

Consortium for studying, evaluating, and supporting the introduction of Open Source software and Open Data Standards in the Public Administration

Project acronym: COSPA



Work Package 5

**Definition of a target ODS to use in the project;
development of bridges from existing documents to
ODS, also using existing tools.**

Deliverable 5.1

Definition of the target ODS

Contract no.: IST-2002-2164



Project funded by the European Commission under the
“SIXTH FRAMEWORK PROGRAMME”

Project Acronym	COSPA
Project full title	A Consortium for studying, evaluating, and supporting the introduction of Open Source software and Open Data Standards in the Public Administration
Contract number	IST-2002-2164
Deliverable	5.1
Due date	02/06/2005
Release date	24/03/2006
Short description	The deliverable 5.1 describes the requirements to a target Open Data Standard and the methodology of how it can be selected from a list of competing Open Data Standards. This method enables public administrations to define their Open Data Standard for a certain type of application (i.e. word-processing,, Calculation, Email) based on collected requirements for that standard.
Author(s)	IBM.
Contributor(s)	Free University of Bolzano/Bozen
Project Officer	Tiziana Arcarese Trond Arne Undheim European Commission Directorate-General Information Society Directorate C - Unit C6- eGovernment, BU 31 7/87 rue de la Loi 200 - B-1049 Brussels - Belgium

Disclaimer

The views expressed in this document are purely those of the writers and may not, in any circumstances, be interpreted as stating an official position of the European Commission.

Neither the European Commission nor any person acting on its behalf is responsible for the use which might be made of the information contained in this document.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission.

Copyright

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/2.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Table of Contents

Deliverables Navigator.....	7
Executive Summary.....	8
1 Introduction.....	9
1.1 Open Data Standard.....	9
1.2 Scope of this Document.....	11
1.3 Structure of the Document.....	11
2 Migration issues.....	12
2.1 Prerequisites.....	12
2.2 Approaches and Strategies.....	13
2.3 Importance of interoperability in eGovernment.....	13
3 Method.....	15
3.1 Research Phase:.....	18
3.1.1 Comprehensive ODS research.....	18
3.1.2 Research into functional requirements	18
3.1.3 Research into non-functional requirements.....	19
3.1.4 Migration strategy.....	19
3.2 Develop Phase.....	20
3.2.1 ODS quality check.....	20
3.2.2 Workshop on further ODS research.....	21
3.2.3 Requirements Quality Check.....	21
3.2.4 Workshop on requirements research.....	22
3.2.5 Categorising requirement priorities.....	22
3.2.6 Filter by essential requirements.....	22
3.2.7 Procedure selection.....	23
3.3 Review Phase:.....	23
3.3.1 Procedure Review.....	23
3.3.2 Quality Check.....	24
3.3.3 Refinement Workshop.....	24
3.4 Usage phase:.....	24
3.4.1 ODS selection.....	24
3.4.2 Applying the result.....	24
4 Using the method in COSPA.....	26
4.1 Research Phase.....	26
4.2 Develop Phase.....	28
4.3 Usage Phase.....	32
4.4 Proposal: How to research requirement-matching for ODS (and OSS).....	32
5 Results.....	34
5.1 The (OASIS Open Office-compliant) OpenOffice.org File Format.....	34
5.1.1 Guidelines.....	34
5.1.2 Technical Details about the OpenOffice.org Format.....	35
5.2 The OpenDocument format.....	37
5.2.1 Technical details.....	37
5.2.2 Extensions supported.....	37
5.3 A final overview.....	38

