EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

9 April 2015
EMA/187520/2015
Information Management Division

Annex to the European Medicines Agency (EMA) Master Data Management Roadmap

Substance, Product, Organisation and Referential data

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Table of Contents

1. Implementation approaches for SPOR master data 3
1.1. Referentials MDM implementation 3
1.1.1. Initial analysis 3
1.1.2. Referential MDM implementation roadmap 3
1.2. Substance MDM implementation 4
1.2.1. Initial analysis 4
1.2.2. Substance MDM implementation roadmap 5
1.3. Product MDM implementation 7
1.3.1. Initial analysis 7
1.3.2. Product MDM implementation roadmap 8
1.4. Organisation MDM implementation 8
1.4.1. Initial analysis 8
1.4.2. Organisation MDM implementation roadmap 10
2. Roadmap development methodology 11
2.1. Description of the current state of data management 11
2.2. Description of the desired future desired state 13
2.3. Gap analysis 18
2.4. Prioritisation and optimisation of actions - approach 20
2.5. Recommendation based on prioritisation and optimisation analyses 20
3. Document history 21
4. Glossary 22
4.1. Key Terms 22
4.2. Systems 23

1. Implementation approaches for SPOR master data

1.1. Referentials MDM implementation

1.1.1. Initial analysis

Analysis of Referential data sources, maintenance and usage by stakeholders focused on the EUTCT system from the business and IT perspectives in the following areas:

- The connection between EMA and external (NCA) systems and EUTCT to consume/provide lists and/or their values to build a picture of the IT architecture.
- Understanding the interaction between EUTCT and stakeholder systems is required to plan the data migration approach.
- Determining the source/reference of the lists in EUTCT.
- Externally sourced lists may require different access and security controls in the new 'Referentials' MDM system.
- Identify which EUTCT users will be impacted by the data migration work, and what effect it will have on their operations.
- Identify the processes to update the EUTCT lists. The processes are often different whether they relate to internally or externally sourced lists.
- Better understand the EUTCT functionality available to the end users.
- This will ensure that the new user interface maintains at a minimum the current functionality and enables identification of improvements for the new system.
- Identify the existing levels of access and security for the different types of end users.

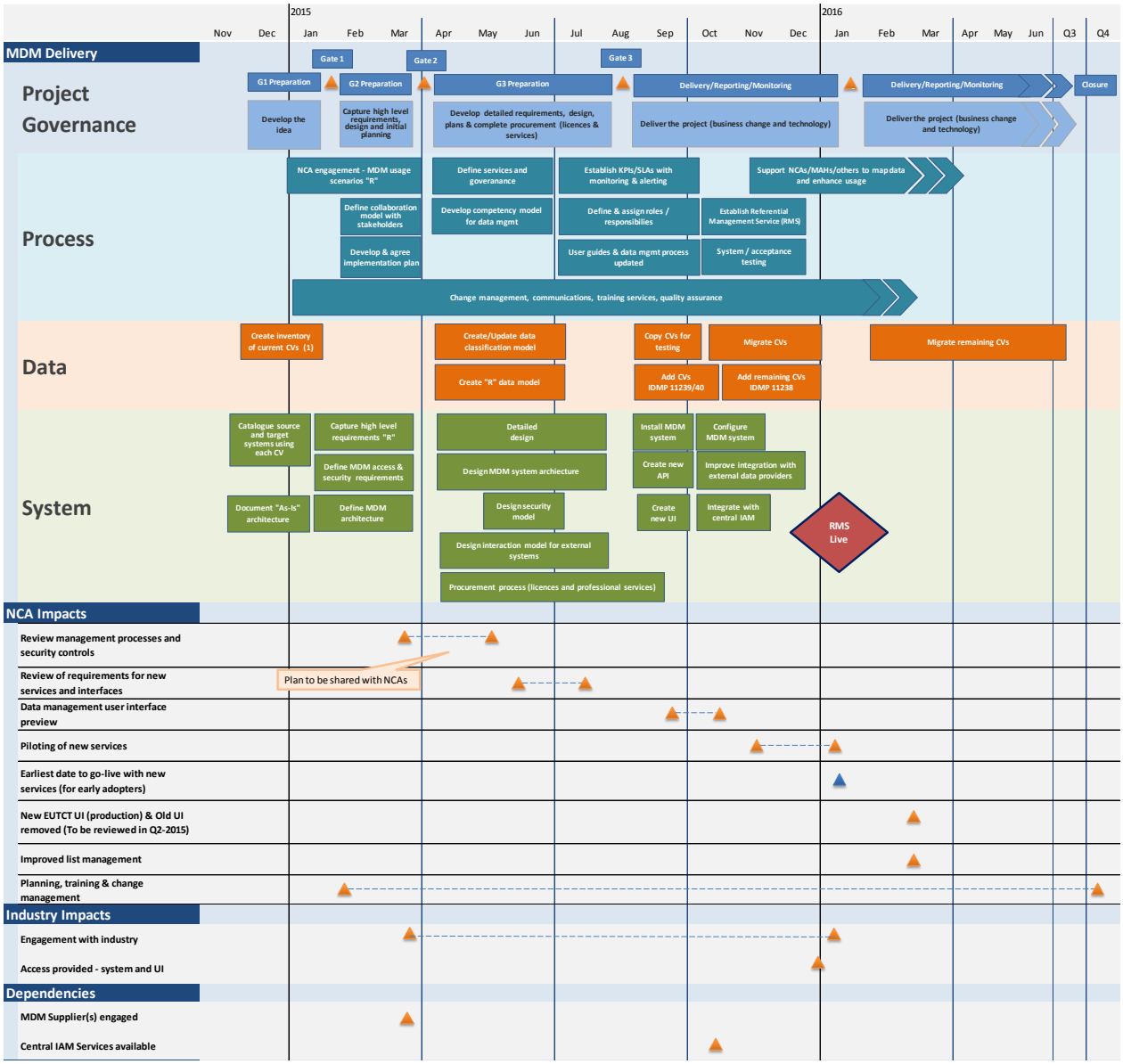
The vast majority of the existing EUTCT lists are created and sourced from within EMA (79%). Analysis of usage of the 77 lists in EUTCT by NCAs and EMA systems found significant variation in usage among and between systems.

The findings show that both the EUTCT graphical user interface and the system interface are currently used in various degrees by a considerable number of National Competent Authorities (NCAs). These findings will be taken into account when designing the migration approach from the current EUTCT system to the future MDM, which will have to be communicated to the stakeholder community involving them throughout the process.

1.1.2. Referential MDM implementation roadmap

Fig. 1 shows the key project activities for the Referentials MDM implementation. The development of new 'Referentials' master lists will continue alongside Substance, Product and Organisation MDM implementation.

Figure 1. Referential MDM implementation roadmap plan



1.2. Substance MDM implementation

1.2.1. Initial analysis

Analysis of substance data focused on human EudraVigilance (EV) systems storing Article 57 data in the eXtended EudraVigilance Medicinal Product Dictionary (xEVMPD) format, where the majority of Substance data for use in human medicines is managed. This system will become the initial source of substance data for the future substance MDM system. The data structures for storage and exchange of substance data in the MDM system will be based on ISO IDMP 11238 specification. Internal data storage data model might be further enriched to support internal EMA business processes.

