EnCoRe Project Deliverable

Title: Technical Architecture arising from the third Case Study

Identifier: D2.3

Version: 1.0

Date: 18 November 2011

Status: Final

Authors: Marco Casassa Mont, Vaibhav Sharma, Siani Pearson, Rizwan Saeed, Martin Filz

Editor: Marco Casassa Mont

Reviewers: The entire EnCoRe project team

Class: Public

Summary

This document is a formal deliverable of the EnCoRe project and contains the definition of the EnCoRe Technical Architecture arising from the project’s third Case Study which is focused on the Cabinet Office/Identity Assurance Programme.
About the EnCoRe project
The EnCoRe project is an inter-disciplinary research project into informational privacy, undertaken collaboratively by UK industry and academia, and partially funded by the Technology Strategy Board (TP/12/NS/P0501A), the Engineering and Physical Sciences Research Council and the Economic and Social Research Council (EP/G002541/1).

Enabling individual citizens and consumers to retain control over their personal data can be achieved in a number of very different ways that are based on different trust models. When an individual discloses his or her personal data to commercial and other entities, he or she also grants consent, sometimes implicitly, for it to be used for one or more purposes. Subsequent control over the storage, use and onward sharing of that data relies on the notion of trust that the given consent will be respected.

Currently, many organisations’ personal data management operations do not justify the trust placed in them by the data subjects. This is for a variety of legal, regulatory, process-related, economic and technical reasons.

The EnCoRe project’s aims are:
• To enable organisations to adopt scalable, cost effective and robust consent and revocation methods for controlling the usage, storage, location and dissemination of personal data.
• To benefit individuals by providing meaningful, intuitive mechanisms which will allow them to control the use of their personal information held by others.
• To help restore individual confidence in participating in the digital economy and so, in turn, benefit the wider society.

The overall vision of the project is to make giving consent as reliable and easy as turning on a tap and revoking that consent as reliable and easy as turning it off again.

The project partners are:
• Hewlett–Packard Laboratories
• HW Communications Ltd
• London School of Economics and Political Science
• QinetiQ
• University of Oxford

The EnCoRe project runs from June 2008 to April 2012.

The project’s website is www.encore-project.info and it tweets at www.twitter.com/encore_project.
Executive Summary

This document is a formal deliverable of the EnCoRe project. It defines the EnCoRe Technical Architecture arising from the third Case Study: a scenario based on the Cabinet Office/Identity Assurance Programme. It also describes a set of related functional use cases. The requirements which guided the design of the architecture were gathered and defined by the legal and social science research groups within the EnCoRe project, through direct discussions with various WorkStreams in the Identity Assurance (IDA) Programme.

The scope of the EnCoRe Technical Architecture for the third Case Study encompasses all the technical functions required for the management of data subjects’ choices, the enforcement of individuals’ consents that are pertinent to the Case Study’s scenario and related monitoring and auditing capabilities.

Specifically this document aims to consolidate and finalise the specification of the EnCoRe Architecture by leveraging previous work done within the EnCoRe Architecture documents D2.1 and D2.2.

The technical architecture is the block-level design of the necessary technical system, at the level of functional blocks (i.e., software and service components) and the data flows between them and to/from humans, other technical systems, compliance and other business processes and regulatory environments. Its goal is to provide the basis for an EnCoRe reference implementation that validates the approach and the technology. To that end this document’s approach is to start with contextual information and overviews, and incrementally refine the level of detail and ground it to the scenario’s usage by presenting functional use cases.

Document History

<table>
<thead>
<tr>
<th>Doc ID</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1.0</td>
<td>Final Document</td>
<td>18 November 2011</td>
</tr>
</tbody>
</table>

Acknowledgements

The editor and authors wish to thank those who reviewed the drafts (Nick Moffat, David Lund, Bassem Ammar, Noel Catterall, George Mourikas, Edgar Whitley and Ioannis Agrafiotis) and also those who contributed with ideas and material to this deliverable. A special thank to HW Communications for suggesting the usage of a Remote Data Registry to support a key use case for Case Study 3.
Table of Contents

Executive Summary ................................................................. iii
Document History ................................................................. iii
Acknowledgements ............................................................... iii
Table of Contents ................................................................... iv
List of Figures ........................................................................ vi
Abbreviations ....................................................................... vii
1. Introduction ....................................................................... 1
2. Changes from the second Technical Architecture ............... 3
3. High Level Principles ........................................................ 5
4 High Level Overview of the Technical Architecture ............ 6
  4.1 Supported Interaction Flows and General Use Cases ........ 9
5 Refined Description of the Technical Architecture ............. 11
  5.1 Consent & Revocation (C&R) Privacy Assistant ............... 14
  5.2 Consent & Revocation Provisioning ................................. 14
  5.3 Data Registry Manager .................................................. 17
  5.4 Privacy-Aware Access Control Policy Enforcement ....... 19
    5.4.1 Policy Enforcement Point (PEP) .............................. 21
    5.4.2 Policy Decision Point (PDP) ................................. 21
  5.5 Event Manager ............................................................. 22
  5.6 Obligation Management System ..................................... 24
  5.7 Audit Logs Component ................................................. 28
  5.8 Compliance Checking System ....................................... 28
  5.9 External Workflow Manager ......................................... 30
  5.10 EnCoRe Configuration Database .................................... 33
6. Management of Privacy Preferences ................................. 35
  6.1 Technical Implications of Privacy Preference Management ... 36
  6.2 Implications for Data Subjects .......................................... 37
7. Specification of Privacy-Aware Access Control Policies and Obligations ............................... 40
  7.1 Privacy-aware Access Control Policies ......................... 40
  7.2 Obligation Policies ....................................................... 43
8. Sticky Policy Management ................................................... 47
9. Design and Deployment Options ........................................ 51
10. Use Cases ......................................................................... 52
  10.1 General Use Cases ..................................................... 52
    10.1.1 USE CASE 1: A data subject discloses personal data along with consent/privacy preferences ........................................ 52
    10.1.2 USE CASE 2: Employees and/or applications try to access data for specific purposes ........................................ 54
    10.1.3 USE CASE 3: Data subject changes their consent/privacy preferences ........................................ 55
    10.1.4 USE CASE 4: Personal data disclosed to a third party ........................................ 56
  10.2 Use Cases for the Third Case Study ............................ 59
    10.2.1 Setting the Context .................................................. 59
    10.2.2 Use Cases .............................................................. 63
      10.2.2.1 Use Case 1 ......................................................... 63
      10.2.2.2 Use Case 2 ......................................................... 64
List of Figures

Figure 1. EnCoRe Third Technical Architecture ................................................................. 6
Figure 2. Refined EnCoRe Architecture .............................................................................. 11
Figure 3. Privacy Consent & Revocation Assistant .............................................................. 14
Figure 4. Consent & revocation provisioning ....................................................................... 15
Figure 5. Data Registry Manager ....................................................................................... 17
Figure 6. Data Registry Database - Scheme ....................................................................... 18
Figure 7. Privacy-Aware Policy Enforcement ................................................................... 20
Figure 8. Interactions involving PEP, PDP and Data Registry Manager ......................... 22
Figure 9. Event Manager Component ............................................................................... 23
Figure 10. Refined Event Manager Component ................................................................. 24
Figure 11. Event Manager Storage - Schema .................................................................... 24
Figure 12. Model of the Obligation Management Framework ........................................ 25
Figure 13. Obligation Management System - Modules ...................................................... 26
Figure 14. Obligation Management System - Interactions ............................................... 27
Figure 15. Audit Logs ......................................................................................................... 28
Figure 16. Compliance Checking System ......................................................................... 30
Figure 17. External Workflow Manager ........................................................................... 31
Figure 18. Interactions between External Workflow Managers ....................................... 33
Figure 19. EnCoRe Configuration Database - Schema ..................................................... 34
Figure 20. Association of Privacy Preferences to Attributes in Personal Data .................. 35
Figure 21. Example 1: Gathering Privacy Preferences from Data Subjects (simple approach) ........................................................................................................................................... 38
Figure 22. Example 2: Gathering Privacy Preferences from Data Subjects (hybrid approach) ........................................................................................................................................... 39
Figure 23. Elements characterising EnCoRe Access Control Policies ............................... 41
Figure 24. Additional details - Elements characterising EnCoRe Access Control Policies ... 41
Figure 25. Elements characterising EnCoRe Obligation Policies ....................................... 45
Figure 26. Additional details - Elements characterising EnCoRe Obligation Policies........... 45
Figure 27. Sticky Policy Mechanism ................................................................................ 48
Figure 28. Sticky Policy Mechanism in multiple EnCoRe deployments ......................... 50
Figure 29. General USE CASE 1 ..................................................................................... 53
Figure 30. General USE CASE 2 ..................................................................................... 54
Figure 31. General USE CASE 3 ..................................................................................... 55
Figure 32. General USE CASE 4 – Enterprise A ............................................................... 57
Figure 33. General USE CASE 4 – Enterprise B ............................................................... 57
Figure 34. Simplified Architecture of Identity Assurance Ecosystem .............................. 60
Figure 35. Simplified Architecture of Identity Assurance Ecosystem .............................. 62
Figure 36. Consent Management for MIDS ..................................................................... 64
Abbreviations

The following abbreviations are used frequently in this document:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>Attribute Provider</td>
</tr>
<tr>
<td>C&amp;R</td>
<td>Consent(s) and Revocation(s)</td>
</tr>
<tr>
<td>DS</td>
<td>Data Subject</td>
</tr>
<tr>
<td>EWM</td>
<td>External Workflow Manager</td>
</tr>
<tr>
<td>GUI</td>
<td>Graphical User Interface</td>
</tr>
<tr>
<td>HCI</td>
<td>Human-Computer Interaction</td>
</tr>
<tr>
<td>ICO</td>
<td>Information Commissioner’s Office</td>
</tr>
<tr>
<td>IDA</td>
<td>Identity Assurance</td>
</tr>
<tr>
<td>IDP</td>
<td>Identity Provider</td>
</tr>
<tr>
<td>IWM</td>
<td>Internal Workflow Manager</td>
</tr>
<tr>
<td>LDAP</td>
<td>Lightweight Directory Access Protocol</td>
</tr>
<tr>
<td>OMS</td>
<td>Obligation Management System</td>
</tr>
<tr>
<td>PD</td>
<td>Personal Data</td>
</tr>
<tr>
<td>PDP</td>
<td>Policy Decision Point</td>
</tr>
<tr>
<td>PEP</td>
<td>Policy Enforcement Point</td>
</tr>
<tr>
<td>PKI</td>
<td>Public Key Infrastructure</td>
</tr>
<tr>
<td>PSP</td>
<td>Public/Private Service Provider</td>
</tr>
<tr>
<td>RDBMS</td>
<td>Relational Database Management System</td>
</tr>
<tr>
<td>XML</td>
<td>Extensible Markup Language</td>
</tr>
</tbody>
</table>
1. Introduction

The aim of this document is to present and discuss the third EnCoRe Technical Architecture. The first EnCoRe Technical Architecture [1] was designed to fulfil the requirements of the first EnCoRe case study, centred on employee data and focusing on an organisational context. The second EnCoRe Technical Architecture [2], based on a Biobank scenario, fulfilled additional requirements including: the need to support more flexible and compelling privacy-aware policies beyond access control such as obligation policies; the need to ensure that data subjects’ privacy preferences are taken into account and enforced when personal data is shared with third parties. This architecture was designed to support future needs such as the ones related to the third case study.

The third EnCoRe Technical Architecture primarily aims at refining and finalising previous specifications in the following areas: flexible expression of privacy preferences (choices); tracking of data whereabouts; privacy-aware access control policies and obligation policies; sticky policies; logging, auditing and compliance checking. These refinements are driven by additional knowledge and requirements gathered in EnCoRe, during the second and third case studies.

This document provides a description of the updated architecture. The reader is directed to [1] and [2] for context, background and some detailed information.

EnCoRe now explicitly defines the concepts of “Privacy Options” and “Privacy Choices” to respectively refer to: the options that data subjects (users) have in terms of how their personal data should be handled and processed; the actual choices made by the data subject (based on the available options). The EnCoRe Technical Architecture explicitly supports the management of privacy options and privacy choices. For back compatibility to previous architectural documents [1,2] this document still uses the generic “privacy preferences” term: depending on the context, it clarifies the meaning, i.e. if the focus is on privacy options or choices.