6 Conclusions.....	40
Appendix A - Abbreviations.....	41
Appendix B - Glossary.....	42
References.....	43

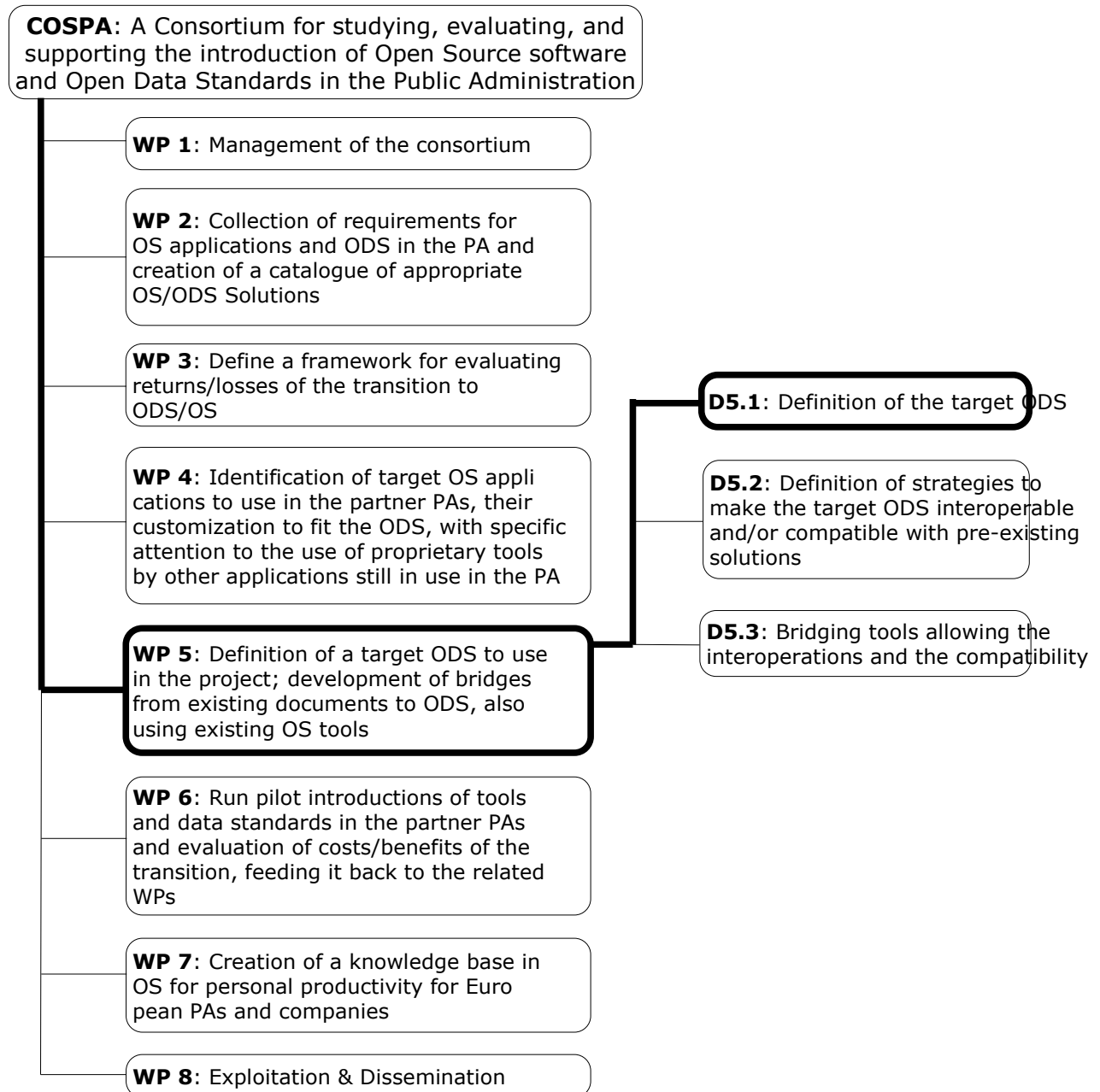
Index of Tables

Table 1 Migration strategies and examples.....	13
Table 2 Impact of migration strategy.....	20
Table 3 Word processing software standards	26
Table 4 Word processing, data standard requirements.....	28
Table 5 Word processing, data standard requirements and categories.....	29
Table 6 Available document formats compared with PA's essential requirements.....	30
Table 7 Standards compared with PAs' conditional requirements.....	31
Table 8 Sample OpenOffice.org manifest.xml file.....	36
Table 9 Different extensions available for the OpenDocument data format.....	38

Index of Figures

Figure 1 Method overview.....	16
Figure 2 Structure of the OpenOffice.org file format.....	35

Deliverables Navigator



Executive Summary

The COSPA project provides fact-finding, analysis, knowledge and possible implementation scenarios for the adoption of Open Source Software (OSS) solutions and Open Data Standards (ODS) in EU Public Administrations (PAs). The overall objective is to provide potential OSS migrants with actionable knowledge that will simplify and streamline any decision process about making the transition to OSS. Drawing on fieldwork and an analysis of the literature, COSPA benchmarks the effectiveness of deployed OSS solutions, and compares and pools knowledge from an analysis of user requirements. The entire project findings are collected in a dynamic knowledge base available to PAs and citizens seeking an objective, in-depth account of OS parameters, options, experiences, costs, barriers, and opportunities.