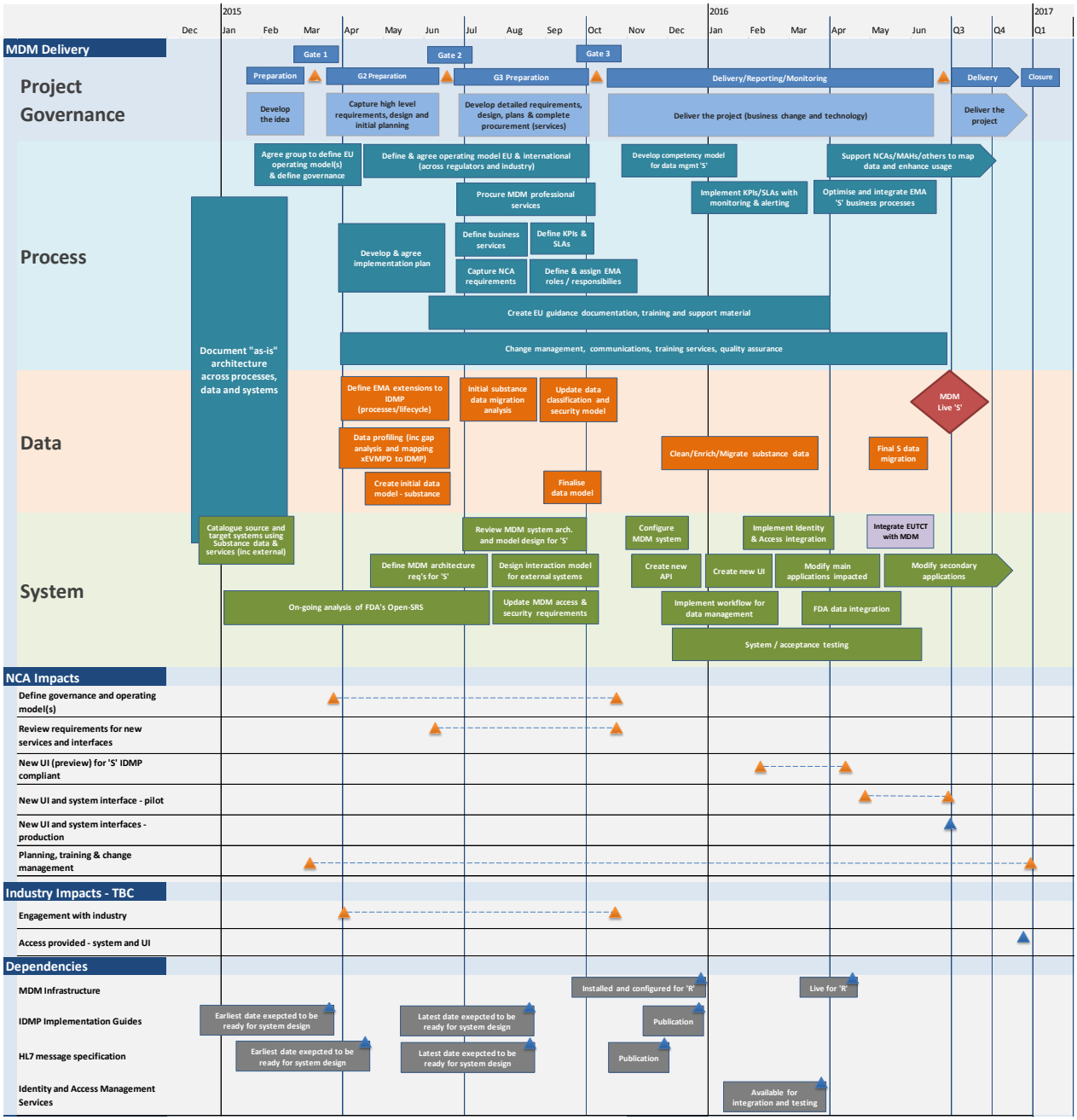
This initial analysis took into account EV and other directly or indirectly connected systems. The following points were identified:

- Several EMA systems depend on the substance data stored in the human EV system. Not all of them will be able to migrate to new MDM solution or to the extended ISO IDMP data model at the same pace.
    - Transition solutions and migration projects need to be designed for each individual system and related business processes.
    - Backward compatibility will need to be provided from the new MDM system to the current consumers of the EV provided substance data. Several implementation methods are technically possible and will be subject to further analysis during the implementation project.
- The xEVMPD data format stores only a subset of data attributes required by the ISO IDMP standard. Migration of the xEVMPD data to the MDM systems will require data enrichment. The extent of data enrichment will be determined during the project analysis and design phase.
- Currently there are different processes for registering authorised and development substances in the xEVMPD. They will need to be harmonised to allow registration in the MDM systems all types of substances.
- There is number of external dependencies and constraints that might influence the implementation of the substance MDM system:
    - Processes, roles and responsibilities to enrich xEVMPD data to the ISO IDMP data format are no yet defined.
    - ISO IDMP implementation guidelines are still under development. Although they are very advanced, there still is a risk of changes that will influence the implementation.
    - Operating model for registering and maintaining substance data in Europe still needs to be decided. This includes EU specific ISO IDMP implementation guide, decision on centralised or federated model for substance registration, assignment of maintenance organisation of substance data for Europe and globally.

1.2.2. Substance MDM implementation roadmap

A draft roadmap plan is outlined in *Fig.2*. It describes the major activities required to deliver the Substance MDM service, expected dependencies, and envisaged impacts on the NCAs.

Figure 2. Substance MDM implementation roadmap plan



Notes

The approach to enriching substance data to become IDMP compliant will be identified during the process modelling.

The IDMP data model implemented will be designed to support compliance with IDMP but, depending on the EU region implementation guidance, may not include all possible IDMP data fields.

1.3. Product MDM implementation

1.3.1. Initial analysis

The scope of this analysis covered the two primary EMA systems that store product data, namely EudraVigilance (EV) and Siamed. Other product data containing systems (e.g. Orphan db., Paediatric db., Scientific Advice db. and Veterinary systems) will be analysed as part of the Product MDM implementation projects.

xEVMPD (as part of EV) includes National as well as Centrally Authorised Products. Siamed stores Centrally Authorised Products only and supports the MA regulatory procedures. EV, on the other hand, supports the Pharmacovigilance activities such as the ADR monitoring and clinical trials activities.

Siamed product data is registered and updated based on the MA application submission data and based on the outcomes of their procedures. Whereas, the product data in xEVMPD is registered and updated following Art. 57 data submissions by MAHs.

This preliminary analysis of Siamed and xEVMPD systems was conducted to assess the current state, the possible impact of implementing the product MDM for these systems and other systems connected to them, as well as identifying potential constraints and dependencies from business and IT perspectives.

