The eDoc Building Blocks[†]

[†] This document is intended purely as a discussion paper exclusively for the internal, non-public use of the recipients. Although every effort has been made to ensure the accuracy and relevance of the contents, they do not in any way represent the official position or policy of the European Commission

TABLE OF CONTENTS

1. Context and Background

1.1. Ministerial declaration[‡]

The Ministerial declaration of 24 November 2005, approved unanimously in Manchester UK, focused on delivering clear social and economic benefits to citizens, businesses and governments, through 4 key challenges to governments:

1. No Citizen Left Behind
2. Efficient and Effective Government
3. Delivering High Impact services designed around customer's needs
4. Key enablers for eGovernment (*Simple and Secure access to online public services*)

The declaration identified the implementation and use of eID and the recognition of electronic documents and their electronic archiving as key enablers.

1.2. Signpost document, Building Blocks, Roadmap

The document "Signposts towards eGovernment 2010"[§] published in 2005 by the European Commission Information Society and Media Directorate General, discussed in more detail a number of objectives devised to improve Europe's social and economic development.

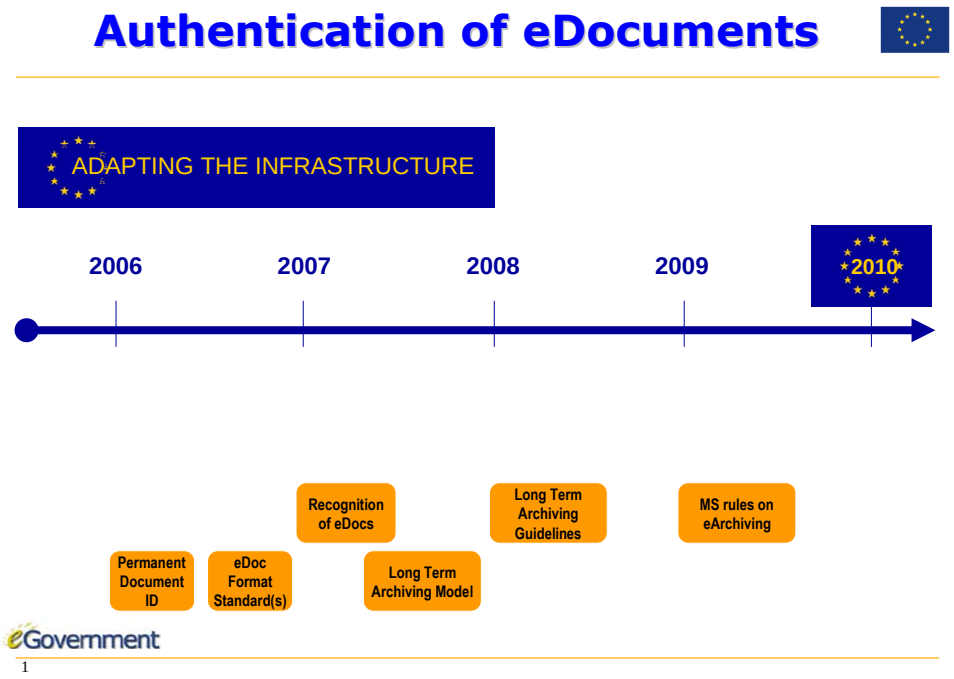
The document points out that as online transactions become more widespread, so will the uses of electronic documents. In particular, it sets out the goal that by 2010 Member States will have agreed a framework for reference to, and use of, authenticated electronic documents across the EU. The Signpost document represents a first step in establishing a roadmap for the achievement of this goal.

[‡] Available at: http://archive.cabinetoffice.gov.uk/egov2005conference/documents/proceedings/pdf/051124declaration.pdf

[§] Available at: http://europa.eu.int/information_society/activities/egovernment_research/doc/minconf2005/signposts2005.pdf

The document identified a number of "Building Blocks" laid out in time for efforts related to eID (electronic Identity) and "electronic document authentication" (a.k.a. "eDoc") efforts, which represent areas of activity which need to be conducted during the period leading up to the 2010 target to achieve the desired objectives. The elaboration of these Building Blocks should lead to the establishment of the roadmap for implementation of eDoc.

The context for the use of electronic documents is the delivery of interoperable eServices by EU Member State administrations under the eGovernment program.



