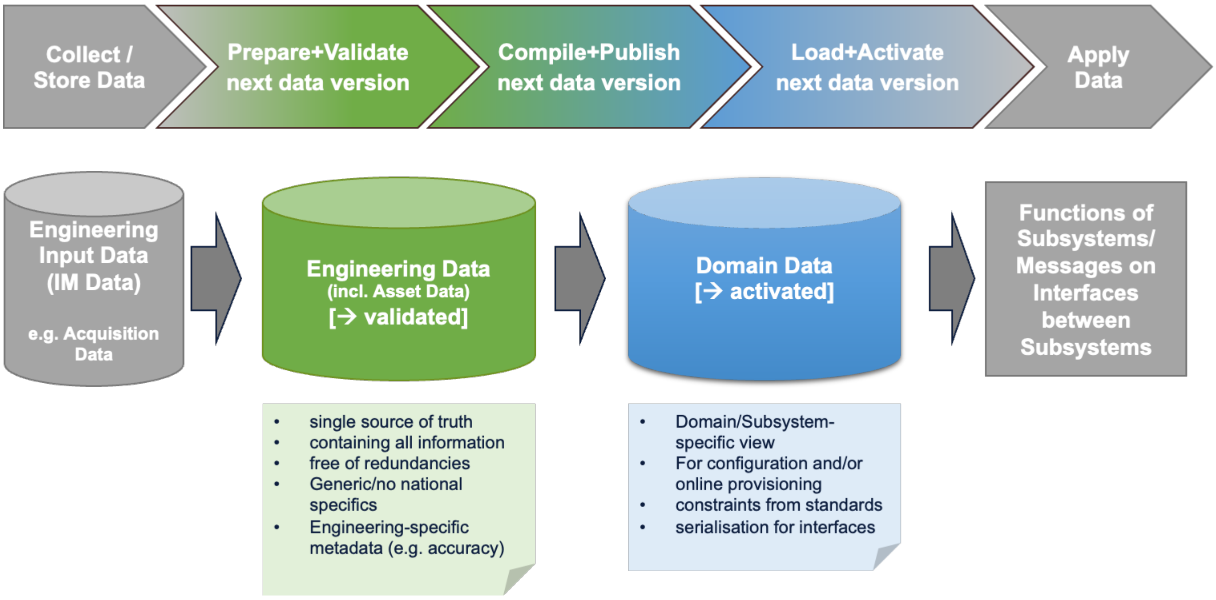

Figure 1 Main activities of high-level data process using the terms "Engineering Data" and "Domain Data"

[Content to be approved**]**

4.2 Motivation

SPT2TS-123984 - Today, Engineering Data-related processes are too often paper/drawing-based, manual activities, inefficient, and not state-of-the-art – compared to the available technical possibilities. Subsystem-specific, propriety formats characterise the Domain Data. Consequently, standardised digital toolchains in the field of CCS engineering are not yet established in the railway domain.
As with national BIM plans by infrastructure managers, a proper digitisation strategy also relies on the standardisation of interfaces within this process to support coordination, data exchange, and close collaboration within the project. The standardisation is also a consequence of the SERA objectives and call for a harmonised solution.**[**Content to be approved**]**

SPT2TS-123985 - The data structures developed in national contexts are a first step towards this digitised process flow. The international standardisation of data formats, rules, and processes improves the business case and ROI due to scale. Higher development costs related to safety, such as CCS, increase the need for a standardised environment with the semi-automated, digitised engineering process, i.e. planning, validation or transformation functions. Today, several initiatives provide interface data formats, such as EULYNX Data Prep, railML, RSM, RCA, IFC-rail, Linx4Rail, and X2R4-PSM. These partially overlapping standards complicate the decision process of IMs or other parties to invest and build toolchains following long-term roadmaps. Also, the coexistence of several data standards in parallel for the same use case is not acceptable for safety, functional or economic reasons if a new, standardized architecture, as intended by the System Pillar, is developed. ERJU System Pillar shall improve the situation by harmonising the input information required from engineering or infrastructure data inputs to operate the System Pillar systems within their area of operation. As it will be shown in the following chapters, the decisions will be driven by the functional needs of the consumers, though the technical and economic feasibility is equally essential for the System Pillar design process. **[**Content to be approved**]**

5 Assumptions Target Architecture

Regarding the overall System Pillar architecture, the following assumptions relevant to SD1 are made:

SPT2TS-2034 - Increased configurability

- The systems (as designed by domains like Traffic CS) will offer generic functionality with parameterisation, so Software and configuration are separated.
- The systems will offer flexible data loading and configuration regarding Topology and Asset Data.
- The Domain Data can be consumed from a single source of truth by clearly defining provisioning interfaces (incl data quality constraints) and responsibilities.
- At least for migration, packaging of SW and configuration is needed before the data is validated.

[Content to be approved**]**

SPT2TS-2035 - Increased data needs

- In the future, the consumers of topology, geometry, and further information will increase, e.g. localisation, IPM/Perception systems, management systems, information systems, and warning systems.

[Content to be approved**]**

SPT2TS-2036 - Support of CCS/TMS data model

- All system interfaces that transport Domain Data will support a common data structure and semantics as defined by the SPT2TS-2040 - CCS/TMS Data Model, incl. interface-specific quality constraints that must be met. However, use case-specific restrictions on the interfaces are expected, e.g. from file loading up to „on-the-wire“ transmission with binary protocols.

[Content to be approved**]**

SPT2TS-2033 - Generic engineering and data validation rules

- Standardised architecture, standardised engineering rules, and harmonised operational rules go together
- Generic engineering rules enable standardised implementation and the economical automation of data generation, validation, and transformation.

[Content to be approved**]**

6 Scope of TCCS SD1

SPT2TS-2037 - A harmonised digital CCS/TMS system requires a shared data language applied at all relevant interfaces with similar exchange items. Furthermore, it requires comprehensive engineering data for planning and building assets according to the System Pillar architecture (i.e. radio-based ETCS only). With the SPT2TS-2040 - CCS/TMS Data Model
the System Pillar provides a data structure for standardised engineering and to align a data structure for standardised interface specifications within CCS including CCS-TMS. This data structure shall be suitable across all relevant use cases of the System Pillar such as engineering and communication for traffic control or automated train operation.