The following key points were identified:

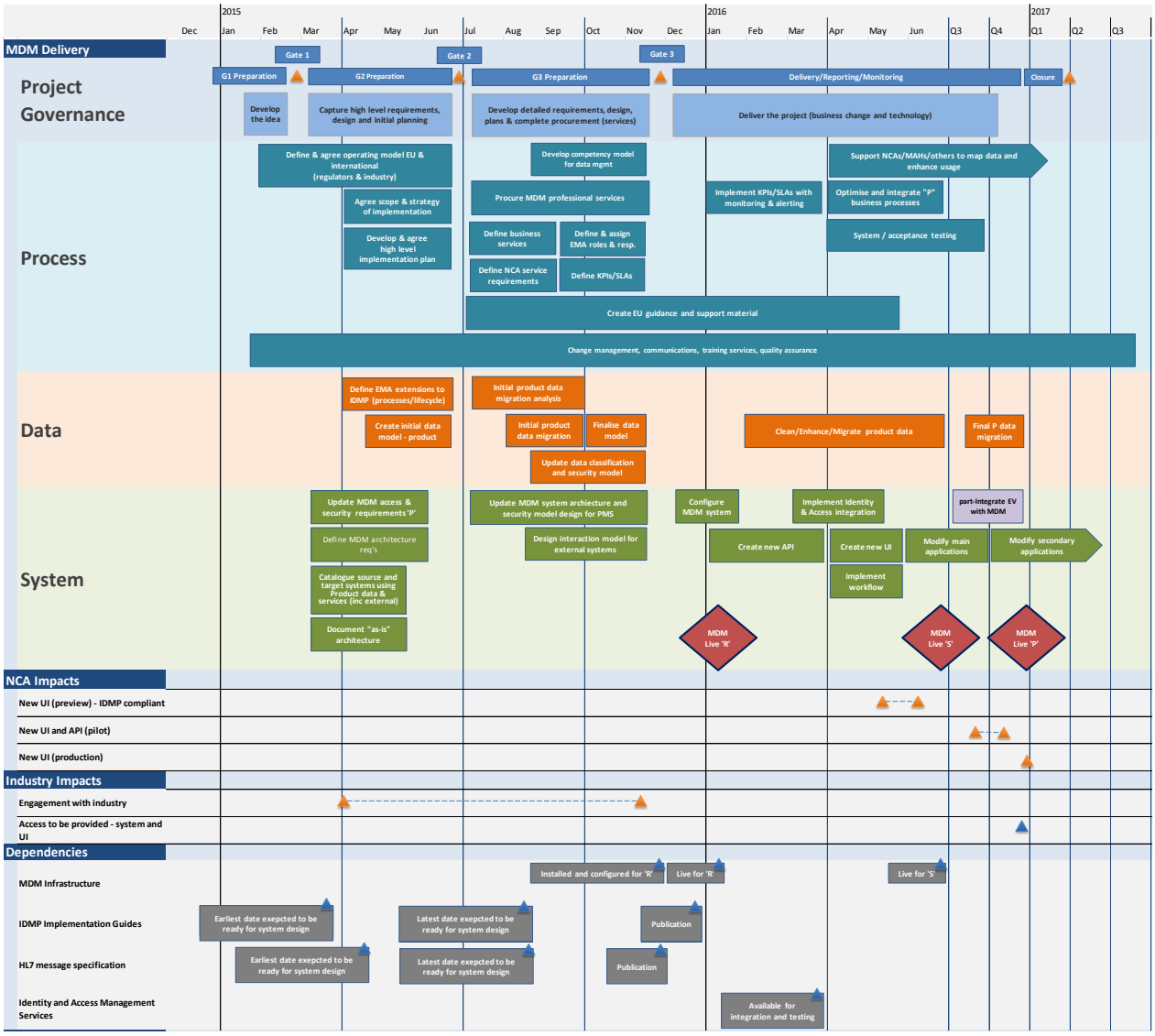
- There will be a considerable number of systems impacted by the product data migration to the future MDM. Some of these systems would require adaptation work such as Siamed and EV. The extent of adaption required for the secondary systems connected to Siamed and EV will depend on data related changes implemented in Siamed & EV.
- Siamed and xEVMPD have different data needs, and manage product data from different perspectives e.g. Siamed stores data such as 'product resources' or 'product contacts' at the level of medicinal product group, which is not a concept that exists in ISO IDMP. During the transition phase, ISO IDMP implementation in MDM should be able to support the existing Siamed and EV approach/functionality in managing product data.
- Data Quality levels between Siamed and xEVMPD vary. Both systems are going through data quality improvement initiatives throughout 2015.
- The authoritative source for Product data to be migrated to the future MDM system can come from Siamed, EV or a combination of the two, facilitated by the use of an MDM tool to compare and consolidate data. Further analysis is required to define the rules for consolidation of product master records within the MDM tool.
- In common with the Substance MDM project there is a number of external dependencies and constraints that might influence the implementation of the MDM system:
    - Processes, roles and responsibilities to enrich xEVMPD data to the ISO IDMP data format are not yet defined.
    - ISO IDMP implementation guidelines are still under development. Although they are very advanced, there still is a risk of changes that will influence the implementation.
    - Operating model for registering and maintaining product data in Europe still needs to be decided. This includes EU specific ISO IDMP implementation guide, decision on

centralised or federated model for product registration and data validation and the approach to assignment of Product IDs and Pharmaceutical Product IDs.

1.3.2. Product MDM implementation roadmap

A draft roadmap plan is outlined in *Fig. 3*. It describes the major projects required to deliver the Product MDM service, expected dependencies, and envisaged impacts on the NCAs.

Figure 3. Product MDM implementation roadmap plan



1.4. Organisation MDM implementation

1.4.1. Initial analysis

Organisation MDM covers the broadest spectrum of data management throughout EMA and hence it can be one of the most complex for data integration. The criteria to define the scope of the first project

were to minimise the risks and to deliver quicker benefits to stakeholders. The proposal for the scope of the initial project focuses on supporting the high priority EMA programmes/business activities.

The implementation of Organisation and contact MDM should be based on a single data model and standard. It should be generic enough to support the different business activities and processes taking place throughout the medicinal product life cycle. The implementation approach should be iterative.

Iteration 1 includes organisation data supporting the management of organisations acting as Marketing Authorisation Holders. Through the iterations 2 and 3, the existing list of organisations in MDM will be extended by the inclusion of organisations acting as Marketing Authorisation Applicants and Sponsors. The implementation of Sponsor organisation may also include the management of Sponsor contacts. Further iterations are also possible. The detailed scope will be defined during the first phase of the project.

The full extent of the scope for the first project will be further defined at the beginning of the project.

In addition, the following should be considered:

- The extent of adaptation introduced in the existing systems will be minimised as much as possible.
- New service to access Organisation data (list of organisations) to be provided via a system interface. There is also the possibility to provide a user interface for external stakeholders to query the dictionary.
- Implementation of Organisation MDM should implement a generic organisation data model which will have to be easily extendible to support additional business activities in the future.
- The possibility of using an existing data model for organisations (and contacts) master data. Further investigation should be performed on the different solutions already available in the market and/or implemented by other organisations.
- A codification system which uniquely identifies organisation entities will need to be selected e.g. DUNS, IFA, GS1, etc.
- The consolidation of organisation master data will require the sourcing the data from multiple EMA systems, which adds to the complexity of the implementation.
- The target and transitional operating models for registering and maintaining organisation master data will need to be defined during the analysis and design phase of the initial project. Most likely, the process of mastering data during the transition phase and in the final solution will need to be different to allow incremental and progressive master data implementation.
- It is yet unclear on the dependencies to support the Identity and Access Management (IAM) system which will be development in parallel with the SPOR projects. Furthermore, this analysis will also need to be supported by more clarity with regards to confidentiality of organisation data before it is made available to the stakeholders or the public.
- The full implementation of organisation MDM will be instrumental for the development of the single submission portal which is planned for 2018 under the eCollaboration programme.