As indicated via the timeline above, the essentials identified via the "Building Blocks" for electronic documents include:

- Permanent identifiers for electronic documents so that they are uniquely identified and identifiable
- Permanence of the electronic documents beyond any specific technology, medium, or platform (format standardization)
- Recognition of authentic electronic documents produced in one EU Member State as being authentic in any other EU Member State
- Long-term storage and archiving of electronic documents across the EU according to an established model & guidelines, and uniform rules

1.3. The draft CIP workprogramme and the eDoc Pilot

The draft CIP workprogramme[**] directly addresses mutual recognition and interoperability of electronic documents as follows:

> *Mutual recognition and interoperability of electronic documents is a pre-requisite and key enabler for many eGovernment services. This will require policies, practices and standards on electronic document format, to establish how electronic documents are identified, authenticated and accessible, and also long term archived. An agreed Framework for electronic documents should ensure permanence beyond any specific technology, medium or platform and shall guarantee availability and allow users to identify which representations of any document are considered authentic by a Member State or associated country and recognised as such in another. Pilot actions addressing this goal should deliver and test:*
>
> - *an agreed framework for reference to, and use of, authenticated electronic documents across the EU. Such documents and the infrastructure supporting them shall be able to include text, picture, audio, and video content;*
> - *electronic archives being able to store documents in acceptable formats for as long as is necessary to fulfil specific legal or cultural obligations;*
> - *eServices being able to interoperate across the EU, through identifiable and authenticated official electronic documents;*
> - *openly available Common specifications, for interoperability of electronic documents.*
>
> *The common specifications developed by the pilot shall be publicly available for all Member States or associated countries. Entities responsible for the definition of national standards and systems for electronic documents are expected to exploit the results of the pilot in view of work towards an EU-wide common specification*

1.4. The Ultimate Objectives

The objectives laid out in the above can become a reality by 2010 if the Member States collectively establish a document recognition framework considered as acceptable throughout the EU, which also takes into consideration technologies that will enable the creation, use and long-term storage of authenticated representations of documents.

[**] Available at: http://ec.europa.eu/information_society/activities/ict_psp/index_en.htm

This framework is/can be formulated as follows:

1. An EU agreement on electronic document formats (which may include elements of all media types such as: text, picture, audio, and video), recognized by all administrations as equivalent to their physical representations (paper, cassette, etc.)

    - Identifiable and addressable official documents ensuring that they can be authenticated and assert their official and reliable nature.

    - Authenticated electronic documents. This includes interfaces and methods for electronic identification and authentication, in particular the data, security, encryption, authentication and policy issues relating to the proposed eID framework;

    - due attention will be paid to existing standards in the relevant domains, for example from ISO/IEC, CEN, W3C and OASIS

2. An EU agreement on formats and methods deemed acceptable for long-term storage and archiving, whether signed or not.

3. A common understanding within the EU regarding permanent infrastructure to be made available for long-term document storage and authentication.

2. Detailed Description of the Building Blocks

2.1. Introduction

Until now it has not been set out in much detail what is behind each of these eDoc building blocks. For each of them, it would be useful to have a more detailed view of the different activities that will take place, the issues that are to be tackled, and what are the expected outputs. The sections below set out some thoughts on such potential details. They are intended as a first attempt to explore the Building Blocks in more detail, and to serve as a stimulus to further reflection.

2.2. Permanent Document ID

This building block is concerned with the issues surrounding the identification of electronic documents. More precisely it concerns the establishment of a framework for Permanent ID's (PID), so that electronic documents can be uniquely identified and accessed when needed. The framework would be developed via a consultative process.

Among the principal uses of this framework is the filing and retrieval of these electronic documents. As there is a considerable amount of information and knowledge available about large-scale filing systems (e.g., the Dewey Decimal system for libraries, etc.), and about large scale document repositories (and associated search/retrieval systems), the possibilities for re-use must be seriously considered/studied. How should this information/knowledge be taken into account/factored into the discussion/design of the electronic system?

A number of issues are raised within this building block, such as:

- Who assigns the PID's? What model is to be used? A single assigner or decentralized assignment? What about Namespaces? how to split the multiple namespaces into different parts?
- Is there to be a single registry of documents (or registries)?
- Are there to be multiple registries for different document types
- Is there a difference between "current" documents and "older" ("archive(d)") documents, and if so are they to be handled (stored? searched? etc.) differently (and how so)?
- How are "disposable" documents to be dealt with (documents that while signed have limited interest beyond their immediate intended use?

- How are the documents to be accessed using the ID?
- What are the links with eID (authors, publishers, signers, etc.)
- Could the ID be a unique combination of (some subset of) the electronic document's attributes? Typical attributes could include:
    - Archived bit
    - Where stored/where produced
    - When produced
    - When published
    - Version
    - Format type/spec
    - Sector/Domain type/spec
    - Serial number
    - Title
    - Other???