This deliverable presents a method for selecting an Open Data Standard for a specific application type in a PA undertaking a software migration. The method is designed to enable PA managers to find a suitable ODS from a corpus of standards for their specific process requirements. It will enable different requirements, analysis procedures, and contexts to specify the suitability of different Open Data Standards. The main advantages of implementing a new ODS are interoperability, flexibility, quality and vendor independence.

The method takes as input a set of requirements from the PAs, matches these requirements against a list of ODSs and selects as output the ODS best suited to the PA. The method is based on concepts in the existing literature and similar approaches to solving various problems in Information Technology (e.g. Software Quality Models). The ODS in question will have an impact on the PA's migration strategy, future development and asset plans. It will largely determine the software acquisition process of since new software will have to support the standard in question.

The method is illustrated by an application within the COSPA consortium, using data found elsewhere in the project deliverables – namely, how PAs might choose an Open Standard for the word-processing software used for basic business practices in their organisations. The results of this application in the COSPA context offer a good example of how the method can be used. However, the method is first presented independently of this application so that PAs outside COSPA can use it to simplify their own selection of Open Standards.

1 Introduction

1.1 Open Data Standard

The key term used in this deliverable is Open Data Standard (ODS). An IT standard (whether OSS or not) is created to enable and facilitate the exchange of documents between different programs, program versions, operating systems and hardware platforms.

The Open Standards concept is based on people working together openly to develop solutions that address common requirements and goals in a collaborative environment. Open Standards are usually developed in committees (open standardization bodies) such as the Internet Engineering Task Force (IETF) or the World Wide Web Consortium (W3C). An Open Standard can help to ensure that different software solutions and their everyday business process implementation requiring communication and interconnectivity can interoperate and integrate.

Using a standard such as the Internet TCP/IP communications protocol, for example, almost any component on a network can talk to any other component, creating an infrastructure for collaborating and coordinating resources across the globe.

Open Standards help [1]:

- encourage repeatable processes
- optimise communication
- increase acceptance of products and services

These enable:

- interoperability
- simpler integration
- faster integration

Which in turn, lead to:

- more available resources
- greater return on investment
- more options
- improved vendor independence
- broader vendor choice
- lower vendor cost
- quality
- flexibility
- reduced risk.

Of course Open Standards are not the only way to achieve standardisation. While Open Standards exhibit openness during the collaborative effort to create the standard, and in offering access to its specifications and technology when implementing those specifications, there are nevertheless other types of standards such as “industry de facto” standards. Most “de facto” standards require the adoption of proprietary technology and may require a licence fee to be paid to the suppliers of that technology.[1]

The use of the term “Open Data Standard” in this document follows the definition of the European Commission published in the Final Version (November 2004) of the European Interoperability Framework. This definition mirrors the definition found in the much-cited "Open Standards: Principles and Practice," by Bruce Perens. The IDA¹ definition is consistent with the draft "Principles for Open Source and Open Standards" formulated for the Open Standards Alliance. For further information see the Open Standards Alliance Website at <http://www.openstandardsalliance.org>.

Most importantly, Open Standards will be required to ensure **interoperability** for pan-European eGovernment services. The following check-list of minimal characteristics shows what is required for a specification and its associated documents to be considered an Open Standard:

- The standard is adopted and will be maintained by a not-for-profit organisation, and its ongoing development occurs on the basis of an open decision-making procedure accessible to all interested parties (consensus or majority decision etc.).
- The standard has been published and the standard specification document is available either freely or at a nominal charge. Permission must be granted to everyone to copy, distribute and use it for free of charge or at a nominal fee.
- The intellectual property - i.e. any patents that may be involved - of (parts of) the standard is made irrevocably available on a royalty-free basis.
- There are no constraints on the re-use of the standard.