Although inspired by, and focused on, the specifics of the third EnCoRe case study, this architecture is much more widely applicable than to just that scenario, being suitable for use in other scenarios where an individual (the data subject) discloses his or her personal data to an organisation, which may wish to disclose it to other organisations. Its legal ability to do so may depend on the specific details of the consent, granted by the data subject at the time of disclosure. At that time, the data subject may not be fully aware of the implications of granting consent, and/or may select the simplest consent options offered by the organisation. Later, perhaps after becoming more aware of these implications, or having just changed her mind, the data subject may wish to revoke the previously granted consents and be sure that her new wishes will be respected by all the organisations that have (or have access to) copies of the personal data she disclosed. In order for this to happen, a complex set of interactions, between and within the involved organisations, is required. This architecture provides the framework for these.
For consistency with the first and second EnCoRe Technical Architectures, the organisations considered here are assumed to be non-state organisations, and hence this document uses the terms “organisation” and “enterprise” to refer to these equivalently. However, the points made in this document are also equally applicable to state organisations.

Finally, it is beyond the scope of this document to provide any implementation and deployment details. However, in order to ground some of the key architectural concepts, a few technical details have been added for illustrative purposes only.
2. Changes from the second Technical Architecture

The major changes from the second technical architecture include, but are not limited to:

- **Enhanced Audit Logs**: this becomes a central component to collect audit log information from the other EnCoRe components;

- **Consolidated Compliance Checking Component**: this component now includes a consolidated Compliance Checking component, based on the component defined in Architecture D2.2. It supports Risk Management activities; it leverages the information contained in the Audit Logs component;

- **Enhanced Event Manager**: this component has been fully specified to enable the collection and distribution of events among EnCoRe components, in particular to support the Obligation Management System;

- **EnCoRe System Configuration DB**: this component has been introduced to explicitly store EnCoRe configuration information. It includes: configuration of EnCoRe components; schemas defining how various types of privacy preferences are associated to personal data; mapping of data types to stored personal data (potentially in legacy systems) within and across organisations;

- **Removal of Semantic Configuration Component**: no case study requirements have been provided about this component. It has been replaced with the EnCoRe System Configuration DB which provides a pragmatic, simpler and more general approach to the configuration of the EnCoRe system and related components;

- **Enhanced Data Registry Manager**: this component has been fully specified to describe how to store: data subjects’ privacy preferences; links to associated personal data; data whereabouts;

- **Enhanced Privacy-Aware Access Control Policy Enforcement**: this component (known as Privacy-Aware Policy Enforcement in the second Architecture) has been fully specified along with related privacy-aware access control policies. A reference implementation is provided;

- **Enhanced Obligation Management System**: this component has been fully specified along with related privacy obligation policies. A reference implementation is provided;

- **Enhanced Sticky Policies**: this capability has been fully specified along with a reference protocol to enable the exchange, enforcement and monitoring of sticky policies, within and across organisations;

- **Trust Authority**: this component has been introduced to support the management of sticky policies and related auditing capabilities.

- **Removal of the Negotiation Manager**: this component’s initial role was to support degrees of negotiation between data subjects and organisations. However no specifications and requirements have been provided by the third
case study about how this could effectively be supported by organisations. Its core capabilities are now provided by existing EnCoRe modules, which enable data subjects to make their privacy choices and customise them – based on templates supported by organisations.
3. High Level Principles

The principles that guide the design of this architecture are the following:

- Provide individuals (data subjects) with graphical interfaces which support various levels of details in terms of privacy options. This might range from setting preferences by using graphical sliders (which show predefined privacy levels) to setting fine-grained privacy preferences for each of their personal data items (or groups of data items). The same approach must be provided when individuals decide to change their preferences;
- Provide formal specification of fine-grained privacy preferences (choices) that might be driven by graphical interfaces or by the data subjects setting specific values;
- The expression of (or changes of) fine-grained preferences directly influences how data subjects’ information is accessed, processed and disclosed, within and across organisations (if data has been shared with third parties). Specifically, these preferences have an impact on how privacy-aware access control and obligation policies are enforced and the definition of sticky policies;
- Organisations must have a systemic approach to managing the association of privacy preferences to personal data, within their back-end systems that takes into account the complexity of their environments and potential legacy systems;
- Organisations must also have a systemic approach to tracking the whereabouts of data subjects’ information, within and across their boundaries;
- Organisations must be able to log and monitor the various activities involving the access, management, processing and disclosure of personal data, happening via EnCoRe solutions. This is meant to support the required privacy compliance checking capabilities.
4 High Level Overview of the Technical Architecture

The diagram below provides a high-level overview of the third EnCoRe Technical Architecture.

![Figure 1. EnCoRe Third Technical Architecture](image)

The third EnCoRe Technical Architecture is the final EnCoRe Technical Architecture. It refines the previous two architecture specifications [1,2]. It is heavily based on the second EnCoRe Architecture as it was designed to support general purpose capabilities for consent and privacy management.

This architecture factors in requirements and knowledge gathered from the three EnCoRe case studies. The core components are:

- **Consent & Revocation Privacy Assistant Plug-in**: this is a client-side component. It is in charge of supporting data subjects’ (users’) definition of privacy preferences (for specific sets of data items) and the local storage of historical choices and system decisions. It can be implemented in a variety of ways, including: a web browser plug-in; a set of scripts provided within web pages;

- **EnCoRe System Configuration DB**: this component stores, in a centralised way, the configurations of the EnCoRe components. This includes: the schema, defined by administrators, describing how various types of privacy preferences are associated to personal data (the way this association will be presented to the data subjects); the mapping of internal representations of types of personal data (e.g. names of data items within legacy databases,
LDAP, etc.) to higher-level definitions used by EnCoRe (e.g. names of data as presented to data subjects. NOTE: this is also used to provide a common vocabulary in the case that EnCoRe systems are deployed across organisations); the initialization parameters and various shared configuration parameters of EnCoRe components;

- **Consent & Revocation Provisioning**: this component is the contact point between the organisation’s web server/portal and the EnCoRe framework (i.e. the EnCoRe components). It primarily provides workflow-based coordination and provisioning capabilities to the following cases: data subjects disclose their personal data along with privacy preferences; data subjects change or revoke some of their privacy preferences. The core engine of this component is the *Internal Workflow Manager*;

- **Internal Workflow Manager**: this sub-component of the Consent & Revocation Provisioning component orchestrates the sequence of privacy management tasks needed to be executed on personal data and preferences within the organisation. This includes the initial disclosure of personal data and preferences, along with any subsequent change. This component interacts with various EnCoRe components, including: the internal personal data storage(s); the Data Registry Manager; the Privacy-aware Policy Enforcement component; the Obligation Management System. It also interacts with the External Workflow Manager component when personal data (and related preferences) are transmitted across the external boundaries of the enterprise;

- **Data Registry Manager**: this component is in charge of storing data subjects’ privacy preferences, along with associations to the related personal data. This component also keeps tracks of the whereabouts of personal data, within and across organisations;

- **Data Viewer and Manager**: this component is in charge of supporting data subject’s requests about which types of personal data the organisation stores about them and the whereabouts of this data (i.e. where data has been disclosed to). The required information is retrieved from the Data Registry Manager;

- **Internal Workflow Agent**: this component is a trusted interception point, usually deployed within Applications and Services (e.g. via a software wrapper) or on top of data stores (e.g. an SQL interceptor for RDBMS databases). It is meant to interact with the Privacy-Aware Access Control Policy Enforcement component, to ensure that access to personal data is granted according to policies and data subjects’ preferences;

- **Privacy-Aware Access Control Policy Enforcement**: this component is in charge of enforcing security & privacy access control policies on personal data, driven by data subjects’ preferences. These policies takes into account of preferences such as purposes for accessing data, entities the data may/may not
be disclosed to, etc. Access Control Policy templates have been designed to be suitable for direct implementation;

- **Privacy-Aware Access Control Policy DB**: this is the database storing the definitions of access control policies. These definitions can be modified over time by authorised administrators;

- **Obligation Management System**: this component is in charge of enforcing and monitoring privacy obligation policies, driven by data subjects’ preferences. It is complementary to the Policy Enforcement component as it enforces privacy preferences related to notifications, deletion or minimisation of data, etc. Obligation Policy templates have been designed to be suitable for direct implementation. The Obligation Management System receives relevant events from the Event Manager (e.g. about access and disclosure events, time events, etc.) and uses them to trigger the enforcement of obligation policies;

- **Obligation Policy DB**: this is the database storing the definitions of obligation policies. These policies can be modified over time by authorised administrators;

- **Event Manager**: this is the component in charge of collecting events relevant to EnCoRe and broadcasting them to various EnCoRe components. It supports a producer/consumer model: components need to register as producers, to share events with the Event Manager or register as consumers, to be notified of specific events. In the current specification, the Privacy-Aware Access Control Policy Enforcement component is a producer; the Obligation Management System and the Audit Logs components are consumers;

- **Audit Logs**: this component is in charge of getting and storing log information provided by various EnCoRe components. Each EnCoRe component is instrumented with a configurable agent to provide logging data. The Audit Logs component provides key information to the Compliance Checking component;

- **Compliance Checking**: this component is in charge of providing monitoring, auditing and compliance checking capabilities. It is driven by compliance criteria and related policies (e.g. about how data should be accessed, used, disclosed based on organisational policies and data subjects’ preferences); it monitors for compliance to these policies by leveraging (among other things) the information provided by the Audit Logs component. The Compliance Checking Component supports the organisation in their governance and risk management activities, related to data handling, data sharing and privacy management;

- **External Workflow Manager**: this component manages the interactions and collaborations with third parties. This component is in charge of communicating notifications of updates/changes (about personal data and privacy preferences) to third parties to enable the remote enforcement of
privacy choices and related access control and obligation policies. This component deals, *inter alia*, with the management of sticky policies, e.g., the binding of the personal data with cryptographic mechanisms to ensure that data subject choices are fully respected even when their personal data goes across the boundaries of the enterprise.

- **Trust Authority**: this component is in charge of dealing with: the checking of fulfilment of sticky policies; the release of cryptography keys, subject to the fulfilment of policy constraints; audit logging and forensic analysis.

Section 5 provides a detailed description of the various architectural components.

### 4.1 Supported Interaction Flows and General Use Cases

The main interaction flows involved in this architecture are consistent with those defined in [1, 2]. However, they have been refined to reflect the latest architectural design and capabilities.

The up-to-date, key supported interaction flows can be described by means of 4 general use cases, as follows:

- **An end-user discloses personal data along with consent/privacy preferences**: the system captures these preferences via user-side (data subject’s) plug-ins; the information is sent to the back-end C&R provisioning component for internal configuration (of policies and the data registry). This includes setting privacy obligations in the privacy obligation manager, driven by user preferences;

- **Employees and/or applications try to access data for specific purposes** (e.g. marketing, transaction processing, research, etc.): the privacy-aware access control policy enforcement component intercepts these requests (e.g. via SQL interception and/or specific interception points within applications) and grants or denies access based on the evaluation of access control policies. These policies not only describe security constraints (who can access what) but also privacy constraints based on data subjects’ preferences (e.g. approved/banned purposes for using data, black/white lists of entities that can handle data, etc.);

- **Data subject changes their consent/privacy preferences**: the data subject can change their privacy preferences. This triggers a chain of updates of stored privacy preferences within the organisation, including updates of the data registry, access control policies and obligations. If the updated preferences relate to data shared with third parties, these parties will also receive the updates;

- **Personal data is disclosed to a third party**: the system intercepts the attempt of applications to disclose data to third parties (via locally deployed agents). If
the transfer of data is authorized by the access control component, then the personal data is disclosed to the third party via the external workflow manager, by using the sticky policy mechanisms which bundle data to policies and privacy preferences. The level of stickiness (simple association or strong cryptographic binding) can be configured. The data registry is updated accordingly about with where the data has been disclosed.

Detailed descriptions of these General Use Cases and the specific Use Cases for the third Case Study are provided in Section 10.
5 Refined Description of the Technical Architecture

The following figure shows a refined representation of the third EnCoRe technical architecture, compared to the level overview of the architecture shown in Figure 1. The core principles and underlying mechanisms are the same as for the second EnCoRe technical architecture, but with the differences mentioned in the previous section.