Expected benefits from implementing the initial Organisation MDM services:

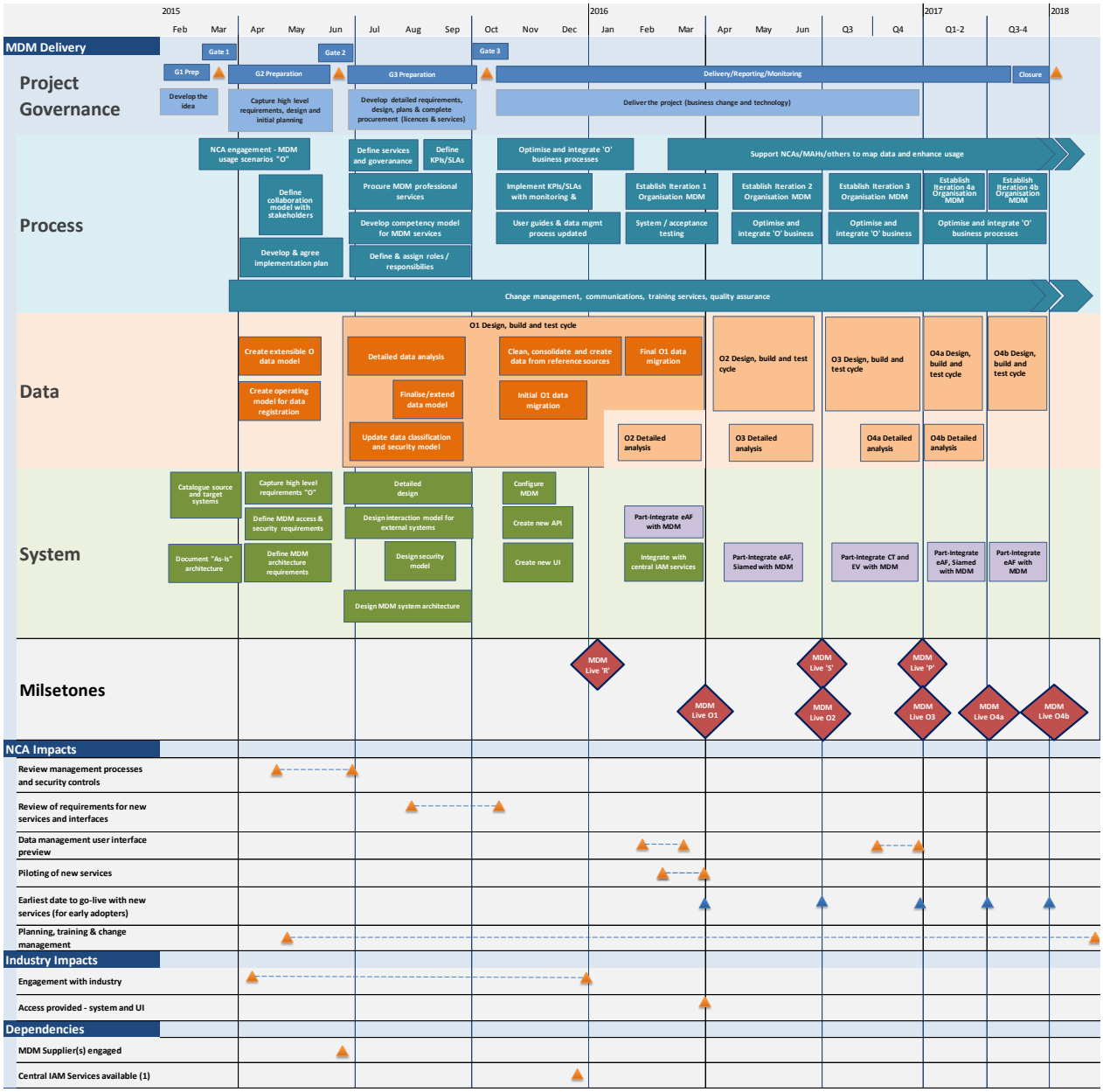
- Provides a single dictionary of organisations (MAHs, MAAs, Sponsors, and Manufacturers) for the whole of EEA, which will be easily accessible for reference and data exchange.

- Improved data quality with the introduction of more structured information on eAF (human) forms. The veterinary eAFs will be supported at a later stage during the implementation of this roadmap.
- Establishes the foundation for future integration between regulatory, clinical trials and Art. 57 submissions and the MDM system.

1.4.2. Organisation MDM implementation roadmap

A draft roadmap plan is outlined in *Fig. 4*. It describes the major projects required to deliver the Organisation MDM service, expected dependencies, and envisaged impacts on the NCAs.

Figure 4. Organisation MDM implementation roadmap plan



2. Roadmap development methodology

2.1. Description of the current state of data management

To understand the current state and known issues of SPOR MDM, a series of interviews were conducted covering EMA business areas, two representative NCAs (Germany and Sweden) and representatives from EMA's strategic programmes. The scope of the interviews at EMA encompassed data management within:

- Pharmacovigilance programme
- Clinical Trials programme
- eCollaboration programme
- Online programme
- Review & Reconnect programme
- Medicinal Product development processes (i.e. paediatric, SA and Orphan)
- Evaluation of medicines processes
- Pharmacovigilance fees (finance processes)
- Procedure management processes
- Veterinary processes
- Data protection

The key themes emerging from these interviews fall under business processes, data, people/organisations and IT. The valuable perspectives we gained from our stakeholders are summarised below.

Business Process Perspective

Interviews identified opportunities for business process improvement relating to SPOR data management:

- There are many disjointed but similar processes followed in the management of SPOR data operating within different business areas and processes across EMA. There is a lack of an enterprise wide approach to implement SPOR data management processes with integrated SPOR lifecycle data management.
- Many of the processes mentioned in the interviews are manual and thus time consuming and inefficient. There is significant duplication in data management processes across the business, particularly in the area of SPOR. There are poor data quality controls for the management of SPOR data in general and in particular for the managing the SPOR data in the reference source systems.

Data Perspective

- SPOR data is stored in different data formats used by different systems, with much SPOR data stored in free text - unstructured format. Differing terminology and business rules are used by

different systems. Data dictionaries are not defined equally across the organisation, leading to conflicting values and inconsistencies across same dictionary types.

- SPOR data originates from many unstructured data inputs. The data quality is known to be poor in some areas. Manual data quality controls and registration are common place. Data security generally has a low profile, with the benefits remaining unknown to many staff.
- There is very little metadata managed on SPOR which is needed to support online publication.

People/Organisational Aspects

There is currently limited governance of data and the benefits of having good data governance. Across the EMA, data management skills and data quality awareness could be improved through learning and development initiatives.

IT Perspective

No system or process is currently in place for overall SPOR MDM. IT systems used for SPOR data management currently operate in silos, and are built to different data models. Some IT systems are no longer supported compounding problems caused due to the silos and lack of interoperability between systems.

Across the EMA, IT systems tend to use local databases with differing approaches for auditing of changes, resulting in a lack of data traceability and presenting an audit risk. In terms of usability, the data management interfaces of EMA systems are perceived to be laborious to work with highlighting that improvements in efficiency could be realised here.

Many systems have embedded logic for managing data quality through their own business rules and lack configuration options allowing data quality consistency. This leads to high dependency on back end IT development needs and little flexibility for business users to adapt the system to business changes.