.Standards which conform to these core requirements will facilitate the interoperability of services and systems at a pan-European level, and therefore enable PAs to provide their own specific user-centric e-services by facilitating the.[2]

In addition to interoperability, PAs also face **archiving issues** that need to be addressed by Open standards. Rapid changes in document formats do not always support archiving, which means that PAs may have problems opening such documents in the future since features are no longer supported and filters will not work. An Open Standard should help in addressing those requirements.

It should also be noted that there is no single definition of an Open Standard. A collection of references to published definitions and formulations for “Open Standards” is available on the Cover Pages Website at <http://xml.coverpages.org/openStandards.html>.

¹ Interchange of Data between Administrations, see Glossary

1.2 Scope of this Document

The main objective of this deliverable is to introduce a method for objectively choosing a target data standard when anticipating or implementing a software migration. This method will be demonstrated by applying it data from the COSPA project on relevant word-processing file formats.

This method aims to show PAs how to find a target Open Data Standard, but it cannot make suggestions about the right Open Data Standard for every possible application in use across all European PAs, since this would involve the evaluation of a vast amount of data. But the method can be adopted by PAs to evaluate their own most suitable standard for a given application type.

The prioritised list of requirements that are used to select the target Open Data Standard also depends closely on the specific context of the migration. In the case of this deliverable, the Open Data Standard eventually selected by the method for the PAs is COSPA-specific. For more information on how to prioritise PA requirements on the basis of an analysis of their user processes, see deliverable D2.8/D2.9.

1.3 Structure of the Document

The first part addresses migration issues. This covers such preconditions as the actual desire to migrate, as well as the need for an eventual migration. It discusses migration approaches and strategies and describes the benefits of a migration to Open Data Standards, underlining the importance of interoperability in the eGovernment context.

The second major section focuses on the method for selecting an Open Data Standard to be used in the migration process. This method consists of four phases - Research, Develop, Review and Usage – and can be applied whenever a new standard is desired independently of the type of application. Each of these four phases is explained in depth. The method naturally has to be applied to each type of application, since there is no single standard that covers all applications.

In the third section, the method is exemplified by showing how it can be used as a decision-support tool for COSPA's own PAs when selecting an ODS in the context of word-processing, the most widespread IT application in use in these organisations. The corpus of data used for the method is taken from project deliverables D2.4/D2.5 and D2.8/D2.9. The results of this exercise (the selection of OASIS Open Office as the appropriate open standard in this specific context) are fully presented and analysed.

2 Migration issues

This chapter addresses migration issues which directly impact the Open Data Standard selection process. Two approaches and their relative impact on the target selection method are explained. The reasons why interoperability and Open Standards are especially important for eGovernment are also discussed.

2.1 Prerequisites

The aim of this deliverable is not to provide a complete migration plan. However, any migration process will impact the Open Data Standard selection process in one of two key ways (explained below), so a number of factors need to be brought into the discussion from the start:

- Dependency on application type
- Training effort/costs (when a new software is to be used)
- Existing know-how
- Hardware dependencies
- User acceptance
- Functional restrictions at the coexistence of CSS and OSS products
- Parallel operation issues
 - Duplicate user management
 - More hardware
 - Administering and taking care of data on two different systems
 - Extended know-how from administrators
- Concepts may not be adoptable one-to-one. (security concepts and access-rights)
- Software dependencies (house standards)
- Obtaining old applications (several clients or emulations) necessary because migration may be inefficient.

Before migrating, it is also important to answer the following questions:

- What is technically feasible and/or where are the key pain points?
- How can known problems be resolved or by-passed ?
- What is important from a technical point of view when it comes to migrating a component?
- Which functionalities can continue to be used after migration, and/or what restrictions will this lead to?

2.2 Approaches and Strategies

Migration strategy plays a key role in selecting the right Open Data Standard. There are two major approaches to undertaking a migration.

The first is a 'hard' or 'complete migration'. In this approach, the complete migration will be carried out at a fixed date, and everything will shift over to the new solution. In such cases, there may be no need for a long-term bridging tool, since all the old applications and/or formats will disappear from use. Converters can then be used to transform old formats / standards into the new one.

The second approach is to carry out a 'soft' or 'partial' migration. This ensures that old applications and formats still exist and continue to be used. The migration will be carried out piecemeal, one department or selected client and/or server at a time. This means that old and new Open Data Standards and possibly applications will co-exist in the workplace. If the old software does not support the new standard, PAs may have to choose new software for the type of application in question or use a transparent bridge.