![Refined EnCoRe Architecture](image)

**Figure 2. Refined EnCoRe Architecture**

The following list of assumptions, made for the first and second EnCoRe Technical Architectures applies to the third architecture also:

1. The EnCoRe Technical Architecture must be flexible, extensible and general purpose. It must be widely re-usable for (and beyond) EnCoRe Case Study 3.

2. All components of the architecture are secured following good security practice and can be used as standalone service components;

3. The data subject’s system (e.g., web browser) that accesses EnCoRe is secured against cross-site scripting or cross-site request forgery attacks;

4. The platforms and systems running the back-end EnCoRe components are secured and observe good security practices;

5. An EnCoRe-compliant system shall conform to the EnCoRe architecture and high-level EnCoRe principles;
6. Proper operational security mechanisms are in place to provide secure communication channels between EnCoRe components and for communications with third parties;

7. Data security mechanisms are in place to ensure different degrees of binding of policies to data. This will be dictated by business, security and privacy needs. It might involve the use of cryptographic solutions for confidentiality and non-repudiation purposes;

8. A data subject can express fine grained consent and revocation preferences dictating obligations and duties to the organisations;

9. Data subjects will manage their data consent and preferences only by interacting with the system to which they originally disclosed their information;

10. The data subject’s preferences (and a trace of consent/revocation changes) will be stored locally within his/her web browser (using an HTML5-enabled browser) and/or in any trusted location (e.g., network file system, service in the cloud, etc.) chosen by him/herself;

11. A data subject will be given privacy advice and support by their local Consent and Revocation Privacy Assistant component regarding potential violations (by organisations) of agreed privacy preferences and also historical information about consent and revocation choices. Specifically, when the data subject interacts with the organisation (e.g. to visualise or update their personal data in a web registration form), the Privacy Assistant checks for the state of personal data against expressed privacy choices. For example, the data subject might have requested that a specific data item should be deleted at a specific point in time. If the data item is still there after the deadline, the Privacy Assistant will highlight the violation;

12. System-wide events can be used for compliance checking and to support risk management;

13. The interaction is limited to an EnCoRe-compliant system or a legacy system which uses an EnCoRe-compliant external workflow manager.

The additional assumptions introduced in the third EnCoRe Architecture are:

1. Each EnCoRe component supports configurable and secure logging capabilities to track various activities of the component. The locally retrieved log information has to be shared with the centralised EnCoRe
Audit Logs component to further support compliance checking and risk management;

2. An EnCoRe-compliant system needs to support interactions with at least a Trust Authority, in order to leverage the sticky policy capabilities. The Trust Authority capabilities can be provided by a trusted third party or by the EnCoRe system;

3. The EnCoRe system needs to be configured, via the EnCoRe Configuration database, by authorised and competent personnel and administrators.

4. The EnCoRe access control and obligation policies need to be defined, configured and deployed in the EnCoRe system by authorised and competent personnel/administrators. The data subject might be notified about these changes if this was a privacy choices expressed by the data subject (and supported by the organisation).

**Architectural Security Principles**

Consistently with what was described in the Second Technical Architecture [2], EnCoRe makes a distinction between two different aspects of security:

- Operational Security
- Data Security

*Operational security* ensures that: secure communication channels (e.g., SSL/TSL) are present to secure communications between internal components of an EnCoRe system and across the involved parties. This is necessary for the transmission of data and policies. Confidentiality and non-repudiation techniques (e.g., encryption and digital signature) need to be in place when transmitting information across organisations’ external boundaries. It also ensures that security best practices are implemented, including proper setting of access controls to the hosting systems, patching and updating of the underlying software solutions, etc.

*Data Security* primarily builds on top of operational security and relies on the additional sticky policies mechanism to ensure that data do not leave the system without their associated policies. While the term sticky policies might seem to imply cryptographic binding, this is not always so. As discussed, a weak sticky policy can simply be a logical binding of policy, data and metadata, (the glue is a simple logical binding).

Refer to [1,2] for further information about this.
The remaining part of this section provides the details about the refined components of the architecture and related concepts.

5.1 Consent & Revocation (C&R) Privacy Assistant

Figure 3 shows that the C&R Privacy Assistant is a client-side component (e.g. a web server plug-in) that supports data subjects in defining and storing their privacy preferences. It keeps a local copy of the choices and preferences expressed by the data subject, for different EnCoRe-enabled services, in order to identify discrepancies that may occur between expected behaviours and actual behaviours. When the data subject interacts with the organisation (e.g. to visualise or update their personal data in a web registration form), the Privacy Assistant checks for the state of personal data against expressed privacy choices. For example, the data subject might have requested that a specific data item should be deleted after a predefined time period. If the data item is still there after the deadline, the Privacy Assistant will highlight the violation.

Moreover, depending on the implementation it could be used in the context of providing synchronous and asynchronous notifications to data subjects. The component can store data subjects’ choices and preferences on a variety of different media, including: the local file system, external storage within the enterprise (such as network folders) and a cloud service provider. The latter is designed to handle the case of using multiple client devices (e.g., laptop, home PC and smartphone) and therefore it is necessary to keep synchronized all the client’s history by leveraging a cloud synchronisation service.

Figure 4 shows the internal modules of the Consent & Revocation Provisioning component.
EnCoRe: Ensuring Consent and Revocation

This component interfaces data subjects’ interactions with the back-end EnCoRe system components by means of web pages or web forms provided by the organisation. All its actions and activities are logged. Log information is sent securely to the EnCoRe Audit Logs component.

It consists of four parts:

- **Internal Workflow Manager (IWM)**
  This module receives all the incoming requests from outside the system and routes them to the relevant components. This includes:

  - A new data subject’s registration request coming from the website is delegated to the “C&R Provisioning Daemon” along with the data subject’s personal data and preferences attached to it. Once the processing is completed it receives a “success or failure” response from the “C&R Provisioning Daemon”. It either stores user data to the Enterprise Data Repository or rejects the request based on the response received and sends an appropriate response back to the website to be displayed to the data subject.

  - After receiving a data access request from an existing data subject’s web browser (e.g. a returning data subject wanting to check/change their personal data and/or preferences), the IWM sends a request with the requestor’s identification to the Data Registry Manager to get the privacy preferences. It also sends an access request to the Enterprise Data Repositories via the local Policy Enforcement Point (PEP). *It is important to notice that the PEP is a distributed component. The IWM component (as for other EnCoRe components) uses a local PEP.* In case of successful authorization, the relevant personal data and privacy preferences are returned to the data subject via web pages/web forms.

  - A request to update personal data and/or privacy preferences by an existing data subject (via their web browser) is delegated to the “C&R Provisioning Daemon”. This request includes the updated personal data and/or preferences. Once all the processing is completed by the relevant EnCoRe components, it receives a response as a “success or failure” back from the “C&R Provisioning Daemon”. Based on the received response, it...
either stores the data subjects’ data in the Enterprise Data Repository or rejects the request and sends an appropriate response back to the website to be displayed to the data subject.

- In case an employee/administrator sends a request, via the web portal, to retrieve data of a specific data subject, the IWM uses its PEP capabilities to obtain the authorization from the Privacy-Aware Access Control Policy Enforcement component; if authorization is granted, it queries the Enterprise Data Repository and returns the response to the requestor.

- Whenever an application/service sends a request to the IWM to share personal data (about a data subject) and related privacy preferences with an external organisation, the IWM uses its PEP capabilities to obtain the authorization from the Privacy-Aware Access Control Policy Enforcement component; if authorization is granted, it passes the request (along with the involved information) to the External Workflow Manager, for further processing.

- In case the External Workflow manager sends a request to the IWM to store personal data (about a data subject) and related privacy preferences coming from an external organisation, the IWM proceeds in a similar way to the case of “a new data subject’s registration request”, by sharing the additional contextual information about the provenance of the data (i.e. external to the organisation).

- **C&R Provisioning Daemon**

  This component receives personal data and data subject’s preferences from the IWM in the cases of “new data subject registration” and “existing data subject update”. It sends a call to the C&R “Verification Manager” to check for the correctness and consistency of the involved personal data and preferences (including whether they are based on templates defined by the organisation). On receiving a positive answer from the “Verification Manager”, it sends the data subject preferences to the Data Registry Manager. Once confirmation is received from the Data Registry Manager, it provides the response (either reporting success or failure) to the IWM.

- **C&R Verification**

  This is the module that validates the correctness and consistency of the format and values of the personal data and data subject’s preferences contained in a “new user registration” request or “existing user update” request submitted from the front end.

- **C&R Registration Process**

  This module is the interface with the Data Registry Manager. Whenever any new data subject’s preferences or existing data subject’s update is received, this module sends the relevant preferences to the Data Registry Manager and receives the response.
5.3 Data Registry Manager

The details of the Data Registry Manager are shown in the figure below.

The Data Registry Manager is the component in charge of managing a repository of the following information:

- data subjects’ preferences (determined by their consent and revocation choices);
- associations of privacy preferences to the actual personal data, stored in organisations’ data repositories;
- information about the whereabouts of personal data, i.e. where the personal data has been disclosed to;
- additional information about personal data, such as provenance.

In this document we also refer to the above information as metadata.

The Data Registry Manager consists of three key modules:

- Data Registry Daemon: it is the module that listens to incoming requests and orchestrates the interactions with the Data Registry Mapping Process;
- Data Registry Mapping process: it is the module in charge of the processes of storing and retrieving information/metadata from the (Virtual) Data Registry Database;
- (Virtual) Data Registry Database: it stores the metadata/information mentioned above. It is important to note that this database does not store any personal data but only references to their storage locations, within Enterprise Data Repositories.

The storage of metadata/information within the Data Registry Database is critical for the correct functioning of the EnCoRe system. It is beyond the scope of this document to describe the implementation details. However, the following figure shows, as an
example, a reference scheme of the Data Registry Database, for illustration purposes only:

![Data Registry Database - Scheme](image)

This scheme reflects a potential flexible and configurable approach to associate privacy preferences to personal data. More details are provided in Section 6.

In this example, a few tables have been introduced to provide the required storage capabilities of the metadata associated to personal data:

- **User**
  
  This table stores references to data subject’s identities, as defined within the organisation

- **DataAttributes**
  
  This table stores the type/name of each Data Attribute but not its value (which defines the type of data subjects’ personal data. It contains the mapping between the logical name of the attribute, used in EnCoRe policies and systems, and the actual name used in the Enterprise Data Repositories. As discussed in Section 6, it also provides links to the metadata, i.e. data subjects’ privacy preferences, external disclosure locations and data provenance

- **SinglePreferences**
This table stores the values of data subjects’ privacy preferences that are defined by a single value, such as deletion dates, yes/no preferences, etc.

- **ListPreferences**

  This table stores the value of data subjects’ privacy preferences that are defined by a list of values, such as the list of purposes for using personal data, the list of allowed/disallowed sites with which data can/cannot be shared with, a list of notification methods to be used, etc.

- **InternalDBLocation**

  This table stores information about the actual Enterprise Data Repositories used to store the data subject’s personal data. It also stores the actual physical location of each attribute in an Enterprise Data Repository e.g. a database or LDAP directory.

- **ExternalDataLocation**

  This table stores the whereabouts of each item of personal information i.e. the list of locations of direct third parties with which data has been disclosed.

- **DataProvenance**

  This table stores information about the provenance of personal data i.e. either if it comes directly from the data subject or from a third party.

When a new data subject registers with the organisation, the C&R provisioning component sends the metadata (described above) about their personal information to the Data Registry Daemon module. This module interacts with the Data Registry Mapping Process which implements algorithms to manipulate the metadata and store it in its database.

The Data Registry Manager provides the stored information to EnCoRe components on a need-to-know basis, after authorization is granted from the Privacy-Aware Access Control Policy Enforcement Component. Its activities are logged locally. Log information is sent securely to the Audit Logs component.

### 5. 4 Privacy-Aware Access Control Policy Enforcement

The Privacy-Aware Access Control Policy Enforcement component is a core EnCoRe component, shown in the figure below:
This component is in charge of ensuring that personal data stored in a variety of Enterprise Data Repositories (e.g. RDBMS databases, LDAP directories, etc.) is accessed based on the fulfilment of privacy-aware access control policies and data subjects’ privacy preferences.