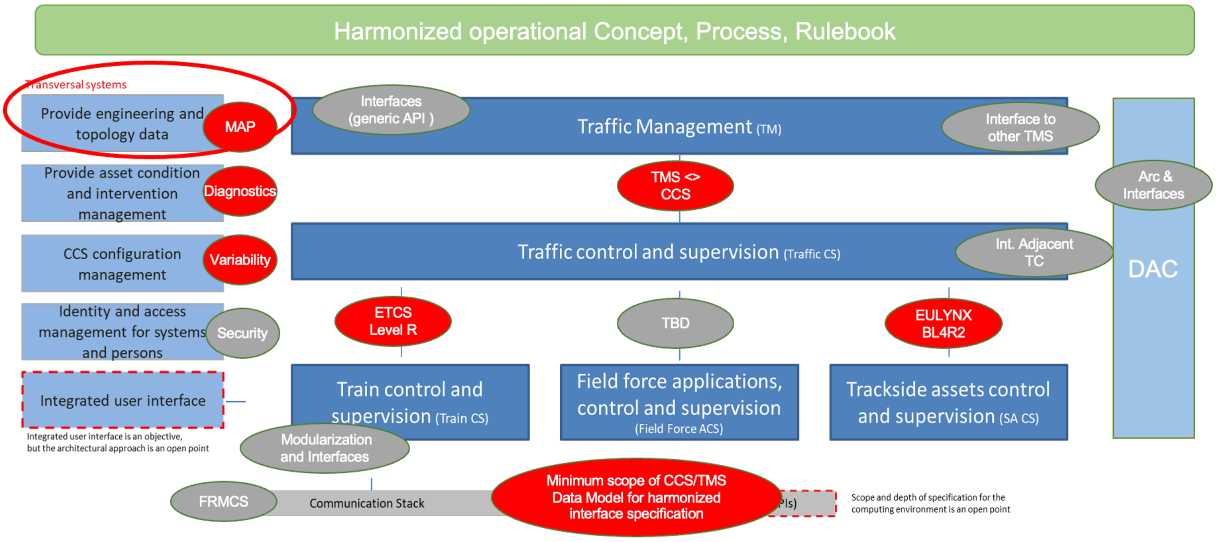

Figure 2 Interfaces between domains and minimum scope of CCS/TMS data model

The Data Model and interface specification must be comprehensive for SERA and therefore sufficiently detailed (not on a high conceptual level), at least on the Functional Interface Specifications (FIS) level within the System Pillar architecture. The development must be according to SEMP. The modelling approach is selected to support these goals.

While the TCCS domain is located in Task 2 of the System Pillar, it offers the inclusion of other tasks with the same and consistently applied CCS/TMS Data Model, as it is already applied for the interface between Task 2 and Task 3 TM.

Besides the static aspects of a data model, the dynamic aspects related to the life cycle of the data are considered by TCCS. This includes the specification of an end-to-end process for data preparation and provisioning (configuration), starting from the required output from the engineering process (SPT2TS-2030 - Engineering Data) up to the compiling of use case-related data (SPT2TS-2031 - Domain Data) for distribution and configuration (e.g. functional use case of Maintenance/Diagnostics or ATO).

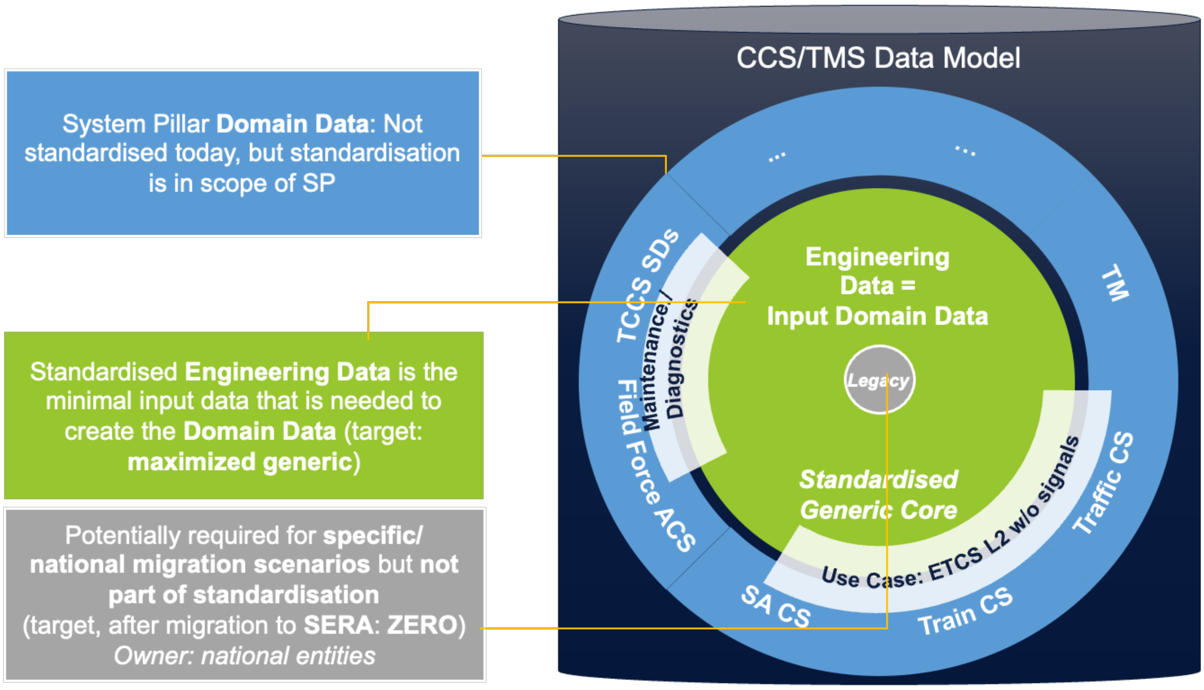

Figure 3 : Scope of harmonised CCS/TMS Data Model for the end-to-end process from preparation/distribution to application of data

[Content to be approved**]**

6.1 Data Model

CCS/TMS Data Model

The CCS/TMS Data Model defines the harmonised language to generate and transport the Domain Data at System Pillar interfaces. The Transversal CCS Subdomain 1 (SD1) is responsible for the specification of the CCS/TMS Data Model in collaboration with

- the System Pillar domains which apply the defined data structures in interface specifications
- the Innovation Pillar which proves the applicability of the data model by demonstrators.