The issues identified can be further summarised as:

- Missing EMA enterprise perspective on master data and processes
- Silo processes
- Silo applications
- Silo databases
- Low levels of data quality
- Duplicated data stores resulting in multiple versions of the truth
- Lack of clear ownership of data, data quality and processes

Table 1 provides some indicative volumetric on core EMA systems storing SPOR data.

Table 1. Strategic EMA systems currently storing and managing SPOR master data

System	Manage (create, update, store)				Volumes of records ('000s)			Human / Vet
	S	P	O	R	S	P	O	
EudraVigilance 7	*	*	*	*	39	594	5.4	Human
EudraCT 8	*	*	*	*	51	484	3,880	Both
SIAMED	*	*		*	>0.8	>3.6	N/A	Both
ECD			*	*	N/A	N/A	52	Both
SAP			*		N/A	N/A	9	Both
EUTCT	*			*	39	N/A	N/A	Both
ESVAC	*	*	*	*	10.5	8.7	N/A	Vet
EudraVigilance (Vet)	*	*	*	*	20	11	0.3	Vet
EudraGMDP	*	*	*	*	>18	N/A	>14	Both
Orphan Drugs	*	*	*	*	>4.3	>2.7	>1.4	Human
Paediatrics	*	*	*	*	19	0.4	1	Human
Scientific Advice	*	*	*	*	>4.9	>4.9	>2.3	Human
All Other Systems	Over 40 applications which 'Manage' and/or consume SPOR data[1]				N/A	N/A	N/A	Both

2.2. Description of the desired future desired state

The interviews with the EMA's business and NCA representatives identified several high level goals, specific actions and requirements and proposed actions to bridge the gaps. All these inputs were taken into consideration during work on this roadmap.

Business Process Perspective

- Interviews clearly stated the need for integration and optimisation of existing business processes combined with the introduction of standardised SPOR MDM processes.
- Existing processes should be standardised to reduce duplicities (regarding to SPOR data management).
- Processes for collecting, validating, storing, maintaining and reusing SPOR data should be implemented from an enterprise perspective and automated. This would provide considerable efficiency gains, reduce registration errors and duplications during the Product development phase and improve collaboration with NCAs.
- New data quality monitoring processes should be established. These processes should monitor 'data quality', 'how data is used', 'how data is shared, created and discarded' as well as retention policy aspects.
- New processes need to be defined to allow NCAs to easily access and maintain SPOR master data in order to support their own regulatory procedures. Additionally other users (individuals and organisations) should be able to have direct and easy access to their user/personal details and manage it directly as much as possible.

[1] This figure does not include media such as Excel spreadsheets, which are used in silos to manage SPOR data in support of specific business activities.

- Process governance should be clearly documented and communicated where accountability, responsibility, transparency and performance monitoring is established.

Data Perspective

- The main data related goals identified by the interviewed business representatives are to improve data quality, reliability and availability in order to deliver business benefits.
- The SPOR master data was indicated as a key element to underpin the improvement of business processes. This data should support more efficient work and meeting business and regulatory timelines.
- Interviewers suggested that data quality could be improved by wider use of structured data inputs accompanied by data quality checks and business rules enforcement.
- High quality and widely used Referential data was seen as important to simplify business specific processes and minimize the need for additional quality checks and controls.
- It is necessary to define what a master data is within SPOR (i.e. define SPOR master data model), taking into account requirements of specific business processes. Business interviewers appreciate the fact that master data can be enriched in individual systems to support their specific business purposes.
- The interviews indicated a possible need for ISO IDMP data model extensions to meet specific EMA requirements and the differences between business rules and data sets used by human and veterinary processes.
- There is a strong need for data lifecycle management processes. Metrics and dashboards for data quality need to be defined taking into account different data quality requirements based on various criteria (e.g. type of Product, risk level, etc.).
- It is a necessity for data to be shared with internal and external users, taking into account data protection requirements. There is also a need to allow external parties to manage their data sets, while applying security and privacy policies. These are to be supported by identity and access management mechanisms combined with data annotations (e.g. data sensitivity classification) to understand value and possible use of data, which in turn will help to protect data and help defining retention requirements.
- Specific requirements indicated in the interviews included points such as backward compatibility of future ISO IDMP compliant data with current Art. 57 data sets, mechanisms to track data changes and management of historical terms, support for multilingual data sets and others.

People/Organisational Aspects

- The people working with the SPOR master data must support and implement new and modified processes. This requires clear definition and assignment of roles in these processes (e.g. dedicated staff to manage SPOR master data)
- Communication of procedures and training is required to further develop competencies in MDM and in data governance in general. Several interviews suggested centralisation of SPOR master data governance for key regulatory processes.

IT Perspective

- Technology related comments in the interviews focused mainly on providing tools to realise goals identified in other areas. Amongst others, new tools should support automation of business processes, improvements in data quality at inputs and updates and provision of data quality monitoring. Interviews welcomed the idea of implementing an automated and integrated MDM solution for SPOR data capture, creation and updates.
- One of the interviews clearly stated that: "There should be enterprise wide approach taken on the architecture to be put in place to manage all SPOR data whether it is master data or not"
- Several comments highlighted the need to remove the silo management approach between the systems and a reduction in the number of systems used and the overhead of managing them.
- There is a necessity to ensure proper reference of master Referential data and integration of the MDM solution with specialised systems supporting business.

2.2.1. Architectural Vision

The future desired state should be built around a centralised MDM solution to manage master SPOR data. Its introduction and follow-up maintenance and usage will require significant change impacting all aspects of EMA architecture: from reorganisation of business processes, through changes to the data and applications organisation, to the underlying technology implementation. The envisioned high level target data architecture is summarised in *Fig. 5*.

Business Architecture

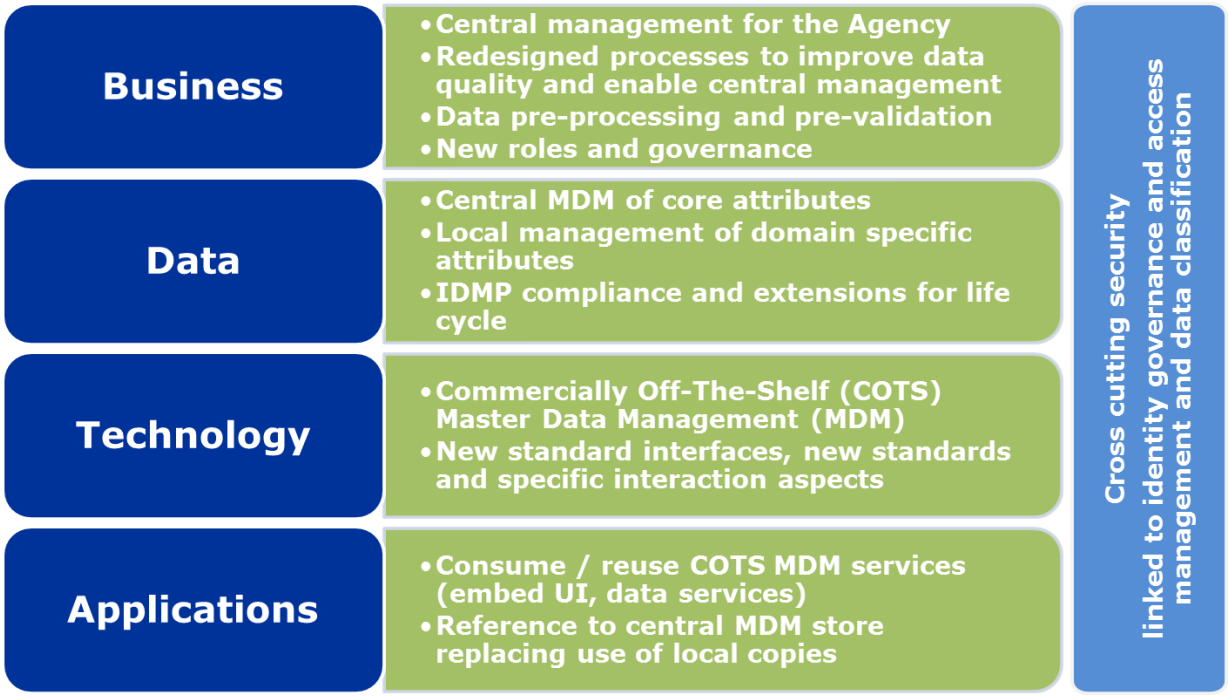
The existing business processes within EMA will need to be integrated and optimised to use and benefit from using the SPOR master data. This will include harmonisation between processes, elimination of duplicated process steps, new data pre-processing and pre-validation process steps and higher process automation. New business services will be created for the management of the SPOR master data and to support all the users of this data. These business services will be underpinned by new data governance processes and roles.