It primarily consists of the following components:

- Policy Enforcement Point (PEP): this is a trusted and secure gatekeeper which allows/disallows access to personal data based on decisions made by the Policy Decision Point (PDP);
- Policy Decision Point (PDP): this is the component that, based on access requests and contextual information, grants/denies access to data based on the evaluation of access control policies and relevant data subjects’ privacy preferences;
- Contextual Handler: contextual information is related to an access request, such as the purpose for accessing the data, the requestors’ role and Identification, location, etc.

It is important to note that the PEP component is a distributed component. Multiple local instances of the PEP are deployed within the EnCoRe components and organisations’ applications/services, to locally enforce access control decisions – as shown in the EnCoRe detailed Architecture. These components securely interact with the centralised PDP component.

The PEP/PDP architecture adopted in EnCoRe is consistent and compliant with current access control standards, such as the ones mandated by IETF [3] and OASIS/XACML [4].

A detailed description of the involved privacy-aware access control policies is provided in section 7.

More details follow about the PEP and PDP components and their overall capabilities.
5. 4.1 Policy Enforcement Point (PEP)

The PEP is a trusted and secure agent or an interception point to ensure that access to personal data is consistent with stated access control policies and data subjects’ preferences.

The PEP acts as a gatekeeper. It handles data access requests coming from a variety of requestors: people (employees), applications, services, EnCoRe components, etc.

The access request usually contains information about:

- The requestor's identity and/or role;
- The identity of the data subject for which an access to their personal data has been requested;
- The list of Attributes (related to the data subjects’ personal data) that the requestor wants to access (e.g. credit card, address, etc.). Potentially this could include the entire set of Attributes associated to a data subject. These Attributes are identified by using the EnCoRe logical names for attributes and automatically mapped (by the Data Registry Manager) to the actual physical attributes stored within the Enterprise Data Repositories;
- The purpose of the request.

As soon as the PEP receives an access request for data pertaining to a particular data subjects (identified by a suitable organisation’s internal UserID) it retrieves the data subjects’ preferences from the Data Registry Database.

It then sends the request along with all the contextual information (i.e. data attributes requested, access purpose, privacy preferences, etc.) to the PDP and waits for the PDP allow/deny decision (which is based on the evaluation of the relevant access control policies and the specific data subjects’ privacy preferences).

Based on the PDP decision response, the PEP either queries the Enterprise Data Repository and returns the response to the requesting component (PDP decision: allow) or rejects the request (PDP decision: deny).

5. 4.2 Policy Decision Point (PDP)

The PDP is the core EnCoRe component that makes access control decisions. These decisions are based on the defined access control policies and contextual information, including the purpose of the requested data access and data subjects’ privacy preferences.

It implements a Policy Evaluation Engine that specifically supports enhanced, privacy-aware access control policies. These policies enable the specification of security and privacy constraints on how personal data may be accessed. A full specification is provided in Section 7.
At the time of initialization, the PDP component reads the policy definitions stored in the Access Control Policy database (or repository).

EnCoRe recommends representing access control policies in a human readable language, such as XML [5].

The PDP processes these policy representations and stores them within an internal Policy Repository, to enable their efficient evaluation by the Policy Enforcement Engine.

Once the PDP receives an access request from the PEP (inclusive of all the additional contextual information) it accesses the internal Policy Repository and evaluates them, based on a priority list (e.g. depending on the sequence in which they were defined).

During the policy evaluation process, if any policy evaluation returns a positive result (i.e. allow access) the evaluation process stops and a positive response is sent back to the PEP. Otherwise, if none of the policies provides a positive response, an access denied response is sent back to the PEP.

The following picture summarises the key interactions between the PEP, PDP and Data Registry Manager components, along with the exchanged information.

![Figure 8. Interactions involving PEP, PDP and Data Registry Manager](image)

### 5. 5 Event Manager

The Event Manager, shown in figure 9, is the component in charge of managing events that are relevant for the correct functioning of other EnCoRe components, in particular the Obligation Management System.
The Event Manager uses a producer/consumer model:

- Event Producers need to register with the Event Manager by declaring who they are and the types of events that they are going to produce. An example of producer is the Privacy-aware Access Control Policy Enforcement;
- Event Consumers need to register with the Event Manager by declaring who they are and which types of events they wish to receive. An example consumer is the Obligation Management System.

After the registration process, involving relevant EnCoRe component, producers send their events to the Event Manager. The Event Manager then broadcast them to the consumers that are interested in those types of events.

The Event Manager does not store the events locally. It logs various related activities and securely shares the log files with the Audit Logs component.

The following list of key events must be supported by the Event Manager:

- Time events (e.g. hourly-based time events);
- Successful Access Control events: successful access request along with contextual information (who is trying to access what, and for which purpose)
- Unsuccessful Access Control events: unsuccessful access request along with contextual information (who is trying to access what, and for which purpose)

The following figure provides a refined description of the modules involved in the Event Manager:
The Event Manager component uses a local database to store information about the involved consumers and producers and the types of supported events.

The following figure shows a possible instance of the schema of this database, provided here for illustration purposes:

5. 6 Obligation Management System

The Obligation Management System (OMS) is in charge of implementing privacy-aware lifecycle management of data.

As detailed in the Second EnCoRe Technical Architecture [2], the OMS is in charge of enforcing constraints and duties that have been defined both by organisational policies and data subjects’ preferences. This includes obligations about:

- notifying data subjects about usage or disclosure of their personal data;
- dealing with transformation and minimisation of personal data;
- dealing with data deletion, etc.
The architecture of the OMS component (and underlying principles) leverages the capabilities and functionalities developed by HP Labs in previous work [6,7]. Figure 4 illustrates the conceptual obligation management framework described in the Second EnCoRe Technical Architecture [2]:

![Diagram of Obligation Management Framework](image)

**Figure 12. Model of the Obligation Management Framework**

Parametric obligation policies are defined up-front defining the various kinds of constraints and duties that the organisation can enforce on the stored data. These obligation policies are parameterised by data subjects’ preferences, in the sense that these preferences instantiate the behaviour of the obligation policies (e.g. parameterised by the preferred time by when data will be deleted, situations when notifications should be sent, etc.).

EnCoRe recommends that these obligation policies are represented in a human readable way, for example in XML [5].

A detailed description of these obligation policies is provided in Section 7.

The triggering of obligation policies happens via an event-driven mechanism. For example, obligation policies might be activated by time events or access events. The OMS is a consumer of events shared by the Event Manager.

The obligation policies handled by this system do not necessarily only depend on access control events, but could also depend on time events, business process events (e.g., disclosure of data), security events (e.g., detected attacks), etc. This is required to take into account and enforce a wide range of situations and requirements, driven both by organisations and by the needs and preferences of data subjects. As such, the
OMS framework and system used in EnCoRe do not suffer from the limitations of existing approaches and frameworks, such as the XACML framework [4].

The following figure shows the internal modules of the OMS component:

![Obligation Management System - Modules](image)

**Figure 13. Obligation Management System - Modules**

The OMS includes the following core modules:

- **Obligation Scheduler**: this component keeps track of the instantiated obligation policies and the relevant preferences and constraints. It schedules the execution/enforcement of obligations, based on their triggering conditions, i.e., based on occurrence of specified events. In case an obligation is triggered and needs to be enforced, the Obligation Enforcer component will be activated;

- **Obligation Enforcer**: this component is in charge of enforcing triggered obligations, i.e., enforcing the set of actions specified in the obligation policy, consistently with the privacy preferences expressed by a data subject. This might include dealing with notification actions, data transformation or deletion actions, etc. The system tracks the status of these actions and their success in being enforced;

- **Obligation Monitoring**: this component is in charge of monitoring the enforced obligations, i.e., ensuring that the involved actions are fulfilled effectively fulfilled and/or that there is no change of the desired state. In case of detection of violations (e.g., failure in notifying a data subject), the obligation monitoring system will take the necessary steps to remediate them, as specified in the obligation policy.
- **Obligation Policy Repository**: this contains a list of obligation templates as well as instantiated obligations, i.e., obligations associated to specific data items and driven by data subjects’ preferences.

The various activities carried out by the Obligation Management System are logged locally and shared securely with the Audit Logs component.

The following figure shows the overall set of entities that are involved in the OMS component, along with its basic interactions:

![Figure 14. Obligation Management System - Interactions](image)

During the initialization phase, the OMS loads the parametric obligation templates and transforms them into an internal representation.

When a new data subject registers with the organisation, the C&R Provisioning component interacts with the OMS component and provides the data subject’s preferences.

The OMS component stores these preferences locally (for performance reasons) but keeps them aligned with the Data Registry Manager. These preferences identify a suitable set of obligation policies that need to be instantiated and scheduled.

Once an event happens, the Obligation Scheduler checks if any obligation (for one or more data subjects) needs to be triggered. If so, it enforces the obligation i.e. all the involved obligation actions are executed (see Section 7.2) and then it monitors it. If during the enforcement process there are failures, the OMS component (Enforcer) will try to solve these failures by repeating the execution of actions/activities that had failed. In case of these attempts still fails, related audit logs are generated and captured by the EnCoRe Audit Logs component.

All these steps are logged.
5. 7 Audit Logs Component

The Audit Logs component is a new component introduced into the EnCoRe architecture to provide a centralised point for collecting various information logged by EnCoRe components and potentially by external components, as shown in figure 15.

![Audit Logs Diagram](image)

Figure 15. Audit Logs

This component can be implemented using standard commercial products and solutions available in the market, e.g. [8].

It provides a consolidated set of logging information to support the Compliance Checking System component.

5. 8 Compliance Checking System

EnCoRe defines a set of high-level principles that various system implementations (based on the EnCoRe Architecture) should be compliant with:

1. **Consent Management**: Data subjects must be supported with clear explanation of how their data will be handled and provided with the facility to manage their consents throughout its lifetime whether the consent is originally implicit or explicit. The service provider must respect and act within those consents, and provide at least the EnCoRe minimal consent functionality.

2. **Revocation Management**: Data subjects must be provided with a facility to revoke previously given consents (implicit or otherwise). Service providers must provide at least the EnCoRe minimal revocation functionality and respect and act upon all revocation requests except to the extent that the law mandates otherwise.

3. **Openness and Support**: Data subjects must be provided with access to appropriate levels of information about the handling of their personal data and at least the minimal event transparency set, clear and accessible guidance on how to manage their personal data consents through the options provided, and
general advice on the potential risks involved in the choices available from the EnCoRe service.

4. **Service Responsiveness:** Clear commitments must be made with regards to availability of service, and the speed with which changes in preferences (new consents and revocations) will be implemented. Data subjects must be offered the facility to be informed when a service doesn’t meet pre-specified commitment levels, and the nature of any resulting unexpected data exposure.

5. **Choice Flow-Down:** Data passed between EnCoRe-compliant systems will retain the data subject’s consent and revocation choices, which will be respected by the receiving party. Where data is to be passed to third parties who are not EnCoRe compliant, explicit prior consent must be obtained from the data subject and protection be sought for the data being shared which recognises the data subject’s consent and revocation choices.

6. **Best practice:** EnCoRe services should adhere to and demonstrate compliance to EnCoRe best practice by:
   - protecting information security as documented by internationally recognized standards;
   - the EnCoRe Information Risk Management methodology;
   - being compliant with relevant data protection related legislation and following guidelines and recommendations from relevant data protection authorities;
   - ensuring employees are provided with appropriate training and support to effectively deliver the EnCoRe service;
   - adopting the appropriate EnCoRe template architectures and policies to ensure service integrity;
   - using the EnCoRe run-time compliance methods;
   - adopting EnCoRe complaints response processes;
   - operating a data minimisation policy.

Some of these principles dictate criteria and processes that need to be followed during overall risk assessment and human-driven auditing activities, such as based on ISO 2700x [9] and Privacy-by-Design specifications.

Other criteria and principles can potentially be monitored and checked against violations by analyzing log files and behaviours of the various EnCoRe components. This includes detecting violations/failures in the following areas:

   - handling personal data, accordingly to data subjects’ preferences;
   - updating or propagating updates/changes of privacy preferences;
   - operational activities carried out by the EnCoRe components;
   - granting/denying access to protected personal data.