While a complete migration replaces all old (proprietary) standards and their dependent software, there is a problem of non-compliant standards in the case of a partial migration. Bridges and tools have to be developed to ensure the coexistence of old and new standards. Within COSPA, most PAs have chosen the soft migration approach to avoid the higher cost of a hard migration.

Experiences from other projects suggest that a hard (transition-free) migration requires intensive high-level technical resources over a short period of time, since hardware and software have to be replaced prematurely with new products, and standard products have to be replaced by self-developed software since they are unavailable on other platforms.

<i>Migration Strategies</i>	<i>Example</i>
Data converters/bridges for office files	Sun's StarOffice Migration Kit (current used formats to OpenDocument format)
Standard interchange formats	XML
Standard interchange format for office applications	OASIS Open Document Format
Common features in a RDBMS	Translation tools like ora2pg (Oracle PL/SQL to PostgreSQL) and inf2pg (Informix SQL to PostgreSQL)
Working with pre-existing solutions	Development of tools and bridges

Table 1 Migration strategies and examples

2.3 Importance of interoperability in eGovernment

When speaking of “interoperability”, the prime focus is on *technical* interoperability. This involves the technical issues of linking computer systems and services together, and includes such key features as open interfaces, interconnection services, data integration and

middleware, data presentation and exchange, accessibility and security services.[2]

At the most general level, interoperability simply means “making things work together”, the ability to combine information and functionality together from different systems. This requires a common framework for the information to be shared, which implies that we know what features systems actually share.

Not all information from PAs needs to be shared. For example, internal administrative business processes which use proprietary formats or self-developed applications will never leave the PA's intranet. Other public administrations in Europe will not need this type of information so there is arguably no need for an interoperable format for such information or for the software that processes it. However, depending on the size of the PA, there may be far more information that needs to be shared than doesn't. Interviews with PA partners show that larger-sized PAs share far more information than smaller PAs.

The EC IDA European Interoperability Framework has underlined the benefits of open standards. The EC emphasizes the strong link between Open Source software (OSS) and the application of open standards: "OSS² products are, by their very nature, publicly available specifications, and the availability of their source code promotes open, democratic debate around the specifications, making them both more robust and interoperable. As such, OSS corresponds to the objectives of this Framework and should be assessed and considered favourably alongside proprietary alternatives." [2]

For more information on interoperability please consult deliverable D2.7, “Detailed description of accessibility and interoperability requirements”.

2 Open Source Software, see Glossary

3 Method

This method has been developed to enable PAs to find the relevant Open Data Standard for a specific application type. This will be called the 'target' standard. A graphical representation of the method is shown in Figure 1.

The **Research** phase covers the investigation of Open Data Standards, together with the functional requirements and non-functional requirements associated with such standards. If there is a migration strategy on the table, it can impact several aspects of the process. A checklist of essential requirements is used to filter out possible Open Data Standards, leaving those candidate standards that enter the final evaluation process. In this project, this requirements checklist comes from COSPA Work Package 2.

The **Develop** phase covers all tasks required in quality-checking and filtering potential Open Data Standards. These include:

- ensuring the collected Open Data Standard list is complete,
- quality-checking the list of **functional** and **non-functional** requirements,
- categorising these requirements into **essential**, **optional** and **conditional** requirements, and
- verifying that the selected Open Data Standard satisfies all (especially the essential) requirements(. (See chapter 3.2)

Finally a selection process is used to find a suitable procedure for determining the target Open Data Standard that best matches the requirements, using a list of all qualifying standards.

The **Review** phase is used to make a final check on whether the method can produce the desired output.

The **Usage** phase refers to final selection procedure. This phase outputs a table of Open Data Standards and their corresponding scores when evaluated against the optional requirements. The Open Data Standard with the highest score is the most appropriate standard for the PA's requirements. This standard will usually be deployed during the transition, and implemented in line with the PA's overall migration strategy.