SPT2TS-123560 - SD1 plans to focus on the following goals regarding data model development:**[**Content to be approved**]**

- Analyse the functional data needs: which data the system requires to realize a particular use case (e.g. ATO).
- Support all use cases that rely on Domain Data along the life cycle, e.g. ATO, Train Protection, Localization, Diagnostic Systems, Security, etc.
- Establish clear responsibility and a single source of truth for required domain data
- Ensure feasibility of data model requirements regarding economic and technical aspects
- Connect to configuration management so that HW, SW, and domain data fit together
- Focus on the users: Design a common data model that is „close to implementation“, ready for usage without complex transformation
- Comply with restrictions of „on-the-wire“ operational use cases (dynamic data, e.g. serialisation for ATO Segment Profile)

- consider security aspects as part of the Domain Data and the overall architecture design.
- Derive required information to generate Domain Data. This minimum input model describes the Engineering Data, which is actually a part of the overall SPT2TS-2040 - CCS/TMS Data Model.

6.2 Data Life Cycle

SPT2TS-2041 - Besides the data model, SD1 covers the whole life cycle of data, i.e. the process of providing new data versions and deactivating obsolete data must be considered. The more dynamic aspects along the life cycle of data comprise the following topics:

- Validation: Ensure Data Quality by a clear definition of quality parameters and generic rules to prove in an automated way
- Compilation: Automatically transfer the information from the Engineering Data into the Domain Data
- Provide: Accelerate data provisioning process, incl. loading & synchronous activation of new data versions. Support flexible data provisioning methods fitting to the applications, depending on the use case and the characteristic of the Domain Data (e.g. update frequency/triggers, safety-criticality of information, time-criticality of information,...)
- Version Management: include a central versioning approach for all use cases and support simple compatibility checks, i.e. during runtime.
- Change Management: ensure development process that maintains backward compatibility in case of data model updates
- ID Management: define responsible entities for creating ids and the lifecycle of ids and create a definition for unique identification that can be used for operational, safety and security use cases

[Content to be approved**]**

6.3 Examples

SPT2TS-2042 - Based on the previous scope definition, the following *data flows are within the scope of TCCS SD1*:

- Compile Domain Data (for specific use case, e.g. see next bullet points) from Engineering Data
- Provide Domain Data (topology incl. elements) to the protection system
- Provide Domain Data to configure and connect to trackside assets
- Provide Domain Data for onboard configuration.
- Provide Segment Profiles as Domain Data for Automatic Train Operation
- Provide current topology as Domain Data to traffic management
- Provide map data as Domain Data for localisation or other onboard functionalities

[Content to be approved**]**

SPT2TS-2038 - In contrast, the following data-related tasks *are not part of TCCS SD1*:

- Develop semantical data models (e.g. provided by ontologies)
  Note: Semantic dictionaries or linkings as provided by ontologies are valuable to reuse data definitions and connect to different data models but are not sufficient for a specification on FIS level as intended here.
- Detailed data models to document national- and legacy-driven engineering processes properly.
- Internal data structures, e.g. to store topology data within systems such as automated train operation, protection systems or localisation

- Topology-related functionality of the systems, e.g. write the state of an asset (e.g. switch left, right), create movement authority based on topology.
- Remaining supplier-specific configuration data to parameterise generic systems

[Content to be approved**]**

6.4 Constraints and applied Modelling Approach

6.4.1 Constraints

SPT2TS-123929 - As long as System Pillar Domains are in early development phases such as Operational or System Analysis, the possibilities for a collaborative interface design and approval based on the CCS/TMS Data Model are very limited. At the same time the parallel Innovation Pillar is expecting the CCS/TMS Data Model as a performant basis for their specifications and demonstrators.
TCCS defined a working mode compatible with the System Pillar's SEMP which supports the evolution of the other domains without missing the contributions to the Innovation Pillar as an important resource.**[**Content to be approved**]**

6.4.2 Top-Down and Bottom-Up

SPT2TS-123930 - Hence, TCCS has already provided the first releases to the Innovation Pillar to test the CCS/TMS Data Model in actual applications and ensure sufficient maturity before proceeding with standardisation. The feedback from the Innovation Pillar is continuously processed to improve the CCS/TMS Data Model, under permanent consideration of a standardised data structure for all use cases of the System Pillar. This process part is called "bottom-up" - since it starts with the requirements and feedback from the implementation and application of the CCS/TMS Data Model. In order to achieve standardisation within System Pillar, the "bottom-up" approach is combined with a parallel "top-down" approach of the System Pillar. During the starting phase, this process works with requirements derived from high-level objectives and assumptions regarding the needs of other domains. These needs are based on already well-defined use cases that remain stable (such as radio-based ETCS only, Trackside Assets, ATO GoA 2), prework from input documents coming from LinX4Rail, EULYNX, OCORA, RCA and others, and ongoing discussions with the System Pillar domains and Flagship Areas projects of the Innovation Pillar.
To sum it up, TCCS chose the following combination of principles for effective development and collaboration:

- **Top-down** development based on objectives, assumptions up to approved requirement specifications with other domains following the rules of System Pillar SEMP (System Engineering Management Plan)
- **Bottom-up** development based on practical experiences i.e. of demonstrators from Innovation Pillar, which is supported by the provided CCS/TMS Data Model formats for implementation (e.g. json, xml)

So, while the CCS/TMS Data Model might face pending domain approvals in the early stages, the **profound basis of the System Pillar foundation (LinX4Rail, RCA, EULYNX, OCORA) as well as the practical collaboration with the Innovation Pillar and other early adopters (ERA, PRIME, ...) allow quick and mature development of model increments** (see also development history in SPT2TS-122465 - CCS/TMS Data Model Revisions)**[**Content to be approved**]**