The Compliance Checking System primarily targets these kinds of monitoring and compliance checking activities, as shown in the following figure:
The Compliance Checking System gathers logging information from the Audit Logs component.

It consists of the following two core modules:

- **Monitoring**: this is the module that, driven by compliance checking rules, continuously monitors for potential suspicious events/alerts and triggers violations/failure notifications;
- **Violation Manager**: this is the component that, based on detected violations and failures, orchestrates workflows of activities, such as notifying risk management teams, further protecting data, temporarily disabling access, etc.

Compliance checking rules need to be defined to explicitly describe the kind of checks that need to be done on audit logs and which situations can be classified as a violation and failure.

Commercial solutions, such as [8], already provide these kinds of capabilities. The EnCoRe logged information, collected via the Audit Logs components, can be plugged into these solutions, as additional log feeds.

### 5. 9 External Workflow Manager

The External Workflow Manager (EWM) Component is the key component in charge of interfacing and interacting with other deployments of EnCoRe systems. Its primary objective is to share personal data along with privacy preferences (by leveraging the sticky policy mechanisms) and ensure consistency of this information, across all the involved parties, for example when data subjects change/revoke their preferences.

This component is primarily based on the specification provided in the Second Technical Architecture [2]. The following figure provides the details:
Figure 17. External Workflow Manager

This component is activated by the Internal Workflow Manager when:

- Personal information needs to be shared with another organisation, along with related privacy preferences (and potential organisational policies);
- A data subject changes their privacy preferences which then need to be propagated to all the third parties their personal data has been shared with.

Requests for sharing data or updating privacy preferences are sent to the EWM and handled by its Workflow Orchestrator. This module checks (via the PEP) for authorisation to disclose data, based on access control policies and data subjects’ preferences.

If the authorization is granted then, based on settings provided by administrators, the sticky policy mechanism is used to disclose data and associated preferences (choices) to the third party. This involves the Data Processing, Crypto Engine and Flexible Sticky Policy Manager modules to support the various sticky policy steps, detailed in Section 8. Eventually data is disclosed to the other party, using the Data Disclosure module.

The Flexible Sticky Policy Manager module is instructed by the External Workflow Manager about the type of binding to be enforced between the personal data and meta-data (inclusive of privacy preferences and a suitable abstraction of the access control and obligation policies). We envisage at least two degrees of stickiness: weak and strong. The latter would involve mediation by a trusted authority/trusted third party (which may also be the organisation disclosing the data) to enable the access to the actual data, based on the fulfilment of agreed policies, and the former would not.

As discussed in the Second EnCoRe Technical Architecture [2], the actual level of stickiness of policies required is dictated by business, security and privacy requirements that might vary, based on the organisation’s needs, legal framework, contractual agreements with third parties and trust. The Flexible Sticky Policy Manager supports different levels of configuration to support different types of stickiness. More details follow.
Weak Stickiness:
The implementation of sticky policies does not necessarily rely on a cryptographic mechanism; a weak version of stickiness can be provided just at the logical level to bind together a set of two elements: the data and meta-data (e.g., abstracted policies, preferences, etc.) [10]. The logical binding can be provided by means of a standardised data format to convey the association of these elements in order to form a single package containing all the required information. For example, this might use a blob (Binary Large Object) file. These packages should be transmitted using best security practices to ensure that the communication channel is secure. It is important to note that weak stickiness is primarily meant to support the association and links between all relevant data, and not really mitigate any threats related to the possibility of breaking them. After having received the personal data, nothing prevents the third party from using these data and yet not enforcing the associated privacy preferences. Additional support in this direction is provided by Strong Stickiness.

Strong Stickiness:
Strong stickiness mandates usage of cryptography. This version makes sure that only the intended recipients can access the data and the policies. This assures the data subject that the data receiver is going to fulfil all the stated requirements mandated by the associated sticky policies. This includes the requirement, for the data receiver, to issue non-repudiable statements that certify its compliance to the sticky policies before getting any access to the data.

Different mechanisms can be used to achieve this. Cryptographic mechanisms, such as encryption, can be used to bind personal data to privacy preferences [11]. If the binding can only be unbound with an encryption key that is disclosed to the third party only after it has (non-repudiably) agreed to fulfil the privacy preferences, then it is possible to obtain a proof that the third party has agreed to enforce the preferences before giving it the access to the personal data (by disclosing the encryption key). But then again, after the decryption key has been issued nothing prevents the third party from using the personal data without enforcing the privacy preferences. If stronger mechanisms are needed, approaches such as those proposed by Zuo and O'Keefe [11] could be considered. However, this type of approach is currently beyond the scope of this architecture.

Section 8 discusses a reference approach and algorithm for handling sticky policies, based on standard PKI.

It is important to note that, to be effective, the External Workflow Manager has to be deployed at both organisations that interact and exchange information, as shown in the following figure. The second organisation might not necessarily have to use the EnCoRe system, as long as it can support the EnCoRe capabilities (as well as compliance principles) and can interface with the EnCoRe External Workflow Manager.
At the data receiving side, the local External Workflow Manager might need to interact with one or more Trust Authorities [11] to get access to the protected data, after the organisation has provided evidence of compliance (or reassured that it will comply) with the stated sticky policies. The interaction details are provided in Section 8.

Once the personal data (along with privacy preferences) has been decrypted (i.e. it can be normally accessed), the External Workflow Manager will interact with the related Internal Workflow Manager, to support the relevant requests i.e.:

- Provide personal data and preferences for a new data subject, OR;
- Change privacy preferences related to an existing data subject.

### 5.10 EnCoRe Configuration Database

The EnCoRe Configuration Database is primarily a support database to store information about configurations of various EnCoRe components.

This includes:

- Information about types of preferences supported by the organisation;
- Information on how to associate types of privacy preferences to personal data, driven by administrators’ specifications;
- Information about the names used in EnCoRe to identify personal information’s attributes, the association with the actual stored attributes (location in the Enterprise Data Repositories) and their physical names. This information pragmatically replaces the Ontological descriptions of data introduced in the Second Technical Architecture [2];
- Information about various EnCoRe components, in particular their configuration and locations, when run as standalone services.
The following figure shows a reference schema for the EnCoRe configuration database, for illustration purposes:

![EnCoRe Configuration Database - Schema](image)

Specifically, the following tables are used:

- **Attribute Relation**
  
  This table associates the logical name of an “Attribute” (i.e. an item in the set of personal data) to its physical name (the name used in the place it is stored) and its physical location.

- **Preferences**
  
  This table defines the set of supported privacy preferences. It specifies the type of preferences and groups them into sets, as described in Section 6. Two core types of privacy preferences need to be supported:
  - Flag or single-valued preferences, generally a value (e.g. deletion date).
  - List-based, generally consisting of preferences that have multiple values (e.g. list of chosen purposes for using data or list of allowed/disallowed third parties).

- **Lists**
  
  This table stores the possible values for “List” type preferences.

- **Resources**
  
  This table maintains a centralised repository of all the locations of distributed EnCoRe components.
6. Management of Privacy Preferences

A core capability of the EnCoRe Architecture is to support a flexible and configurable management of privacy preferences.

Data subjects are encouraged to provide their privacy preferences when disclosing personal data, dictating a variety of constraints in the form of choices, e.g. on how data should be accessed by the organisation, on notification and disclosure criteria, etc.

The first and second EnCoRe Technical Architectures [1,2] provided the key guidelines about how to define, collect and handle privacy preferences. The core set of privacy preferences supported by EnCoRe include:

- Allowed/disallowed purposes for using personal data
- Consent for disclosing data to third parties
- Allowed/disallowed lists of entities with which data can/cannot be shared
- Notification preferences
- Deletion Preferences
- Other preferences related to data handling (e.g. data minimisation, etc.)

The third specification of the EnCoRe Architecture further refines the above concepts and provides additional technical details. Section 5 has described the basic principles regarding how privacy preferences must be collected, stored and used in conjunction with access control and obligation policies.

This section provides additional details of how we envisage managing the association of privacy preferences to personal data. The following figure shows two relevant cases:

![Figure 20. Association of Privacy Preferences to Attributes in Personal Data](image-url)
The data subject’s personal data consists of a collection of Attributes (e.g. Name, Surname, Address, Credit Card number, etc.) and related values. A set of privacy preferences can be associated to attributes in a configurable way, for example:

- **CASE-1**: A single Attribute (group with only 1 attribute). This is just a special case of Case-2 where the number of attributes = 1;

- **CASE-2**: A group of Attributes (a subset of or all the Attributes defining the data subjects’ personal data)

Each group of attributes is uniquely identified in the EnCoRe Configuration Database.

This approach provides great flexibility: depending on the business needs of the organisation and privacy requirements, privacy administrators can configure the EnCoRe system accordingly.

For example, in the third case study (Cabinet Office/IDA Programme) different sets of preferences might need to be associated with different types of attributes, specifically at the Identity Provider side (see Section 10 for more details). In this context, a specific set of privacy preferences might be set for the “Minimal Data Set” (MIDS) attributes (e.g. name, surname, address, etc.) to be used for matching purposes. These preferences will define specific constraints on how data must be handled by the Hub. Another set of preferences can be associated with all the attributes, dictating how the Identity Provider and Service Providers should handle this data.

To summarise, the approach described in this document ensures that configurable groups of attributes can be associated to configurable groups of preferences.

### 6.1 Technical Implications of Privacy Preference Management

Section 5 discussed how the updated EnCoRe architectural components implement this model for handling privacy preferences.

Specifically, this approach is supported by the following EnCoRe components:

- **Internal Workflow Manager**: orchestrating the collection and storage of privacy preferences;
- **EnCoRe Configuration Database**: explicitly storing the association of types of data attributes to types of preferences, as supported by an organisation;
- **Data Registry Manager**: storing the actual privacy preferences and the associations to personal data;
- **Privacy-Aware Access Control Policy Enforcement**: using stored privacy preferences to evaluate and enforce access control policies;
- **Obligation Management System**: using stored privacy preferences to evaluate, enforce and monitor obligation policies;
- **External Workflow Manager**: orchestrating the disclosure of privacy preferences. The supported associations of privacy preferences to attributes can be represented in a variety of ways. These include:

1. Storing them into databases, as described for the EnCoRe Configuration database;
2. Exchanging these representations in specific formats (e.g., XML, JSON, etc.) between EnCoRe components, when sending requests that involve personal data (attributes) and privacy preferences.

When implementing systems based on the EnCoRe architecture, we strongly encourage following the suggested approach to handle privacy preferences, as it has been tested in various case studies and provides the required level of flexibility [19,20].

### 6.2 Implications for Data Subjects

The way privacy preferences are associated to personal data also heavily impacts how they are collected, via GUIs, from data subjects.

EnCoRe aims to support fine-grained association of privacy preferences to personal data. The approach described in this section fully supports this requirement.

In addition, EnCoRe aims to support different levels of abstraction in the way privacy preferences are collected from data subjects, enabling its adoption by a wide range of people with different skills.

Different paradigms can be supported, including:

- Graphical Sliders defining different levels of privacy (e.g., MIN, MED, MAX): the data subject can choose the appropriate level of privacy. Each level is associated with a specific set of values for privacy preferences;
- Allowing data subjects to specify values of privacy preferences, for specific data attributes/groups of attributes, based on interests;
- Hybrid approaches involving a mixture of the above.

The following figure shows a simple example of how privacy preferences can be collected from data subjects by simply using graphical sliders in web interfaces, associated with groups of attributes.

The following figure illustrates an EnCoRe-enabled web page where the data subject registers with a Service Provider by providing their personal data and privacy preferences:
This approach might be suitable for those people that do not want to be exposed to underlying details and who trust EnCoRe to provide suitable default settings.

The following figure shows a related example (a hybrid approach) where the data subject can further refine (if desired), in a fine-grained way, the privacy preferences associated with groups of attributes:

Figure 21. Example 1: Gathering Privacy Preferences from Data Subjects (simple approach)
In this example, the data subject can fully specify the values of specific privacy preferences by using “privacy preference customisation” capabilities.

The EnCoRe Architecture and the recommended model to manage privacy preferences support a wide variety of approaches to graphically gather privacy preferences from data subjects.
7. Specification of Privacy-Aware Access Control Policies and Obligations

This section provides a refined overview of the different types of policies that are handled by the components of the third EnCoRe Technical Architecture. These policies include access control policies and obligation policies which are described in separated subsections below.