Data Architecture

The overarching data approach is to (over time) have a high-quality single source of "the truth" for SPOR master data. This means that there will be a single view of SPOR master data, which is up to date and accurate according to an agreed standard.

The data model for the SPOR master data will be compliant with the ISO IDMP standards. Depending on the analysis of the master data requirements of the EMA and its partners, the master data model may not incorporate the entire ISO IDMP model (e.g. some optional elements may be excluded initially). On-going efforts to define the ISO IDMP implementation guidance will help to determine it. The relevant ISO IDMP standards for Substance, Product and Referential data are described in the glossary. The data model for organisations will be based on data elements from the relevant ISO IDMP standards, and further refined in co-operation with the partners to meet EMA and their requirements.

Figure 5. Target enterprise architecture for SPOR MDM



It is expected that the master data records stored by the EMA will contain additional attributes or specific business rules to accommodate needs of internal EMA processes. Specific EMA systems will reference the SPOR master data and will internally enhance it with attributes specific for a particular business process and related systems, but not considered as part of the master record.

As a result the data model for the SPOR master data will in its majority be based on the ISO IDMP standards which are required for the exchange of data with the partners and the industry. Additionally, extended attributes used internally by the EMA's business processes will complement the master data model.

The unstructured data is outside the scope of this roadmap. However, it is expected that the SPOR master data will be used to annotate and tag unstructured data to allow for better organisation and search.

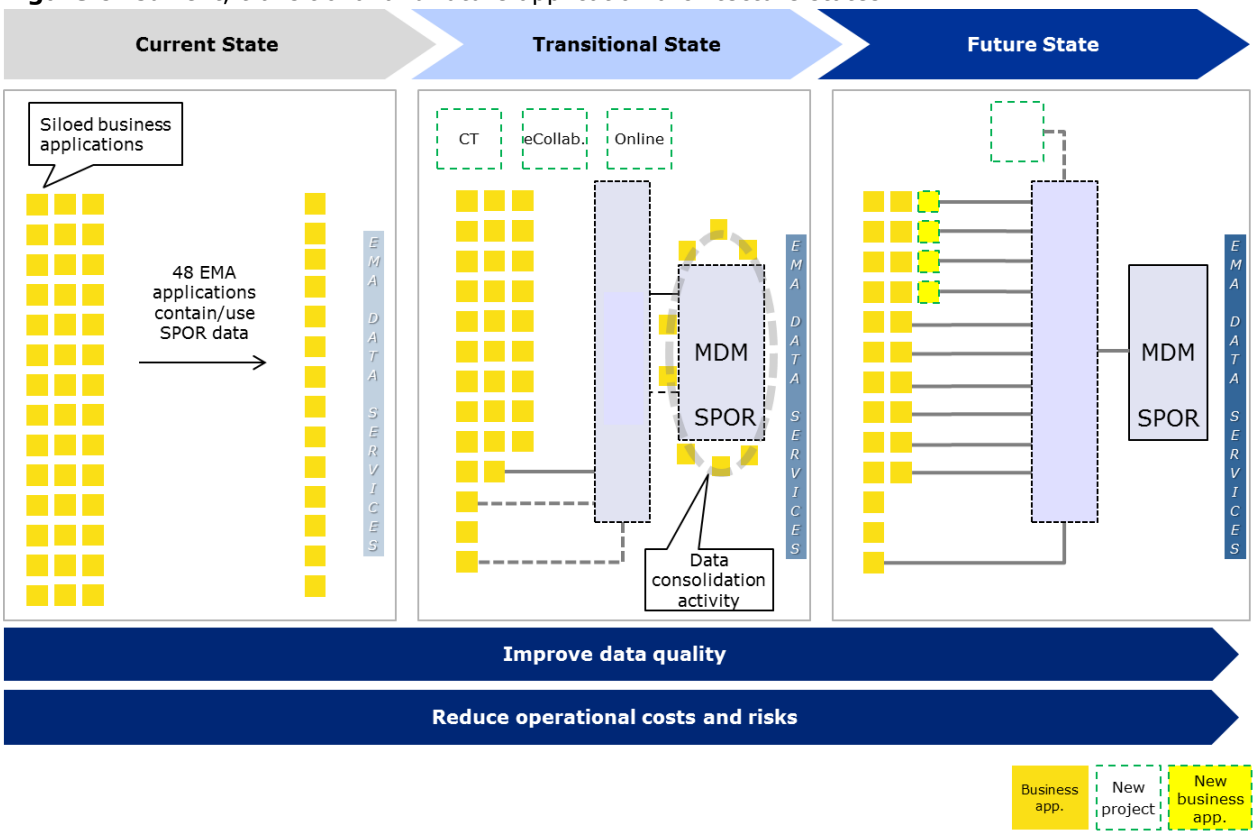
Application Architecture

Existing EMA applications will need to be adapted to work with the new MDM solution. The goal is to remove from applications all functionality related to the management of the SPOR master data and integrate it in the MDM solution. Ultimately, the applications should only reference master data in the MDM solution. If a specific business process supported by a given application requires additional data fields that are not part of the master data model, the application can store and manage these specific data elements locally, but correctly referencing to relevant master data. *Fig*. 6 highlights the transformation from the current application architecture to the future application architecture will need to go through a transition state. Applications managing SPOR data currently will progressively be integrated with the MDM system and will become MDM service users via the 'BUS' interface instead.

Candidate key applications for integration with MDM have been already identified in this roadmap. Further applications will be selected after future analysis. It is expected that the number of applications managing the SPOR data will be reduced and applications with duplicated functionality will be

eliminated or reduced considerably. Newly implemented applications will be designed to natively work with the MDM solution.