6.4.3 Functional-driven data model design

SPT2TS-123931 - Independently of the applied principle (top-down, bottom-up), the resulting CCS/TMS Data Model is always derived from functionality, which means each data structure shall be linked to a functional use case (e.g. gradient profile for ETCS supervision, infrastructure objects for radio based ETCS only asset engineering, ATO, localisation, ,..) and its constraints. Consequently, **each model part has a trace to the system functions or interfaces that use the defined structures**, enabling all advantages of model-based system design

and ensuring the completeness and consistency of data structures. The impact of possible model adaptions can always be evaluated regarding affected functions, systems, and interfaces. Non-functional aspects can be inherited from the function to the linked data structures (e.g. accuracy of location data).
The principle of functional-driven is applied for all kinds of interfaces and along the complete data life cycle. Hence, the engineering data as input for planning, building and configuration will also be function-driven instead of process-driven. It is explicitly prevented to build up the data model based on (digital versions of) existing engineering/ preparation processes based on particular national IM or supplier-specific procedures or guidelines. The resulting CCS/TMS Data Model shall be minimal, while the structure for engineering and configuration should already be comprehensive for the actual communication during operation (e.g. the model contains lightweight, high-performant sampled geometry descriptions (e.g. ETCS, ATO,..), and very accurate, lossless track alignment model part is provided for first operational use cases requiring it).

The principle can be applied and documented with the Polarion tool (e.g. see example SPT2TS-122498 - Timing Point (Demo Structure) as well as Capella or similar MBSE tools.**[Content to be approved]**

6.5 Expected Application of CCS/TMS Data Model

SPT2TS-123933 - Based on the described scope and approach the CCS/TMS Data Model is expected to be applied to the following activities:

- System Pillar:
  - data structure for communication interface specifications, i.e. within CCS but also including CCS-TMS interface
  - data structure for maintenance interface specifications, i.e. for providing new versions of SPT2TS-2031 - Domain Data
    - data structure for describing the end-to-end process from aggregating and validation of SPT2TS-2030 - Engineering Data up to compiling and distribution of new versions
  - data structure for diagnostic interface specifications
- Innovation Pillar:
  - data structure for Domain Data Specifications and Implementation for different use cases and demonstrators, e.g. ATO, Moving Block, Localisation, e.g. FA2 WP13, WP27, WP45, WP21.
  - data structure for Engineering Data and process/ tool specifications and developments, e.g. FA1 WP 27
  - validation tooling for enabling the uptake of the CCS/TMS Data Model in all related projects of the Innovation Pillar FAs

No licensing model (i.e. for external usage) has been decided yet (e.g. European Union Public Licence EUPL, Version 1.2).**[Content to be approved]**

6.6 Linking to other Data Models with Ontology

SPT2TS-2052 - The CCS/TMS Data Model is a specific Data Model which is directly used for data and interface specification. Specific data models describe far fewer data objects (ignoring all unneeded information) but are precise and optimized for specific functions. Central-specific data models try to find "the common precise data objects" of all functional input/output to avoid many conversions and redundant data creation between systems. They try to reduce the scope to the real needs and purely focus on the function.

While the CCS/TMS data model will be the base for platform-specific schemas to comply with all the restrictions (transmission/interface "on-the-wire"/"airgap", existing regulations, safety standards, System Pillar system needs,...), an integration into the ERA Vocabulary is pursued for the purpose of legislation and semantical linking to further data models (e.g. BIM/IFC). Besides reusing semantic definitions and obtaining consistent use of terms within the railway domain, this approach also offers the possibility to benefit from existing links of other data models which already are or will be connected with ERA vocabulary (e.g. railML, BIM IFC).

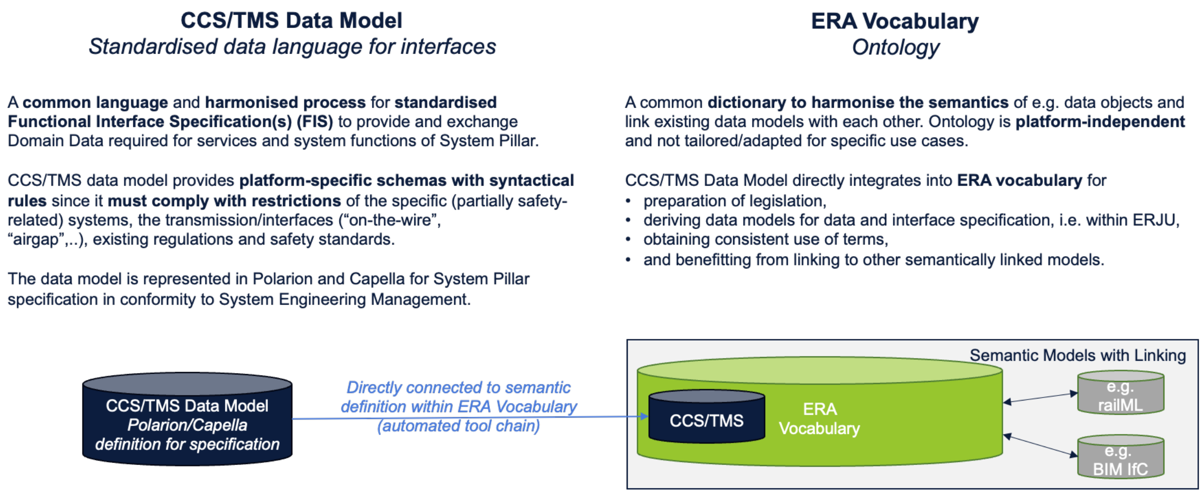

Figure 4 CCS/TMS Data Model bidirectionally linked to ERA vocabulary

In the first step of the integration into ERA vocabulary, an initial linking has been performed by identifying the matching classes and properties/attributes of ERA vocabulary and the CCS/TMS data model and applying some changes to the CCS/TMS data model for optimal integration. The result is automatically exported into ontology language (ttl files), which are used for technical integration of CCS/TMS data model classes into ERA vocabulary. After the integration process has been finished, all further developments in the System Pillar regarding data models are reflected in ERA vocabulary. Official releases of the data model will always be derived from ERA vocabulary by an automatized process, as shown by the following process overview:

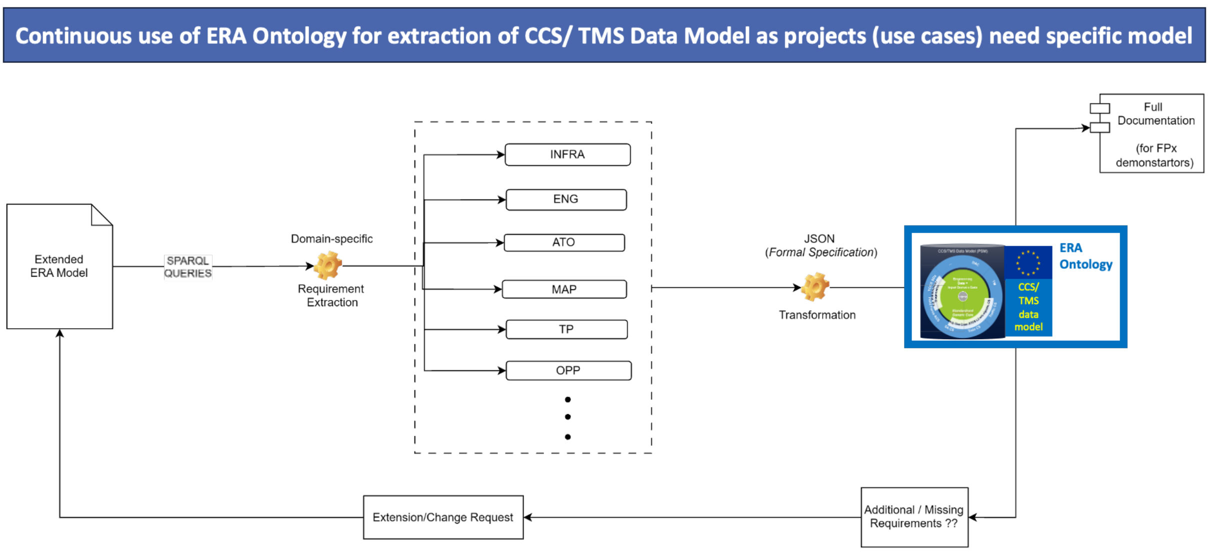

Figure 5 CCS/TMS Data Model derived from ERA Vocabulary to serve interface specification and demonstration needs

To support the specification process, the representation of the CCS/TMS data model in the toolchain of System Pillar (Polarion and Capella) is maintained by an automated toolchain, which keeps the connection to ERA vocabulary.**[**Content to be approved**]**

6.6.1 Example BIM IFC

6.6.1.1 SD1 Linking to external BIM model

SPT2TS-122316 - One important example regarding linking to external models is the connection to BIM (IFC).

Many Ifc objects relate to railway business and are likely to *partially* overlap objects that SD1 will define. This is business as usual because Ifc focus is building objects whilst SD1 is about functional object information with CCS focus (incl. CCS/TMS interface). This said SD1 and Ifc must avoid evolving in silos. The aim is to minimise redundant overlap and avoid potentially conflicting information.
Best practice is to expose object (type) information, such that users can freely navigate and retrieve needed data, e.g. to use alignment information structured in ifc as input to Engineering or Domain Data as defined by SPT2TS-2040 - CCS/TMS Data Model. The premise for finding and accessing information defined in different models is that objects and their definitions are identified, findable and linked. This is typically done by exposing the model in a way that users can navigate to the definitions. When given an object the user can navigate to the definition. This makes accidental model overlaps unlikely and optimises the reusability of data between SD1 and Ifc.
Exemplary linkings between SD1 and ifc are:

SD1	**ifc**
SPT2TS-64107 - Track Edge Geometry e.g. Horizontal Segment	Ifc alignment segment with a set of IfcAlignmentParameters
SPT2TS-49087 - Applicable direction	Ifc Signal with railway signal types
SPT2TS-63866 - Track Vacancy Proving Section (TVPS)	Axle counting equipment
...	

The actual linking of structures from the SPT2TS-2040 - CCS/TMS Data Model to the IFC data model will be applied as a joined initiative with FP1 WP30 (IFC is already included in the CDM of FP1). **[**Content to be approved**]**

6.6.1.2 SD1 Linking to external BIM data

SPT2TS-122318 - In addition to semantic linking of data models (SPT2TS-122316), the linking between SD1 and ifc after instantiation of objects during a real project shall be defined as well.

Ifc objects all inherit from IfcRoot and by this virtue have a IfcGloballyUniqueId. The easiest solution is to equate the SD1 identity to this IfcGloballyUniqueId. However, this could create dependence during track design (alignment) and subsequent phases of the lifecycle. Uncoupling the processes and model can be achieved by

defining references with keys.
The linking on the object instance level requires further elaboration and definition during the semantic linking process.
[Content to be approved**]**

7 Consequences for the System Pillar Domain Design

SPT2TS-2039 - The SPT2TS-2040 - CCS/TMS Data Model and life cycle aspects will impact interfaces, which must be specified by the System Pillar architecture and domains. To support domain-driven extensions and maintain control of the data model, the following working mode shall be applied from the SD1 perspective (see also 8.1- Domain Data Specifications).

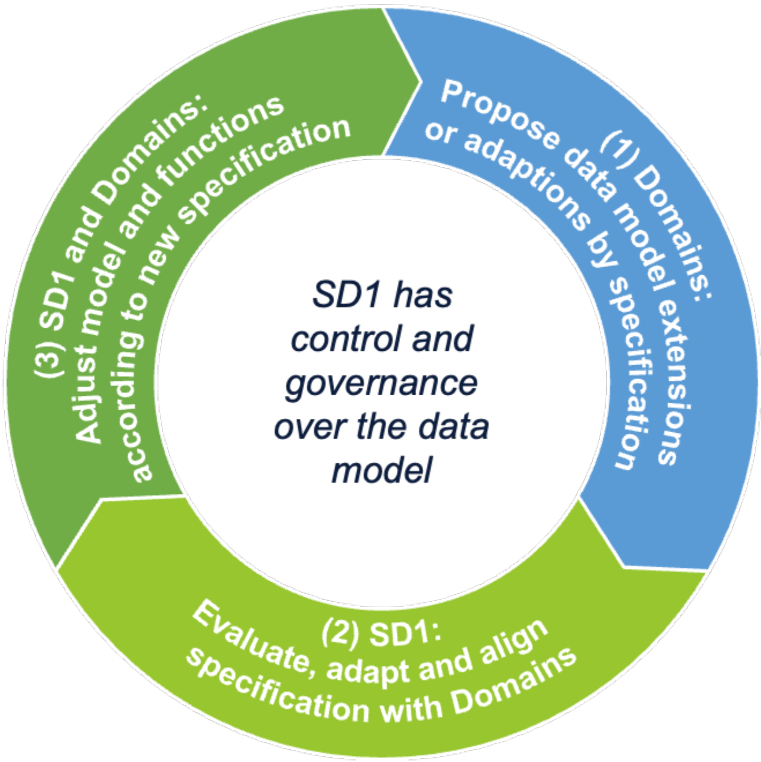

Figure 6 : Cycle of domain-driven data model extensions and collaboration with SD1