The key principles and concepts underpinning these policies have been discussed in great detail in the first and second EnCoRe Technical Architectures [1,2]. This document aims to further refine them by providing more details and examples.

It is beyond the scope of this document to fully specify the access control and obligation policies. However, the aim is to create awareness about the types of constraints that are managed and how they relate to data subjects’ preferences and different types of enforcement and monitoring requirements.

7.1 Privacy-aware Access Control Policies

Privacy-aware access control policies are used to specify the security and privacy constraints that data requestors (employees, applications, services, third parties, etc.) need to satisfy in order to access personal data, associated with data subjects.

These policies take into account data subjects’ privacy preferences. Their evaluation, described in Section 5, changes depending on the specification of privacy preferences and any change of their value that could happen over time, determined by data subjects’ requests.

Figure 23 shows the conceptual, key elements that need to be part of privacy-aware access control policies and their relationships:
Privacy-aware access control policies must specify:

- A logical name and unique identifier of the policy;
- The target of the policy, i.e. specific attributes or all attributes in a data subject’s personal data;
- Triggering rules/context for activating the evaluation of policy rules;
- A set of rules, consisting of:
  - Conditions, defined via expressions;
  - Actions, defined as a list of activities to be done, including the access control decision (allow/deny).

Additional details about policy expressions and conditions are shown in the following figure:
The definition of an “access control policy” is parametric. An access control policy applies to any data target (set of attribute types, related to personal data) and to any data subject. It is instantiated with the right context at the access request time, by the PDP, as discussed in Section 5.

This high-level style of definition of access control policies is compatible with standard policy definitions, such as XACML. However, it specifically targets the specification of security and privacy constraints on accessing of personal data.

A simple example of an access control policy using all the above concepts is shown below. An XML-based notation has been used to represent it, for illustrative purposes only.

```xml
<policy>
  <name>policy-ID1</name>
  <target>ALL-ATTRIBUTES</target>
  <trigger>
    <expression>
      <and>
        <condition>Request.Obj.Location=="PII.DB"</condition>
        <condition>Context.Request.Purpose contained in Context.PrivacyPreferences.Purpose</condition>
      </and>
    </expression>
  </trigger>
  <rules>
    <rule>
      <expression>
        <or>
          <condition>Request.Role=="Admin"</condition>
          <and>
            <condition>Request.Purpose=="Marketing"</condition>
            <condition>Request.Role=="Marketer"</condition>
          </and>
          <and>
            <condition>Request.Purpose=="Research"</condition>
            <condition>Request.Role=="programmer"</condition>
          </and>
        </or>
      </expression>
      <action>
        <decision>yes</decision>
      </action>
    </rule>
  </rules>
</policy>
```

This example of access control policy (policy-ID1) is associated to all attributes in the data subjects’ personal data that is stored in the Enterprise Data Repository (PII
database). It is triggered when the requestor’s purpose for accessing data is compatible with the list of purposes defined by the data subject (privacy preference). *NOTE: access control policies could be defined for a specific subset of attribute names. In this case the policy specifies the list of relevant attribute names.*

If the rule is triggered, then access is granted if the conditions (checking the requestor’s role and the related purposes) are satisfied, i.e. any administrator, a marketer doing marketing or a programmer doing research.

We recommend that systems implementing the EnCoRe Architecture support both:
- a human readable representation of access control policies (to be graphically authored by an administrator) and
- an internal representation within the PDP component that is optimised and efficient, to support high levels of evaluation requests.

### 7.2 Obligation Policies

The second EnCoRe Technical Architecture [2] introduced the concept of Obligation Policies, and specifically the core concepts at the base of Parametric Obligation Policies, to be managed and handled by the OMS system.

As described in [2], a Parametric Obligation Policy could be represented conceptually as:

```
FOR: Target
WHEN Events(Refs)
THEN EXECUTE [Actions(Refs)]
ON VIOLATION:
    EXECUTE [Violation-Actions(Refs)]
```

For a given target (personal data), when specific (parametric) set of events happens then a set of (parametric) actions are executed. In case of violation of the policy a set of on-violation actions are specified to remediate the issues.

More specifically:
- A Parametric Obligation Policy can be associated with a potentially large set of personal data (i.e., no multiple instantiations of the policy is required on the single privacy preferences) and, at the same time, it can dictate customized obligation constraints (based on data subjects’ privacy preferences) over each data item;
- A Parametric Obligation Policy does not embed privacy preferences in its Events and Actions sections. Instead, it contains explicit references to these preferences, which are stored elsewhere (in data repositories);
• The Target section of a Parametric Obligation Policy explicitly describes the data repositories that will contain preference values pointed at by these references - in addition to repositories containing personal data;
• An On-Violation section automates the process of remediation of violated obligations.

The key feature of parametric obligations is implied by the second bullet above. Privacy preferences are stored separately from parametric obligation policies: references in the obligation policies are used to retrieve the privacy preferences. This ensures that a parametric obligation policy can apply to a potentially large set of personal data – as defined in its Target element – and, at the same time, allows the “customization” of its Events and Actions based on references to external privacy preferences.

A set of parametric obligation policies can be created by a privacy administrator to dictate the criteria by which personal data should be handled. The referencing mechanism (coupled to appropriate data descriptions in the Target section) ensures that these policies are instantiated on-the-fly by the Obligation Management System, based on associated privacy preferences, and enforced and monitored for a potentially large set of managed data.

As specified by [2], different types of obligations need to be supported:

• **One-off obligations**: these are executed once, when triggered by specific events. For example the deletion of data attributes at a specific time and date;
• **Recurrent obligations**: these are periodically executed, such as notifying data subjects every X days, if their personal data is still retained by the organisation.

This document provides additional details and clarification about obligation policies. The following figures show the key conceptual elements that need to be part of obligation policies and their relationships:
The structure of the obligation policies is similar to the one used for access control policies. However, the triggering of an obligation policy is caused by events. The policy describes the relevant set of events (e.g. access events, time events, etc.) that will trigger it.

In EnCoRe, the OMS component receives events from the Event Manager: they are used by the scheduler to check if any obligation policy can be triggered.

If an obligation policy triggers, then its actions are executed by the OMS system, based on the policy definition. This might include notifications, data deletion, etc. In case of failures in executing these actions, the Violations section describes how to proceed, i.e. which remediation actions to execute.

A simple example of an access-control policy using all the above concepts is shown below. An XML-based notation has been used to represent it, for illustrative purposes only.
This is an example of a “one-off” parametric obligation. It is triggered when access control events are received by the OMS, either related to access granted or denied.

The triggering part of the event checks (via the associated contextual information) if the “access event” relates to an access request the purpose of which is compatible with data subjects’ chosen purposes (as specified in the privacy preferences). This ensures that the obligation is relevant for a data subject, i.e. it is triggered only when their personal data could have been potentially accessed by the requestor.

When the policy is triggered, the data subject is sent a notification containing additional contextual information. In the case of failure (in enforcing an obligation policy), an audit log is created and the OMS shows the failure to the administrators via its graphical GUI.

As with access control policies we recommend that the systems implementing the EnCoRe Architecture support both:

- a human readable representation of obligation policies (to be graphically authored by administrator) and
- an internal representation within the OMS component that is optimised and efficient, to support high loads of evaluation requests.
8. Sticky Policy Management

The second EnCoRe Technical Architecture [2] introduced and described how we envisage using the sticky policy mechanism [11] to ensure secure exchange of personal data between EnCoRe systems and provide degrees of assurance.

Section 5 discussed various potential approaches to sticky policies, including supporting weak and strong stickiness of policies to personal data.

At the minimum, sticky policies could just consist of the set of privacy preferences defined by the data subject. EnCoRe wants to ensure that these preferences propagate to the third parties with which personal data is shared: this is to ensure that personal data can be consistently managed, based on data subjects’ preferences, wherever it is disclosed.

A more advanced version of sticky policies can include also sanitised/abstracted versions of access control and obligation policies, consistent with confidentiality, security and business criteria defined by the organisation.

The approach discussed in this Technical Architecture is flexible and supports a variety of definitions of sticky policies. It is beyond the scope of this document to discuss how to tailor sticky policies, as this depends on organisational choices and decisions.

It is also beyond the scope of this document to discuss the details of how to implement sticky policies. Full information can be found online [11, 13].

However, this section provides a reference description of a mechanism and approach to support strong sticky policies in EnCoRe, using standard PKI [14] capabilities.

The following figure shows the various steps involved in the sticky policy mechanism and related protocol, used in EnCoRe:
We consider the generic case where two entities (e.g. two EnCoRe systems deployed in two different organisations) want to exchange personal data: specifically, Entity 1 wants to send personal data to Entity 2 (e.g. because of a business transaction) and have degrees of reassurance that this data will be managed and handled according to a few specified policies. In the figure, for simplicity, we refer to personal data also as PII.

Entity 1, Entity 2 and the Trust Authority use Public Key Cryptography (PKI) certificates. They each have a private key (protected) and a public one, associated to a public certificate, signed by a trusted certificate authority.

The following basic steps take place:

1. Entity 1 defines the sticky policy, policy 1, to be associated with the personal data;
2. Entity 1 generates a symmetric key $K$, to encrypt the personal data, also referred as PII in the figure (symmetric keys are used as more efficient and support one-off usage). The Private Key of Entity 1 is used to sign a package of data including: policy 1, a reference to the required Trust Authority (specified via its public key and PKI certificate) and the encrypted symmetric key $K$, by using the Trust Authority’s public key. The sticky policy package is created, including the encrypted data, the encrypted symmetric key $K$, the above signed information and policy 1. This package is sent to Entity 2;
3. Entity 2, in order to get access to the encrypted data, needs to understand the associated policy 1 and agree to fulfil it. In the case that it does, it has to interact with the recommended Trust Authority. It passes to the Trust Authority the policy package and
4. Checking and challenges
5. A. $\text{Enc}(\text{PubE2}, K \| \text{policy1})$
Authority the package containing the signed information, policy 1 and the encrypted symmetric key K;

(4) The Trust Authority engages in a set of checking and challenges with Entity 2 to get signed assertions that Entity 2 understands policy 1 and will comply with its constraints and requirements. It logs all these assertions and interaction outcomes, for future audit and forensic analysis;

(5) If the Trust Authority is satisfied with Entity 2’s assertions, it will release the session key K to Entity 2, encrypted with Entity 2’s public key;

(6) Entity 2 gets access to the session key K and decrypts the personal data.

As shown in Figure 27, there are two options for the policyKeyMapping function that sticks the policy to the key that is used to encrypt the data, and thereby can be thought of as performing the sticky policy functionality. The first option is signcryption, and the second is encrypt-then-sign. These correspond to the following functions respectively: policyKeyMapping(PrivU, PubTA, K||policy) = Sigenc(PrivU, PubTA, K||policy) or policyKeyMapping(PrivU, PubTA, K||policy) = (U, V), where U = A.Enc(PubTA, K||policy||IdU) and V = Sig(PrivU, U||IdTA) and the following notation is used:

- “||” denotes concatenation
- $S.Enc()$ denotes a symmetric encryption algorithm, e.g. specified in ISO/IEC 18033-3
- $A.Enc()$ denotes an asymmetric encryption algorithm, e.g. specified in ISO/IEC 18033-2
- $Sig()$ denotes a signature algorithm, e.g., specified in ISO/IEC 14888 or ISO/IEC 9796
- $Sigenc()$ denotes a signcryption algorithm, e.g. specified in ISO/IEC 29150

Alternatively, $K||policy1$ can be replaced with $K||h(policy1)$, where $h()$ is a hash-function, e.g., as specified in ISO/IEC 10118-3.

This mechanism provides the core capabilities for enabling the sharing of personal data in a secure way and for providing degrees of assurance to Entity 1.

Based on implementations and the context, the exchanged sticky policies can dictate a variety of constraints, at different levels of abstraction, including that the trustworthiness of Entity 2’s IT infrastructure and compliance to EnCoRe principles have been checked. This is handled by the Trust Authority.

In the EnCoRe Architecture, when the EnCoRe system is deployed in two different organisations, the roles of “Entity 1” and “Entity 2” is played by the EnCoRe External Workflow Manager components, as shown in the figure below:
Section 5 described the overall interaction flows in the EnCoRe architecture, associated with the disclosure of personal data to a third party and detailed the capabilities provided by the EnCoRe External Workflow Manager component.
9. Design and Deployment Options

The design and deployment options for this third EnCoRe Technical Architecture are the same as for the second [2].