The method therefore delivers a single target Open Data Standard, which means that PAs will have to repeat the method for each type of application for which they wish to select a new Open Data Standard. Typical application types and their corresponding software which will have to undergo an ODS evaluation using the method include:

- word-processing, (Microsoft Word, OpenOffice Writer, K-Office)
- Calculations / Spread-Sheets (Microsoft Excel, OpenOffice Calc,...)
- Graphic manipulations (Microsoft Paint, Adobe Photoshop, GIMP, OpenOffice Draw).

3.1 Research Phase:

During this phase, a collection is made of all Open Data Standards capable of replacing the current data standard and ensure improved usage, together with all requirements that accompany the selection of this ODS. The outcome of this phase is a consolidated list of all available Open Data Standards which could be targeted by the migration, plus a list of functional and non-functional requirements that this Open Data Standard has to meet. These requirements cover essential and optional demands.

3.1.1 Comprehensive ODS research

Comprehensive research ensures that all possible and available Open Data Standards are taken into consideration. The first source of research is the application being used currently and its corresponding data standard. All other standards handled by this application also needs to be checked.

In the case of word-processing, applications, this can be carried out by using the “Save as” or the “Export as” feature in the specific software application to reveal a list of available file formats . It may also be useful to check other competing software applications from other vendors, since they may be able to handle even more formats. All these standards (or file formats) can be used in the final evaluation. The Develop phase will then verify whether the standard in question meets all requirements or not.

Another good starting point for seeking out industry or open standards is standardisation bodies. They have websites covering all types of Open Standards. For example the W3 consortium offers a lot of information about XML standards.

3.1.2 Research into functional requirements

Functional requirements describe which functions an Open Data Standard should be capable of performing. Because of this they are also called 'behavioural requirements'. A functional requirement is either met or not met by a given standard.

There are two major approaches: a **requirements inquiry**, or a **requirements inspection**. A requirements inquiry starts “from the native point of view”, obtaining information about requirements by:

- talking to users,
- observing them using the system in a real working situation,
- asking questions verbally or in written form, and/or
- collecting requirements by analysing business processes

Requirement inquiry techniques include the use of ethnographic studies, interviews, workshops, and surveys.

A requirement inspection focuses on observing functional requirements *in situ*, and includes heuristic evaluation and cognitive walk-throughs.

Functional requirements include:

- keeping document content and layout information separate
- small file size
- readable by humans
- transformation-friendly

3.1.3 Research into non-functional requirements

Non-functional requirements tend to be features that are measurable. They are also known as non-behavioural since they address an attribute or quality of the Open Data Standard. Non-functional requirements are divided into three categories: social, economic and political.

To take political requirements, these cover such objects as national and/or international legal terms and conditions that need to be factored in when choosing a standard. Since different countries have their specific laws, documents and strategies, these have to be examined as part of the research process. They mostly aim at ensuring interoperability, re-usability, openness, and cost and risk reduction, as well as cost transparency. Guidelines to researching non-functional requirements are given in chapter 4.4.

Non-functional requirements include:

- an Open Data Standard is controlled by a standardisation body
- it is vendor-independent
- it is royalty-free
- it uses a non-discriminatory license

3.1.4 Migration strategy

The chosen migration strategy can impact the selection of the the target Open Data Standard in two ways. Migration may influence the requirements or even create new ones; or inversely the selected Open Data Standard may affect different phases of an already-planned migration.

Migration only matters in the selection process if there is already a strategy, and Open Data Standards can be selected before any migration strategy is planned. However, strategic issues also need to be addressed. If you decide on an Open Data Standard before planning your migration, this can lead to added complexity and effort, depending on the currently used Open Data Standard.

For example replacing an old data standard with a new data standard without needing to change products usually involves a low level of migration effort. But switching both products and the data standard means developing, installing and using bridging tools or converters. Examples of these are given in Table 2:

Migration Strategy (Requirement)	Current used DS	Potential target ODS	Effort
Continuing Migration	MS Word 2000	Word-processing ML	low
Replacement Migration	AbiWord	Open Office XML	medium
Not yet defined	MS Word 2000	Open Office XML	high

Table 2 Impact of migration strategy

Continuing migration means that the PA's current product line continues to be used, whereas **replacement migration** means changing the whole application/platform. Table 2 shows that if you have no migration strategy, you have a larger choice of Open Data Standards to select from. This may however mean that the final Open Data Standard is not supported by the current software, and that both the application *and the* data standard now need to be changed. This is, of course, a much more complex undertaking.