1. The domains specify the needs regarding data in a unified way. This includes structural aspects such as objects, attributes, relations, and quality criteria (e.g. accuracy, expected update frequency, time-criticality, safety-criticality)
2. The domains discuss the needs with SD1. During this process, the data model extensions are worked out and evaluated by SD1. Besides data structure, the alignment process could also include decisions regarding conflicting targets/quality criteria (e.g. high accuracy requirements vs. always updated data)
3. After that, SD1 and the domains adjust their specifications according to the result of the alignment process.

The above-mentioned cycle starts again for each evolution step (ensuring backward compatibility, i.e. for TSI increments).
A similar procedure is proposed to collaborate with Flagship Areas of the Innovation Pillar. **[**Content to be approved**]**

8 Approach for Collaboration and Specification

8.1 Domain Data Specifications

SPT2TS-2048 - Issues regarding engineering, commissioning/approval and operation arise when the Domain Data being brought into the systems isn't in a format and quality usable by data consumers. A data specification that enforces specific formats, constraints and semantic meanings can mitigate such instances. This specification requires input and buy-in from all stakeholders involved with data consumption. During the specification phase, the "Domain Data Contracts" aim to receive and process the needs for data model extension in a structured way. For the implementation and operation, the specification shall ensure reliable data that all consumers can trust for their realized use cases.**[**Content to be approved**]**

8.1.1 Principles of Domain Data Specifications

SPT2TS-2049 - The following general principles shall be applied for Domain Data Specifications:

- Start by clarification of functional data needs (Domain Data). Communicate with domains representing use cases to understand the functional data needs before deriving the input information.
- Data Specifications are enforced and managed by the SD1.
- SD1 also takes the role of a potential Data Producer that has the need to structure and evaluate these needs, e.g. regarding model consistency or technical/economical feasibility aspects (i.e. accuracy requirements vs technical issues vs costs/timing issues)
- Cover schemas in data specifications. On a technical level, data specifications handle schemas of entities and events. They also prevent changes that are not backwards-compatible, such as dropping an attribute in a central object required by many users.
- Cover semantics in data specifications. If the underlying meaning of the generated data is altered, it should break the specification. For instance, if the entity has distance as a numeric field, and the field is used by delivering data in miles instead of kilometres, this alteration is a breaking change. This means the specification should include metadata about the schema describing the data and adding value constraints for certain fields (e.g., temperature).

[Content to be approved**]**

8.1.2 Content of Domain Data Specification

SPT2TS-2050 - The Domain Data Specification shall at least contain the following information:

- **Data Content:** Which objects and attributes are required for the use case? While the first version delivered by the domain does not require a concrete data structure, the final and agreed version of the specification will define the concrete schema incl. semantics, as defined for the SPT2TS-2040 - CCS/TMS Data Model itself in the TCCS SD1 - Data Model_02_Schema
- **Functional linking**: Which parts of the data are required for which functionality (e.g. segment profile for Automated Train Operation, or map data for map matching/sensor validation in localisation algorithms)
- **Data Quality**: What are the data quality requirements for the required data (i.e. accuracy of distances/coordinates incl. tolerance/uncertainty/confidence, or safety-criticalness of information, time-criticality regarding updates)
- **Data Users**: are the data users known (i.e. trackside systems), or are the number of users unknown before operation (e.g. trains consuming map data).
- **Responsibility:** what is the expected responsible entity to provide the information (e.g. infrastructure manager, railway undertaking, supplier of x...)

- **Volatility**: Expected update frequency and trigger for updates of the required information (e.g. stable until topology change, daily change, seasonal change,...)
- **Structure**: are there any structural constraints, i.e. based on the transmission (e.g. radio, wire), functional needs and/or existing specifications for interoperability, such as Subset 126 conform ATO Segment Profiles?
- **Data Security**: does the data contain information relating to security and governance (e.g. anonymization), and are there access limitations expected?