Also in this architecture, interoperability between EnCoRe systems, deployed across multiple systems, is very important. As discussed in Section 5, we recommend using the EnCoRe Configuration Database to provide a clear definition of the supported types of privacy preferences and types of attributes, along with their mappings to personal data and its physical storage. This is the core information that must be harmonised across multiple deployments of EnCoRe systems.

As for the second EnCoRe Technical Architecture [2], this third architecture strongly assumes compliance with best security practices, as discussed in Section 5. This includes designing and deploying all the core EnCoRe components as secure, self contained services along with a clear definition of their APIs and secured interaction protocols. This approach enables a flexible deployment of the architecture, based on actual needs dictated by the target case study, together with business and security requirements.

Specifically, the EnCoRe Technical Architecture encourages the approach where each architectural component is run as an independent, secured service. Various frameworks could be used to achieve this including the EJB [15] and REST/RESTful [16] ones.
10. Use Cases

The third EnCoRe Technical architecture has been designed by factoring in requirements related to both general use cases and specific use cases for the third case study, centred on the Cabinet Office/IDA Programme [17].

The general use cases described in this document complement and refine the ones described in the first and second EnCoRe Technical Architecture [1,2].

For both categories of use cases, this section describes, in more detail, the various interaction flows involved in the EnCoRe Architecture.

In the specific use cases for the Cabinet Office/IDA Programme we provide more details about how EnCoRe can be effectively used to provide the required consent and privacy management capabilities. This description takes into account the information available at the time of writing this document. It also takes into account confidentiality issues, so all the details provided in this document are based on public domain Cabinet Office/IDA Programme information.

10.1 General Use Cases

Section 4 introduced the generic four use cases supported by the EnCoRe Architecture. These use cases are common to all the case studies we have so far investigated and we believe apply to most situations/scenarios.

This section provides additional details about how the EnCoRe Architecture supports them. The description of the use cases focuses on the aspects of relevance to EnCoRe, i.e. consent and privacy management. We assume that suitable authentication processes are in place and security is deployed based on best practices.

10.1.1 USE CASE 1: A data subject discloses personal data along with consent/privacy preferences

This use case is about a data subject that discloses their personal data to an organisation along with their privacy preferences.

The EnCoRe system, deployed within the organisation, configures various internal components to ensure the fulfilment of these preferences, along with the enforcement of access-control and obligation policies.

The following figure shows the involved steps and the affected EnCoRe components.
Figure 29. General USE CASE 1

The following steps and components are involved:

1. The data subject provides their personal data to the organisation, via a web interface, along with the specification of related privacy preferences;
2. At the organisation back-end, the Consent & Revocation Provisioning component intercepts this activity and activates the Internal Workflow Manager (IWM). The IWM coordinated the various next steps;
3. The IWM, after asking for authorization to the Privacy–aware Access Control Policy Enforcement component, stores the personal data in the Enterprise Data Repository. It retains a reference of the storage location;
4. The IWM interacts with the Data Registry Manager to request the storage of the associated privacy preferences, along with the additional metadata (references to personal data, etc);
5. The Data Registry Manager stores the information in the Data Registry;
6. The IWM interacts with the Privacy–aware Access Control Policy Enforcement component to set any local configuration, for example about relevant access control policies;
7. The IWM interacts with the Obligation Management System (OMS) to activate the relevant obligation policy templates along with the contextual privacy preferences.
10.1.2 USE CASE 2: Employees and/or applications try to access data for specific purposes

This use case is about an employee and/or an application/service trying to access personal data, stored within the organisation, for specific purposes.

The EnCoRe system, deployed within the organisation, intercepts this attempt and enforces related access control and obligation policies.

The following figure shows the involved steps and the affected EnCoRe components.

Figure 30. General USE CASE 2

The following steps and components are involved:

1. An application/service (and/or an employee using them) tries to access personal data protected by the EnCoRe system. This attempt is intercepted by a local Policy Enforcement Point – PEP (e.g. by using local interception point within the application or data repository interceptors) along with contextual information, including the identity of the requestor, the requested information and the purpose for accessing the data. The PEP interacts with the Privacy-aware Access Control Policy Enforcement component to get the authorization to proceed;
2. The Policy Decision Point (PDP) in the Privacy-aware Access Control Policy Enforcement component makes a decision to grant/deny access, based on the privacy-aware access control policies and the context;
3. Based on the decision outcome, the PEP component enables or blocks the attempt to access personal data;
4. The Obligation Management System, based on generated events (received from the Event Manager) might trigger related obligation policies (e.g. sending notification to data subjects).

10.1.3 USE CASE 3: Data subject changes their consent/privacy preferences

This use case is about a returning data subject (customer) that wants to change her/his privacy preferences, associated to their personal data.

The EnCoRe system, deployed within the organisation, configures various internal components to ensure the fulfilment of these preference changes, along with the enforcement of access-control and obligation policies.

The following figure shows the involved steps and the affected EnCoRe components.

![Figure 31. General USE CASE 3](image)

This use case is very similar to USE CASE 1, with the exception that information is updated. The following steps and components are involved:

1. The data subjects retrieve their personal data and privacy preferences, via a web interface; make changes affecting these preferences (and potentially the associated data); submit these changes;
2. At the organisation back-end, the Consent & Revocation Provisioning component intercepts this updated information and activates the Internal Workflow Manager (IWM). The IWM coordinated the various next steps;
3. The IWM, after asking for authorization to the Privacy-aware Access Control Policy Enforcement component, updates the personal data in the Enterprise Data Repository – if any change to personal data has been made. It retains a reference of the storage location;
4. The IWM interacts with the Data Registry Manager to update the associated privacy preferences, along with the additional metadata (references to personal data, etc);
5. The Data Registry Manager stores the updated information in the Data Registry;
6. The IWM interacts with the Privacy-aware Access Control Policy Enforcement component to set any local configuration, for example about relevant access control policies;
7. The IWM interacts with the Obligation Management System (OMS) to update the relevant obligation policy templates along with the contextual privacy preferences.

If the involved personal data has been disclosed to third parties, privacy preferences will be updated also in these parties, by using an approach similar to the one discussed in USE CASE 4. The list of the entities where personal data has been disclosed to is retrieved from the Data Registry Manager by the IWM. The IWM interacts with the EWM to ensure that requests for updates are sent to the relevant third parties.

10.1.4 USE CASE 4: Personal data disclosed to a third party

This use case is about an organisation (Enterprise A) disclosing personal data to a third party (Enterprise B), for specific purposes (e.g. due to a business transaction).

The EnCoRe system, deployed within both organisations, handles the overall process of disclosing data, consistent with the data subject’s privacy preferences and related access control and obligation policies.

The following figures show the involved steps and the affected EnCoRe components within both organisations: Enterprise A (sender of personal data) and Enterprise B (receiver).
The following steps and components are involved:

**Enterprise A**

1. An application/service (and/or employee using them) in Enterprise A needs to share data with another organisation, for a specific purpose. This disclosure attempt is intercepted by an internal interception point and redirected to the Internal Workflow Manager (IWM);
2. The IWM ask for authorization to the Privacy-aware Access Control Policy Enforcement component that makes a grant/deny decision based on data subjects’ privacy preferences and relevant access control and obligation policies;

3. If authorised, the IWM interacts with the External Workflow Manager (EWM) proceed with the disclosure of data based on agreed processes. In particular, the EWM might have been configured to share data with the sticky policy mechanism, associating (at least) privacy preferences to the personal data to be disclosed. We assume this is the case;

4. The EWM interacts with the EWM counterpart, in Enterprise B, to disclose the protected personal data, along with the associated sticky policies;

**Enterprise B**

5. The EWM at Enterprise B interacts with a selected Trust Authority (e.g. Enterprise A’s or a third party entity) to get access to the personal data, after satisfying the various involved policy constraints. If it gets access, the same steps discussed in USE CASE 1 apply;

6. After getting access to the personal data (and related privacy preferences), the EWM contacts the Consent & Revocation Provisioning component which activates the Internal Workflow Manager (IWM). The IWM coordinates the various next steps;

7. The IWM, after asking for authorization to the Privacy-aware Access Control Policy Enforcement component, stores the personal data in the Enterprise Data Repository. It retains a reference of the storage location;

8. The IWM interacts with the Data Registry Manager to request the storage of the associated privacy preferences, along with the additional metadata (references to personal data, provenance, etc);

9. The Data Registry Manager stores the information in the Data Registry;

10. The IWM interacts with the Privacy-aware Access Control Policy Enforcement component to set any local configuration, for example about relevant access control policies;

11. The IWM interacts with the Obligation Management System (OMS) to activate the relevant obligation policy templates along with the contextual privacy preferences.

This mechanism is used also in those cases where data subjects want to update their data and privacy preferences and (part of) their personal data has been disclosed to third parties.

A specific case is when data subjects require the organisation to delete their personal data. In this case the organisation propagates the deletion requests to the relevant third parties, before proceeding to delete the information stored locally.
10.2 Use Cases for the Third Case Study

The third EnCoRe Case Study focuses on the Cabinet Office/Identity Assurance (IdA) Programme [17]. This section provides background information about the IdA Programme, followed by a detailed description of related use cases, illustrating how EnCoRe provides the required consent and privacy management capabilities.

We anticipate that all the involved use cases can be considered as specific instances of the four GENERAL USE CASES described in the previous section.

10.2.1 Setting the Context

A recent UK Government ministerial statement [17] provides an overview of the Identity Assurance Programme and related objectives:

“The Minister for the Cabinet Office and Paymaster General (Mr Francis Maude): The Government agreed on 14 March 2011 to the development of a consistent, customer-centric approach to digital identity assurance across all public services. This will allow service users to log on safely to digital public services in a way that ensures personal privacy, reduces fraud and facilitates the move to online public services.

Today I am setting out the context and vision for this delivery programme and will explain how my Department will draw on expertise from organisations in the public and private sector to agree the design for this new approach. NO2ID and other privacy advocates are being kept closely informed of developments and given the opportunity to contribute and comment. This is a cross-departmental initiative and the approach will be introduced in the near term through major Government initiatives such as DWP’s universal credits, NHS HealthSpace, HMRC’s one click programmes and the Skills Funding Agency.

Online services have the potential to make life more convenient for service users as well as delivering cost savings. However, currently customers have to enter multiple log-in details and passwords to access different public services, sometimes on the same website. This involves significant duplication, is expensive to operate and is highly inconvenient for users. It acts as a deterrent to people switching to digital channels, hampers the vision of digital being the primary channel for accessing Government information and transactions, and provides an opportunity for fraudsters.

Our intention is to create a market of accredited identity assurance services delivered by a range of private sector and mutualised suppliers. A key improvement will be that people will be able to use the service of their choice to prove identity when accessing any public service. Identity assurance services will focus on the key imperative to ensure privacy. My Department is leading the project to develop the design and the creation of the market within the private sector. By October 2011 we expect to have the first prototype of the identity assurance model to test with transactional Departments and public sector identity assurance services, with a date for implementation from August 2012.”
As summarised in the EnCoRe internal report - “3rd Case Study user’s Requirements Workshops”, the principles of the IdA programme include the following statements:

“Customer focus: an identity assurance solution must be based around the needs of the individual otherwise it will not be used or valued. People have different and changing needs. No one ‘big brother’ solution will meet the needs of all customers in all contexts”

and:

“Customer control: the use of a customer’s identity and personal data should be fully transparent and controlled by the customer.”

These are core principles that are fully supported by EnCoRe, in particular when the goal is to enable both data subjects (customers, citizens) to express and change their consent (on their personal data) and organisations to explicitly enforce the management of consent and privacy.

This section illustrates how this can be achieved by using the EnCoRe Architecture and related concepts: a set of use cases, tailored for the Identity Assurance Case Study, are presented and discussed.