The conceptual model of eDoc as it relates to the ID could be conceived of as constituted of 3 interlinking parts:

- The eDoc contents themselves
- The Document ID
- The set of structured/standardized attributes associated with the eDoc

The framework resulting from these activities at a minimum should address the issues raised above.

2.3. eDoc Format Standard(s)

This building block is actually comprised of 2 sub-building blocks:

- consultative activities to develop explicit requirements (for interoperability)
- consultative activities to converge towards interoperable implementations of the conceptual model for electronic documents (selection of standards and options they present in line with the interoperability requirements)

The expected outputs would be common specifications in these areas, such as explicit detailed standards for the use of electronic signatures in different contexts; one possible element of the specifications could be a conceptual model of eDoc, providing for a single "virtual" eDoc (see below on the "nature" of electronic documents).

Issues that could be raised include:

- What is the nature of electronic documents in the new scheme? Could they be self-describing? Could they link revised versions of the same document together? Could an eDoc contain differently formatted versions of the same document (renderings), etc.

- Should a standard format for all documents be set as a convergence/long term target? e.g., XML format for birth certificates, tax declarations, university diplomas, work contracts, attestations of no criminal record, attestations of no bankruptcy, etc.

- What about the migration of document formats or format versions (e.g., if one day the PDF or MS word format, or some versions thereof were to become obsolete or unsupported)

- Are there any common selections with respect to data elements contained in specific documents which will require a common agreement, as prerequisite to achieving feasible interoperability of electronic documents? (e.g., in order to have interoperability of birth certificates, will it be necessary to specify lowest common denominator of common data elements any birth certificate must contain?)

2.4. Recognition of eDoc's

Mutual recognition does not mean the same technology/standards at both ends. However, it does imply the establishment of trust between administrations and government agencies located in different member states. In order to establish this trust and the necessary levels of interoperability, minimal standardization does have to occur. Mutual recognition has to be implemented in a non-discriminatory manner with regard to the MS origin of documents. Much of the activity associated with this building block concerns standardization and harmonization type activities, study activities and consultations between member states in specific forums.

Subjects to be addressed/studied via consultative processes include:

- What is the general target of the activity in functional/security terms, i.e., duplication of the existing system into the digital world, or is some improvement sought; if improvement is sought, what level of security should the new system provide? should there be multiple

levels of security related to documents (concerning multiple ways in which they could be signed), and if so what levels? a consultative process needs to be followed to build a consensus.

- Qualification of signatures: differences in the approaches used/possible need to be resolved to reach equivalence
- What adaptations and changes do administrations and government agencies have to undergo to implement the requirements for mutual recognition (organization, staffing, training, infrastructure, etc.)
- What is the lowest common denominator in terms of data requirements for acceptance of documents of equivalent types (e.g., for a birth certificate, what specific data fields should signed electronic documents produced by any member state contain?)
- What are the legal, procedural, linguistic, organizational and even technological barriers to mutual recognition

Expected outputs include:

- Description/Specification of legal changes required to eliminate barriers, including possibly new directive(s)
- Description/Specification of necessary changes to administrations, including changes in business processes, organization, etc., required to implement mutual recognition
- Definition of the minimum data sets/content required for specific documents (birth certificates, diplomas, etc.)
- Definition of a common framework for authenticating documents and ID's in order to achieve trust

2.5. Archiving

These building blocks are concerned with the infrastructure, technology and processes necessary to keep documents available and accessible permanently, even long past the timescale of their initial intended purpose.

Issues raised include:

- Authenticity is something that will have to be maintained, as conditions change; for example,
- As for the PID BB, one central archive, distributed archives? How linked?

- What kind of access controls in place? How to safeguard privacy and provide robust data protection?
- How do electronic documents end up in the archive? What about the trusted path from on-line repositories to the archive(s)?

2.5.1. Long-Term Archiving Model

The model should address all the open questions with regard to the archiving

2.5.2. Long-Term Archiving Guidelines

Among other things, this also concerns the elaboration of the obligations of the MS with regard to citizens/businesses, and to the other MS, e.g.,

- Maintenance of the authentication strength (including signatures, documents formats, format versions, etc.)
- Shouldn't lose documents
- Use of META for filing so that once filed/archived documents can be efficiently located/recovered when needed
- Access controls mandated
- Requirements with respect to revision of formats (obsolescence problem)

Also seeks to give guidance to the MS who may already begin building archives as part of their eDoc infrastructure and reorganizations, etc.

2.5.3. MS Rules on eArchiving

A new directive could be foreseen concerning archiving, and imposing the uniform rules and obligations; this will take some time to formulate and agree