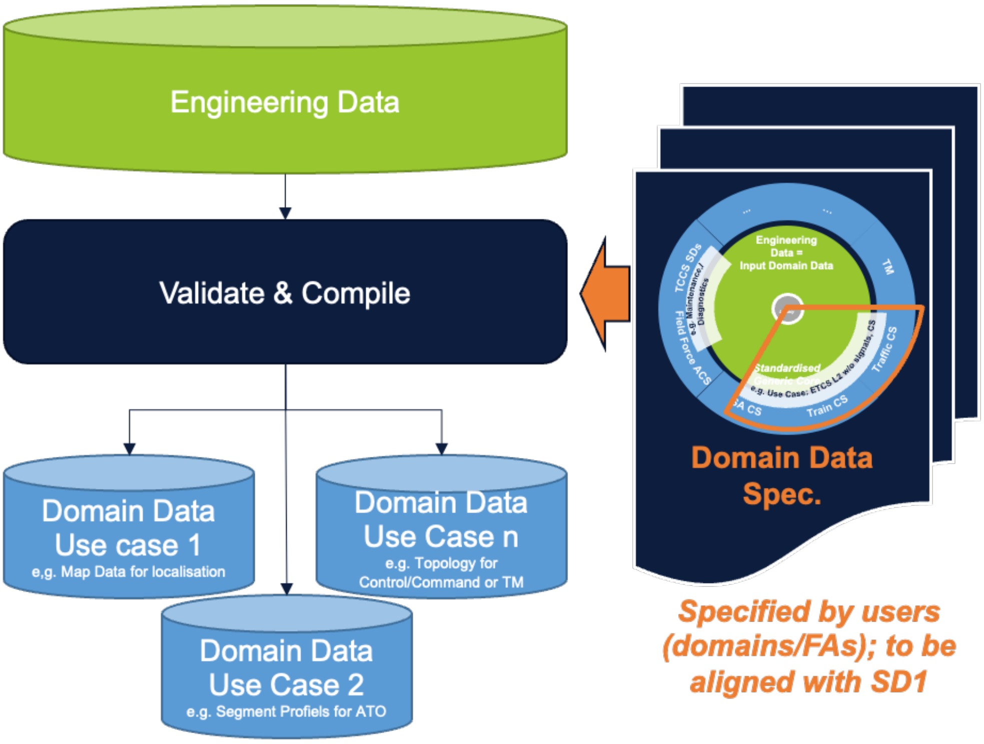

Figure 7 Domain Data Specifications

[Content to be approved**]**

8.2 Formal Specification and Schema Transformation

SPT2TS-2043 - To provide a model that is unambiguous and very close to implementation, the SD1 decided to use a formal, textual specification including meta-information such as semantics and linking to source data models. Based on the experiences in former modelling or implementation projects, a JSON notation has been selected and described as schema TCCS SD1 - Data Model_02_Schema. The data model itself is explained and documented in TCCS SD1 - Data Model_01_Introduction. The definitions of this document are automatically exported into technical schema files (e.g. XML/XSD, json,..) that can be used for the first implementation, demonstration and testing purposes to support a high grade of maturity even in early development stages. **[**Content to be approved**]**

The following excerpt from TCCS SD1 - Data Model_01_Introduction gives an impression of the chosen JSON notation (Example: Topology):

SPT2TS-1697 - Formal specification "Topology":


```
{
 "structs": [
 {
   "name": "TrackEdge",
   "info": "A track edge is an uninterrupted stretch of railway track, without divergence or convergence.",
   "belongsToSubPackage": "topology",
   "see": "http://ontorail.org/rsm12/Common/Topology/LinearElementWithLength",
   "attrs": [
      {"intId": 1, "name": "id", "dataType": "string", "key": true, "info": "Identity of the track edge, needed for referencing"},
      {"intId": 2, "name": "name", "dataType": "string", "info": "User-friendly name. Empty string, if equal to the id"},
      {"intId": 3, "name": "length", "dataType": "uint32", "units": "m", "exp": -3,
        "info": "Distance along the TrackEdge's 3D-alignment. Use zero if not defined."}
    ]
 },
 {
   "name": "TrackEdgeLink",
   "info": "Defines a relation between two track edges along which a train can run.",
   "see": "http://ontorail.org/rsm12/Common/Topology/PositionedRelation",
   "belongsToSubPackage": "topology",
   "attrs": [
     { "intId": 1, "name": "id", "dataType": "string", "key": true, "info": "Identity for referencing, e.g. by points and crossings"},
     { "intId": 2, "name": "trackEdgeA", "reference": "TrackEdge", "info":"Connects to track edge A"},
     { "intId": 3, "name": "trackEdgeB", "reference": "TrackEdge", "info":"Connects to track edge B"},
     { "intId": 4, "name": "startOnA", "dataType": "boolean", "info": "True when linked to the start of track edge A, false when to the end"},
     { "intId": 5, "name": "startOnB", "dataType": "boolean", "info": "True when linked to the start of track edge B, false when to the end"}
   ]
 },
 {
   "name": "TopoArea",
   "info": "Defines a container for rail network topology",
   "belongsToSubPackage": "topology",
   "attrs": [
     {"intId": 1, "name": "id", "dataType": "string", "key": true},
     {"intId": 2, "name": "trackEdges", "composition": "TrackEdge", "multiplicity": "0..*", "sorted": true},
     {"intId": 3, "name": "trackEdgeLinks", "composition": "TrackEdgeLink", "multiplicity": "0..*", "sorted": true}
    ]
  }]
}
```

[Deleted]

SPT2TS-2044 - The formal representation also allows automatic transformation into specific schemata for different applications (protobuf, xml/xsd, or further development/documentation (e.g. UML). The tools for transformations and resulting schemata will be provided as part of the next releases. **[**Content to be approved**]**

SPT2TS-1657 - The model is primarily documented, versioned and stored in Polarion. The model is exported with each release and parsed into the mentioned data schemata. These schemata allow unambiguous standardisation and fast integration into development tools or demonstrators with test data, i.e. Innovation Pillar. This practical usage of the model creates a short feedback loop so that a high level of maturity is quickly attained.

In addition, a translation into UML (e.g. plantUML, XMI) can be used for overall model visualisation, which is automatically created and always consistent with the model itself. The bridge to Capella (if available to Polarion) will be used to achieve a synchronised model view with the System Pillar architecture and all other domain stakeholders working with Capella.