The Identity Assurance programme aims to deliver a rich ecosystem of services. The following figure shows the involved entities and stakeholders:

![Simplified Architecture of Identity Assurance Ecosystem](image)

Specifically, the ecosystem involves these key players:

- **Citizens (data subjects):** these are people that want to engage online with public and/or private service providers in a variety of transactions involving business, financial, government and personal related aspects. In order to engage, they need to be authenticated, i.e. their identity needs to be certified
and assessed, potentially with different level of assurance depending on the type of interaction;

- **Identity Providers**: these are approved and trusted organisations (e.g. the UK Post Office, Banks, Mobile Phone companies, etc.) that provide the relevant processes to certify the identity of citizens, with different degrees of assurance (depending on their usage) and ensure that citizens can use these identities (and related personal information) when engaging, online, with service providers;

- **Public/Private Service Providers**: they provide services online, including government services and private ones. They rely on assertions made by Identity Providers, by using standard identity federation protocols, to gather and assess the identities of their users. In the first instance, the service providers will only by public service providers but the intention is that the framework will generalise to include private service providers as well

- **Hub**: this is a mediation component, run by the UK government to provide trust and transparency in the interactions happening between Identity providers and Public Service Providers. In the context of a business transaction involving a citizen, these two stakeholders are not aware of each other. Interactions are mediated by the Hub. This should reduce the risk of collusion and linkage of information as well as provide degrees of auditing of the IdP and PSP activities. Furthermore the Hub might be requested to add additional personal data (attributes) to information exchanged between the IdP and PSP, based on the nature of the transaction and the availability of this information within Attribute providers (such as the DWP, HMRC, DVLA departments);

- **Attribute Providers**: these are existing sources of personal information for citizens (e.g. DWP, HMRC, DVLA, etc.). They can be indirectly involved in citizens’ business transactions, via the mediation of the Hub, to provide additional personal details to support citizens/data subjects’ interaction needs (e.g. DVLA supplying their driving license number or car registration number in case of interactions with an insurance company or a parking service).

As mentioned above, the Identity Assurance programme aims to use standard federated identity management solutions, such as [18], to enable the relevant interactions.

Specifically, a citizen, when trying to access an online PSP service, is redirected, via the Hub, to a trusted IDP of choice, where they can be identified and authenticated. The citizen does this by providing their authentication credentials (the type of credentials to be used might change depending on the required level of assurance).

Once authenticated at the IdP site, a **Minimum Data Set (MIDS)** i.e. basic personal data such as name, surname, etc.) necessary to identify the data subject is passed to the Hub that might enrich it by adding additional information retrieved from Attribute
Providers. Finally the Hub passes the MIDS data, along with any additional information, to the PSP, for local matching if identities (i.e. local identification/authentication) and to enable the citizen to access the desired services.

The goal is to ensure that the asserted identity of a citizen can be successfully used at the PSP site, to identify the citizen based on the locally available information.

It is important to notice that, in the described scenario, lots of personal data can potentially be exchanged between the various stakeholders, related to authentication, matching (MIDS) and business transactions.

*To make this programme successful, it is important that citizens (data subjects) have control over how their personal data is disclosed between the various stakeholders and subsequently used; they must be allowed to change their consent and related privacy preferences at any time; they must have degrees of assurance that their preferences are enforced by the various stakeholders.*

EnCoRe can help to provide citizens with the desired level of control over their personal data and the involved organisations with mechanisms and solutions for enforcing privacy and consent.

The following figure shows the ideal situation where EnCoRe is deployed within the key players in the ecosystem (IdPs, Hub, PSPs) and provides the required consent and privacy management capabilities:

For the purpose of this document, we primarily focus on use cases involving interactions of data subjects with an IdP as they have been explored in more detail with the Cabinet Office/IDA Programme and for which we have degrees of validation of the value provided by EnCoRe.
It is important to notice that, in the current IDA specification, the Hub is stateless, i.e. it does not store information about data subjects and preferences. Because of this, we cannot assume that the Hub will store privacy preference information about data subjects and the whereabouts of personal data.

An EnCoRe system deployed within the Hub needs to access this information stored in a Remote Data Registry. This Data Registry could be hosted by a third party, by the IdP (assuming that the content is encrypted and cannot be accessed by the IdP) or by a system under the control of the data-subject, via their Privacy Assistant.

In the remaining part of this document we assume that the Remote Data Registry is actually owned by the data subject. The information about how to access can be passed as an additional preference, specified by the data-subject at the IdP level.

10.2.2 Use Cases

The following use cases have been identified:

- **Use Case 1**: a citizen (data subject) provides consent for the use of their personal data as MIDS
- **Use Case 2**: a citizen provides consent for the use of selected Attribute Providers for the MIDS matching process
- **Use Case 3**: a citizen provides consent for sending / using further Verified Attributes
- **Use Case 4**: ensuring privacy in transactions through the Hub by using sticky policies
- **Use Case 5**: changing and propagating data & consent updates
- **Use Case 6**: a citizen revokes consent for an IdP to hold their data at all

The remaining part of this section provides details about how EnCoRe supports these use cases.

10.2.2.1 Use Case 1

This use case is about enabling a citizen (data subject) to provide consent for the use of their personal data as MIDS.

It involves the citizen registering online with an IdP (once their identity has been certified by the IdP and credentials issued), releasing personal data along their consent on using parts of their personal data (e.g. name, surname, address, etc.) as MIDS information i.e. for matching purposes with PSPs.

This use case directly maps to the GENERAL USE CASE 1 described in this section and the consent management criteria discussed in Section 6.
The specific consent information about MIDS data can be captured by using an IdP web form. This form enables the citizen to explicitly give consent on which data items can be used for MIDS purposes or, more realistically, give consent to use MIDS data for matching purposes.

We assume that the IdP uses the EnCoRe system in order to capture, store and handle personal data along with privacy preferences. The following picture shows (as for the GENERAL USE CASE 1) the various involved steps in the EnCoRe Architecture.

The EnCoRe system will enforce privacy preferences at the time of sending MIDS attributes to the Hub, during a citizen’s interaction with the PSP.

Based on the chosen privacy preferences, this might mean having to abort the transaction, if the citizen gave no consent.

10.2.2.2 Use Case 2

This use case is about a citizen providing consent for the use of selected Attribute Providers for the MIDS matching process.

Also this use case directly maps to the GENERAL USE CASE 1 described in this section and the criteria discussed in Section 6. It is strongly related to the previous use case too.

The IdP can enable the citizen to explicitly give consent about which additional attributes and Attribute Providers can be involved during the MIDS matching process.
For example, the citizen might give consent that further interactions could happen with DVLA and the driving license be used for MIDS matching purposes. However, the citizen might give no consent for having interactions with DWP.

In general, the EnCoRe system can support the gathering of a wide set of privacy preferences related to personal data:

- Allowed/disallowed Attribute Providers to be used for MIDS matching purposes;
- Allowed/disallowed set of attributes that can be gathered from allowed Attribute Providers;
- Agreed purposes for using MIDS data;
- Notification preferences;
- Allowed/disallowed PSPs which MIDS information can be sent to;
- etc.

These preferences are captured, processed and deployed in the EnCoRe system as discussed in the previous use case and the GENERAL USE CASE 1.

It is important to note that only some of the preferences associated with personal data can be locally enforced at the IdP site (e.g. sending notification about disclosures).

Other preferences can only be enforced at the Hub side, as only the Hub has the knowledge about the PSP involved in the interaction. This includes preferences about which Attribute Providers can be used and which additional attributes can be retrieved.

The EnCoRe system, deployed within the Hub, will be in charge of enforcement. The EnCoRe sticky policy mechanism will ensure that preferences are associated to the MIDS data as well as provide degree of protection of the data.

The EnCoRe system, deployed within the Hub, uses the Remote Data Registry to store information about data whereabouts (i.e. which PSPs data has been disclosed to). This might include information about which Attribute Providers store data subjects’ personal data) and related preferences. This information will only be accessible by the Hub during the transaction with the authenticated data subject.

10.2.2.3 Use Case 3

This use case is about a citizen providing consent for sending / using further Verified Attributes to PSP.

It is handled exactly as in the previous two use cases. However, in this case the citizen expresses their privacy preferences about the entire set of personal data, as discussed in the GENERAL USE CASE 1. Section 6 discussed how different set of preferences can be associated to different sets of attributes and how EnCoRe fully supports this mechanism.
Also in this case, some of the preferences can be locally enforced at the IdP site, others will be enforced at the Hub level. This definitely includes preferences about which entities the data can be disclosed to, as this information is only available at the Hub level.

Also in this case, the EnCoRe system, deployed within the Hub, uses the Remote Data Registry to store information about data whereabouts and related preferences. This information will only be accessible by the Hub during the transaction with the authenticated data subject.

The EnCoRe sticky policy mechanism will ensure the propagation of personal data and privacy preferences.

10.2.2.4 Use Case 4

This use case is about ensuring privacy in transactions through the Hub by using sticky policies.

We assume that the Hub uses the EnCoRe system. During a business transaction involving a citizen with a PSP, a few initial interactions are required between the IdP and the PSPs for identification and authentication purposes.

The EnCoRe system deployed at the IdP side can send personal data and privacy preferences to the PSP by using the sticky policy mechanism. This can be achieved with the steps discussed in the GENERAL USE CASE 4, in this section.

In this case the EnCoRe system deployed within the Hub has the required information on how to handle the personal data, based on the privacy preferences (contained within the sticky policies). This includes enforcing consent about selection of Attribute Providers, attributes to be used, PSPs to be used, etc.

If the Hub needs to collect additional data from Attribute Providers, citizens’ consents can be locally enforced. In case citizen’s consent was given, the additional personal data can be added to the sticky policies package, by using the same mechanism (i.e. the additional data will be encrypted with the same sticky policies mandating the citizens’ privacy preferences).

The Hub does not necessarily need to obtain access to the actual personal data associated to the sticky policies, but only enforce consent whilst the data is in transit. In this case the interaction with the Trust Authority, to get access to the data in clear, can be disabled.

This approach ensures the confidentiality of the data in transit: the EnCoRe system is purely used to enforce the associated preferences when interacting with the PSP.
The Hub will then, recursively, use the same mechanism, as discussed in the GENERAL USE CASE 4, to share information with the PSP, by means of the sticky policy mechanism.

It is important to notice that the Hub keeps track of the whereabouts of citizens’ personal data only by using a Remote Data Registry Manager (associated to the data subject): no local information is stored. However, the process steps are logically the same as for the GENERAL USE CASE 4.

10.2.2.5 Use Case 5

This use case is about a citizen that decides to change/update their consent and personal data.

We consider the scenario where the citizen does this by interacting directly with the IdP. In this case, basically the same steps discussed in the GENERAL USE CASE 3 apply with the variant where the Hub uses Remote Data Registries. In this specific scenario, it involves a chain of local updates and update notifications to third parties:

- **Local update of information within the IdP and propagation of update requeststo the Hub**: the EnCoRe system deployed at the IDP site locally updates the preferences and data (GENERAL USE CASE 3). By using the Data Registry Manager, it sends notification of updates to the Hub (GENERAL USE CASE 3). The sticky policy mechanism is used to protect data against undesired accesses;

- **Propagation of updates to PSPs**: the Hub, once it receives an update request it uses the Remote Data Registry Manager (associate to the data subject), to retrieve the list of PSPs the personal data has been disclosed to. It then sends the notification of update to the involved PSPs (GENERAL USE CASE 3);

- **Local update of information within the PSP and propagation of updates to other third parties**: as above. If the PSP has disclosed the personal data to other parties (consistent with the citizen’s preferences), then further notification and updates are sent to these parties.

10.2.2.6 Use Case 6

This use case is about a citizen that revokes consent for an IdP to hold their data at all.

It is a specific instance of use case 5, where the citizen changes their preferences and revokes the fact that the IdP can hold their data.

It can be locally handled within the IdP as described in the GENERAL USE CASE 3 which illustrates how the EnCoRe system handles consent changes locally: in this specific case the required action is either to delete personal data or disable any further access and usage of that data.
A variant of this use case might also involve the propagation of these changes to the other entities to whom citizen’s data has been disclosed. The mechanism and approach described for use case 5, above, applies here too.
References


[17] UK Cabinet Office, Identity Assurance (IdA) Programme Statements, [http://services.parliament.uk/hansard/Commons/ByDate/20110518/writtenministerialstatements/part003.html](http://services.parliament.uk/hansard/Commons/ByDate/20110518/writtenministerialstatements/part003.html)


[20] EnCoRe, Case Study 2: Biobank, 2011