Figure 6. Current, transitional and future application architecture states



The current state sees a number of siloed applications containing and using SPOR data independently in support of business processes and with large number of interactions among them. Across the four SPOR data domains, a transitional state will involve the implementation of an interface between the separate applications and a centralised MDM service. The Future state will see SPOR data accessed from a centralised MDM service via a single system interface with improved data quality and reliability. Throughout the stages of implementation a sequential, iterative approach to consolidating data from the local applications into the MDM service provides a lower risk, cost-effective option compared to a one-off 'big bang' approach attempting to consolidate data from all applications in a single project.

Technology Architecture

It is recommended that the technological implementation of this roadmap should follow the Service Oriented Architecture (SOA) approach. The MDM solution will be implemented together with well-defined technical services, programmatic interfaces and interaction models to support automated integration with other systems. This model allows for high level of standardisation and reuse across EMA and stakeholder activities.

2.3. Gap analysis

The current state and desired future state were obtained from interviews with some NCAs and key EMA stakeholders, and an industry position paper[2].

Aims

The objective of this gap analysis is to derive and formalise a series of discrete, tangible actions which will build a strong foundation for the achievement of goals and objectives set out in this roadmap, and support the vision of EMA. Each action should result in or support EMA in meeting the requirements of a diverse group of stakeholders, while recognising the competing priorities between them.

Approach for the roadmap development

The desired future state requirements were grouped into high level distinct areas and mapped against the current state. This enabled the gaps between the current and future states to be identified with a clear link to the stakeholder requirements and vision of the overall roadmap.

Identification of gaps was informed by the vision and goals of the roadmap. The differences between the current state and desired future state were described in terms of business processes (Process), SPOR data management (Data), EMA people and organisational aspects (People) and system and application capability (IT).

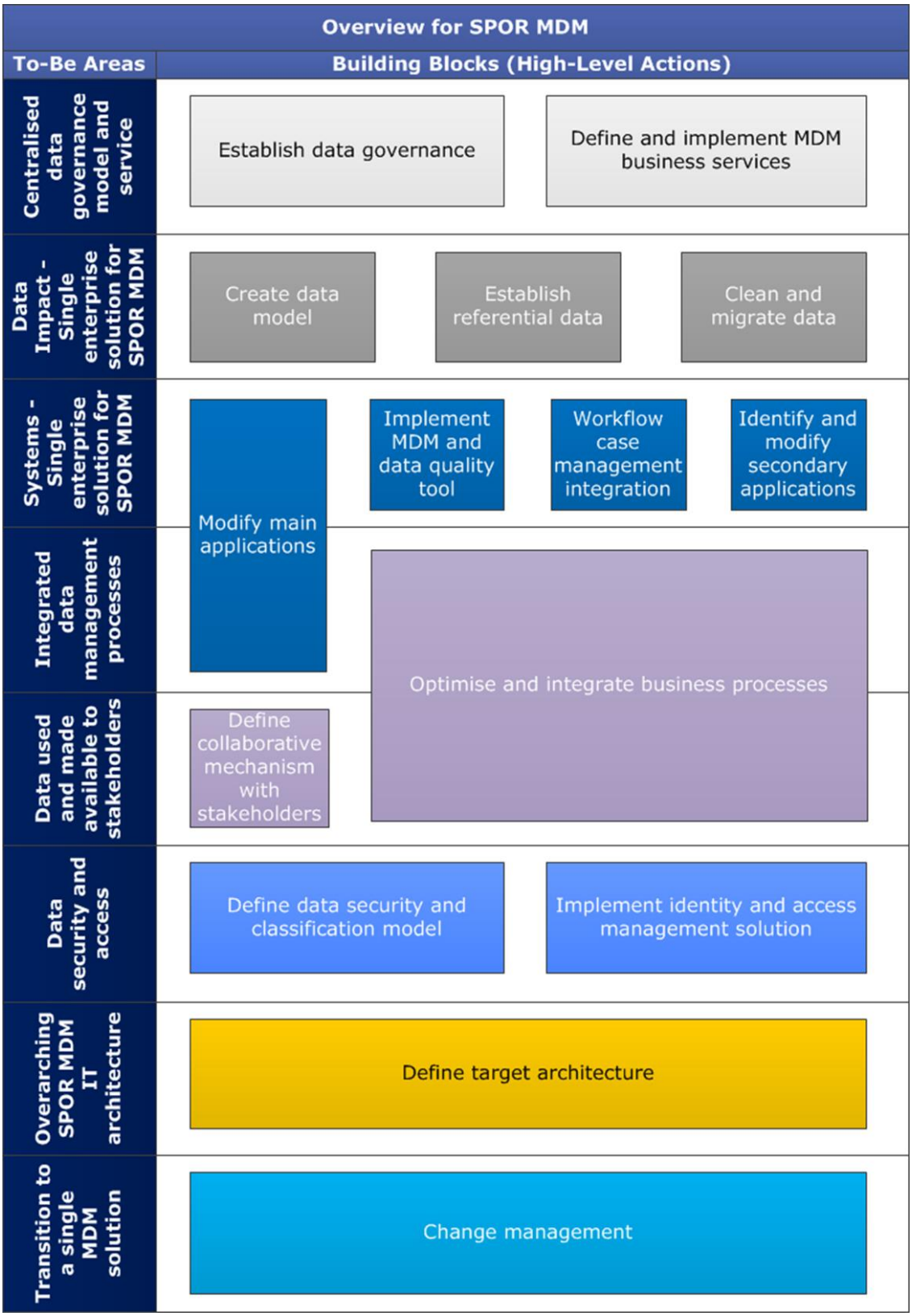
Outcomes

In total 47 high-level action actions were identified to bridge the gap between the current and desired future state.

The vast majority of actions cut across the core business processes and Substance, Product, Organisation and Referential data. Each action was mapped to the related business process and impacted systems, enabling building blocks (high level activities) to be made for the actions, subdivided by Substance, Product, Organisation and Referentials. *Fig. 7* represents these building blocks graphically for SPOR MDM. The building blocks shown here are common to Substance, Product, Organisation and Referentials – however there are some minor variations in building block composition, most notably for the Referential domain.

[2] Fédération Européenne d'Associations et d'Industries Pharmaceutiques: *'Principles for the Implementation of ISO IDMP Standards for EudraVigilance and Development of the Road Map', 6 October 2014*

Figure 7. Overview of the building blocks (summarising high-level actions) supporting the realisation of the desired future state for SPOR MDM



2.4. Prioritisation and optimisation of actions - approach

Prioritisation and optimisation were two complementary exercises:

Prioritisation:

- To define the implementation sequence of the SPOR master data related services
- To highlight which systems and business processes will be impacted as a result and timing of expected benefits.

Optimisation:

- To create the optimum allocation of resources and time to implement the catalogue of logical building blocks (high level activities) from the gap analysis according to their impact, complexity and inter-dependency.