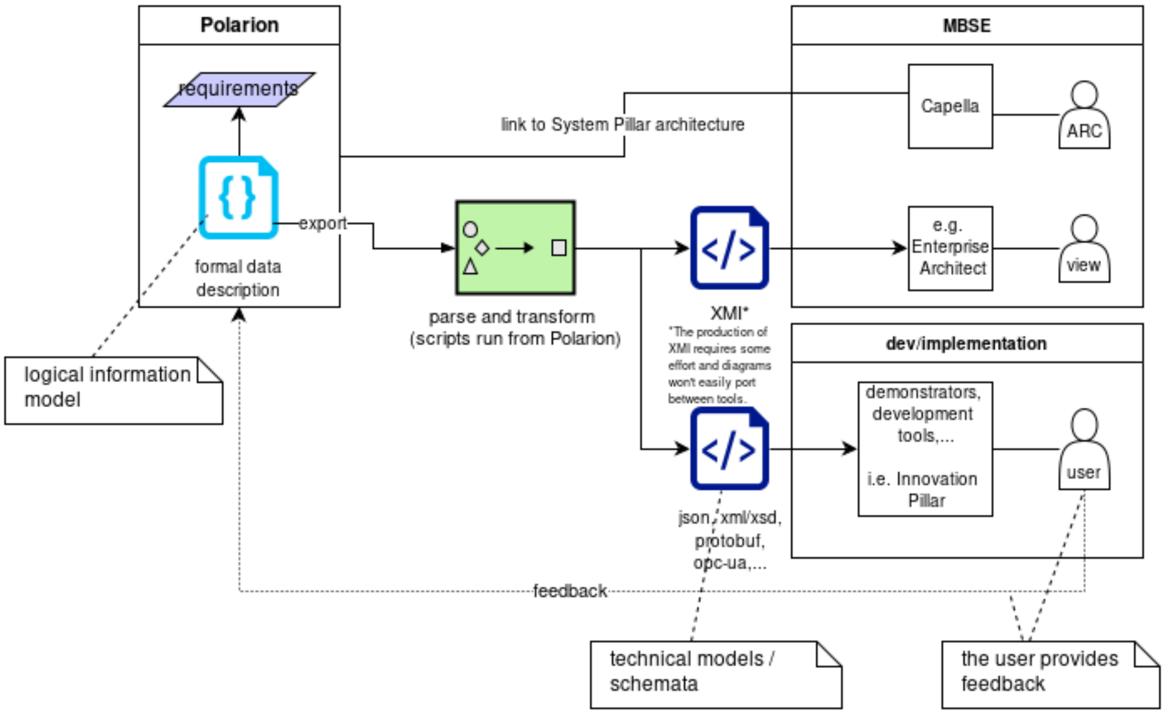

Figure 8: Handling the data model. The data model is a formal description of the object types and their relations. The ensuing technical information models are the schemata for transporting actual data over-the-wire or as files.

[Content to be approved**]**

8.3 Usage of Existing Models

SPT2TS-2045 - As pointed out before, the approach of SD1 is to avoid inventing new models but reuse existing concepts for the specific purpose of SD1 (described in 6- Scope of TCCS SD1). The following figure shows an excerpt of possible candidates, such as RCA Object Catalogue, EULYNX Data Prep, Linx4Rail CDM, or X2R4 data model:

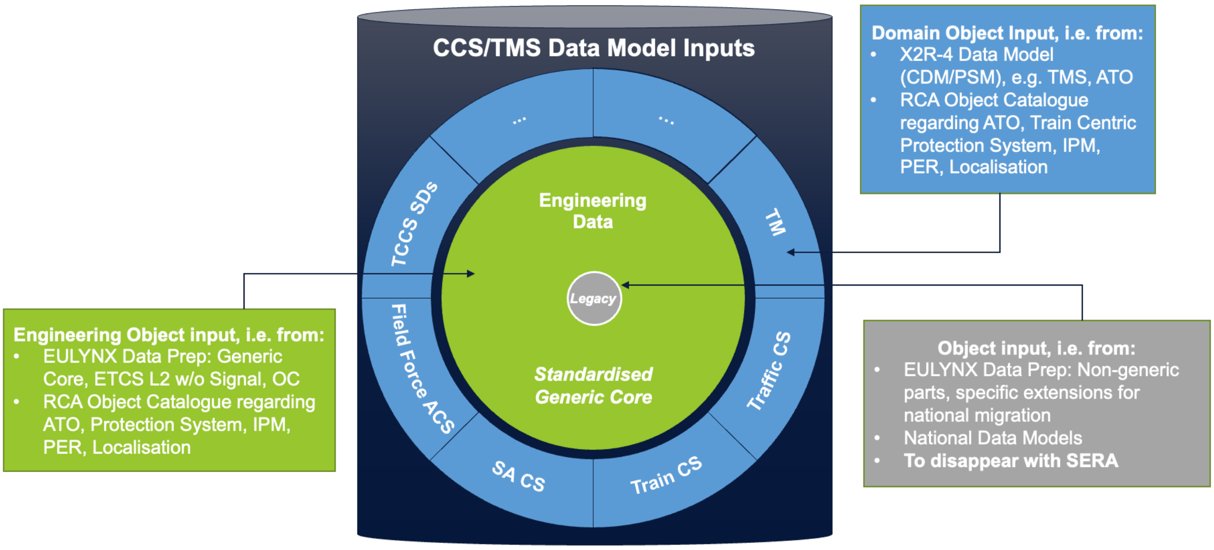

Figure 9 : examples for candidates of possible data model inputs

[Content to be approved**]**

8.4 Bridge from SD1 to other SP tasks and IP FAs

SPT2TS-2046 - Since the SPT2TS-2040 - CCS/TMS Data Model of SD1 is focused on CCS/TMS-related use cases and systems, it must be possible to connect to external data models covering external use cases. It is planned to use the Innovation Pillar FP/FA 1 work package 26 to aggregate additional data needs from the FAs and mirror them against the development of the SPT2TS-2040 - CCS/TMS Data Model within System Pillar. The connection to the external data model can be defined and managed by the support of the ontology approach (ERA Vocabulary/Ontology) with a centralised data dictionary and semantic linkings between models (see 6.6- Linking to other Data Models with Ontology)

In addition, direct exchanges to data-related Innovation Pillar FA Work Packages (e.g. FA1 WP 27 "Digital Asset Engineering", FA2 WP 27 "Digital Register" and connected WPs) are established to get feedback from implementations dealing with the same topic. **[**Content to be approved**]**

9 Evolution to Target Architecture

SPT2TS-2047 - The evolution to a target architecture requires intermediate stages enabling first use cases as early as possible.

The Domain Data Model, including engineering information, is built up in several increments based on the knowledge of the System Pillar systems and the prioritized use cases. At least for the migration to a harmonised data provisioning process, it is expected that several ways to publish new data versions must be supported (e.g. file-based loading and "on-the-wire" provisioning in parallel). However, the shared information needs should be provided by a single source of truth following the principles (data model, versioning, id-management,..) of the interface specifications defined by SD1 together with the SP architecture and domains. To protect investments and enable economic migration scenarios backward compatibility of all future extensions with the next TSI increments becomes crucial. This aspect is considered during the development of the Data Model itself. **[**Content to be approved**]**

SPT2TS-48934 - As starting point of already TSI defined use cases such as ATO (Subset 126) or ETCS (Subset 026 and 036) are used as profound base to build up the data model, even before the System Pillar architecture is evolved yet.**[**Content to be approved**]**