2.5. Recommendation based on prioritisation and optimisation analyses

Prioritisation

During the prioritisation process, a number of criteria were used to determine the sequence reflecting a conservative approach to reduce risk, obtain business benefits sooner and address stakeholder concerns. The criteria for implementing the SPOR domain areas, taking EMA-level priorities and legislative requirements into account are as follows:

- Establishment of foundations for roadmap implementation
- Inherent risk of implementation
- Provision of quick wins early in the project cycle
- Promotion of continuous improvement and learning
- Ensuring consistent data classification
- Creation of referential data model as a repository to share ISO IDMP data with the EU and international regulatory network
- Number of systems involved in migration
- Expectations & priorities of stakeholders

Optimisation

This analysis assesses the direct business impact against relative complexity of the MDM implementation building blocks using the criteria considered when defining this lower level sequencing (see *Fig. 8* for an example analysis conducted for the Substance data domain). In determining scores for these two areas a number of factors were taken into account:

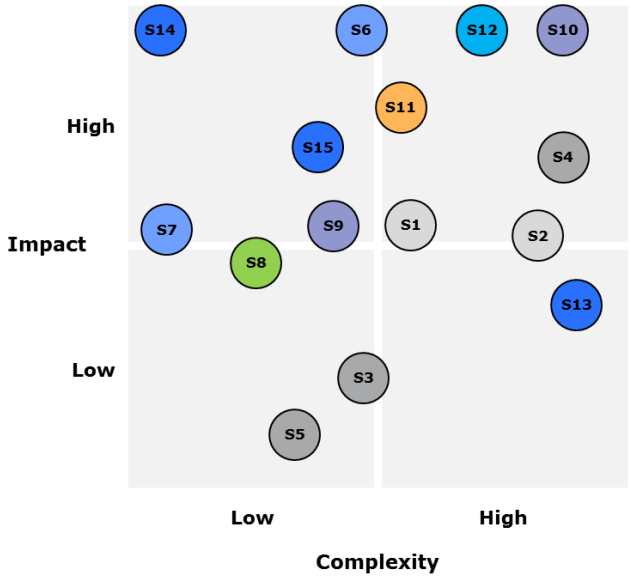
Business Impact (benefits)

- Value to stakeholders
- Direct business benefits in terms of efficiency and effectiveness

Complexity

- Number of applications/systems that must be adapted
- Number of processes affected
- Resources required
- Level of knowledge within EMA
- Technical challenge
- Number of dependencies

Figure 8. Optimisation matrix for building block implementation – Substance data domain example output



S Building Block	Building Block Description	To-Be Area
S1	Establish data governance	Centralised data governance model and service
S2	Define and implement MDM business services	Centralised data governance model and service
S3	Create S data model	Data Impact - Single enterprise solution for SPOR MDM
S4	Clean and migrate S data	Data Impact - Single enterprise solution for SPOR MDM
S5	Establish S referential data	Data Impact - Single enterprise solution for SPOR MDM
S6	Implement identity and access solution	Data security and access
S7	Define data security and classification model	Data security and access
S8	Define collaborative mechanism with industry, NCAs, FDA stakeholders	Data used and made available to stakeholders
S9	Workflow case management integration	Integrated data management processes
S10	Optimise and integrate S business processes	Integrated data management processes
S11	Define target architecture	Overarching SPOR MDM IT architecture
S12	Change management and communications	Support stakeholders through transition to a single MDM solution
S13	Modify main applications	Systems - Single enterprise solution for SPOR MDM
S14	Implement MDM and data quality tool	Systems - Single enterprise solution for SPOR MDM
S15	Identify and modify secondary applications	Systems - Single enterprise solution for SPOR MDM

3. Document history

Version	Date	Action	Authors
V0.1 draft	19/11/2014	Document created to present a proposal for the high level SPOR projects sequence approach to EUNDB and internally at EMA.	Kepa Amutxastegi, Pedro Pina Ferreira, Richard Fautley, Marek Lehman, Ray Power, Adam Stiling
V0.2 draft	13/03/2015	Draft roadmap consolidated for presentation and endorsement internally at EMA.	Kepa Amutxastegi, Richard Fautley, Marek Lehman, Ray Power, Adam Stiling
V1.0 draft	09/04/2015	Draft roadmap finalised for publication on the external EMA website.	Kepa Amutxastegi, Richard Fautley, Marek Lehman, Ray Power, Adam Stiling

4. Glossary

4.1. Key Terms

Term	Description
Article 57 / Art. 57	The submission of data on medicines by marketing-authorisation holders is a legal requirement from the 2010 pharmacovigilance legislation. Article 57(2) of Regulation (EU) No 1235/2010 requires: 1) EMA to publish the format for the electronic submission of information on medicinal Products for human use by 2 July 2011; 2) marketing-authorisation holders to submit information to EMA electronically on all medicinal Products for human use authorised in the European Union (EU) by 2 July 2012, using this format; 3) marketing-authorisation holders to inform EMA of any new or varied marketing authorisations granted in the EU as of 2 July 2012, using this format.
COTS	Commercial Off The Shelf – used within this document to refer to an available Product as opposed to development of a bespoke solution.
IDMP	Identification of Medicinal Products. A set of standards agreed within ISO comprising: • ISO prEN 11615, Data elements and structures for unique identification and exchange of regulated medicinal Product information • ISO prEN 11616, Data elements and structures for unique identification and exchange of regulated pharmaceutical Product information • ISO prEN 11238, Data elements and structures for unique identification and exchange of regulated information on Substances • ISO prEN 11239 Data elements and structures for unique identification and exchange of regulated information on pharmaceutical dose forms, units of presentation, routes of administration and packaging • ISO prEN 11240 Data elements and structures for unique identification and exchange of units of measurement
ISO	The short name for "International Organization for Standardization".
Master Data	Business data entities used across multiple systems, applications, and/or processes. Master data within the EMA is considered to include the primary attributes used to describe Substances, Products, Organisation and Referential data (SPOR) which are

Term	Description
	required to support multiple business processes.
Master Data Management (MDM)	The processes, governance, policies, standards and tools that consistently define and manage the master data providing a single, trusted point of reference.
Referentials/Reference Data	Lists of data used to ensure consistency in describing master data. Examples being country codes, routes of drug administration as well as more structured lists such as MedDRA.
SOA	Service-oriented architecture (SOA) is a software design and software architecture design pattern based on distinct pieces of software providing application functionality as services to other applications. This is known as service-orientation. It is independent of any vendor, product or technology.
SPOR	Substance, Product, Organisation, Referential data – considered to be "master data". Substance: the ingredients of a medicine Product: the medicinal Product itself Organisation: organisational data, e.g. Pharmaceutical companies, their addresses, their plants, distribution centres, the regulatory agencies, and persons related to these organisations. Referential: Reference lists such as dosage, forms, country codes, package codes, weight codes, etc.
Stakeholders	Pharmaceutical industry, EU Network partners and other external parties (e.g. FDA)

4.2. Systems

Term	Description
EudraCT	The EU's electronic database of clinical trials. It contains information submitted by sponsors and marketing-authorisation holders about clinical trials that started in the EU after 1 May 2004 and about clinical trials linked to the development of medicines for use in children.
EudraVigilance / EV	The EU's system monitoring the safety of medicines through safety reports. It is designed to receive, process, store and make available information submitted electronically. In addition this system contains the xEVMPD as used by MAHs to submit and maintain medicinal Product data (relating to human medicinal Products) as required by Article 57.
EUTCT	A central repository and publication system for controlled term lists used in the European medicines

Term	Description
	regulatory network.
SIAMED	The European Medicines Agency's Product information and application tracking system. It is used for managing: • Pre-submission activities • Applications within the centralised procedures • Fee calculation • Product data • MRLs data • Post-Authorisation Measures and PSURs
SAP	The IT system used by EMA to support finance and human resources activities.
xEVMDP	Extended EudraVigilance medicinal Product Dictionary – the database of Products and Substances that are either under development or approved for use within the EEA. This is the repository used to store data submitted in relation to Article 57.