In the final case, the migration strategy is actually driven by the target Open Data Standard. The PA has to migrate to a new application and standard, and can decide whether to carry out their replacement migration completely or partially.

3.2 Develop Phase

This phase usually comprises a variety of quality assurance checks. These ensure that the Open Data Standard meets the given functional requirements and non-functional requirements. The Develop phase checks the quality of the input from the Research phase and when necessary improves it using workshops and/or surveys.

Once all requirements have been quality-assured, they are categorised and Open Data Standards which do not meet the essential requirements are filtered out.

The outcome of this phase is a list of Open Data Standards which have passed all quality checks and may be used to substitute an existing standard wherever reasonable and feasible. All Open Data Standards which pass this phase meet all essential requirements and can enter the evaluation process.

Together with the Open Data Standard list and the optional and/or conditional requirements, a procedure is selected to evaluate the Open Data Standard in the target list that offers the best fit.

3.2.1 ODS quality check

Even though many Open Data Standards may have been collected in the Research phase, this quality check is designed to ensure absolute completeness. This means that all available standards are candidates for the selection procedure. The quality check also verifies that the collected standards can be used for the type of application for which the standard will be selected.

3.2.2 Workshop on further ODS research

To improve the quality of the input data, the the already available Open Data Standards list has to be refined. This can be done either in a workshop or by running a survey. A workshop is the preferred solution since since communication is more direct, making it easier to solve problems in understanding. The aim of the workshop should be to carry out new research into Open Data Standards that were missed in the first round. The workshop can be held virtually via the Internet (e.g. e-Workshop) or by phone.

3.2.3 Requirements Quality Check

This check determines whether the selected requirements meet the required quality standards (section 3.1.3) and whether they are appropriate for the selection task.

To be method-friendly, the list of (functional and non-functional) requirements needs to meet the following minimal quality standards:

- Accessibility of the data standard (e.g. legible, licenses, etc.).
- Completeness (e.g. no partially filled out questionnaires, or several different answers).
- Coherence (e.g. no contradictions within one questionnaire, level of detail in the data is sufficient,).
- Diversity (e.g. questionnaires filled out by different types of users, not different copies of a single answer,).
- Definiteness (e.g. each entry in the input refers only to a single requirement).
- Process coverage (i.e. all requirements imposed by the administrative business processes to be covered are listed),
- Quantity (e.g. too small set of requirements makes it irrelevant).

The quality check should be carried out by all parties. If the resulting requirements set is unsuitable, or if the requirements seem to be incomplete, action is taken to improve the quality of the merged input to obtain a better set. The cycle of selecting requirements, checking their quality and improving them where necessary is repeated until the quality of the requirements set is deemed sufficient.

Typical reasons for non-quality in a given list of requirements:

- all the requirements come from a single person and are therefore not representative.
- A special requirement has legal implications and cannot be used in the PA.
- The quantity of the collected requirements is too small.
- Only essential requirements have been gathered.
- Only non-essential requirements have been gathered.

3.2.4 Workshop on requirements research

Like the Open Data Standard research workshop, the requirements workshop aims to improve the quality of the list of requirements. As suggested in section 3.2.2, this can be done by:

- refining the available data
- holding workshops or running a survey
- holding e-Workshops and/or conferences

3.2.5 Categorising requirement priorities

Once the requirements have passed the quality check, they have to be prioritised into essential, conditional and optional requirements. All requirements - social, technical, economic and political requirements - are important and can be categorised in this way. They are all vital to the administrative business processes/tasks, and the Open Data Standard has to address them.

Priority is ranked on a scale of three categories. This scale is subjective and imprecise, so everyone involved has to agree on the meaning of each level on the scale they use. In the COSPA project, meanings were established as follows:

- **optional** – functions that may or may not be worthwhile
- **conditional** – enhances the standard, but the standard is not unacceptable if the requirements are absent
- **essential** – the standard is not acceptable unless these requirements are satisfied

For example there might be two PAs in different countries which need to share some type of data, but happen to use different types of software. An essential requirement for sharing that data is a fully described (open) data standard. Without a complete definition of that data standard, the PAs will be unable to develop the bridges that will allow them to share these types of data.

An optional requirement could be that the Open Data Standard must be self-describing. This would simplify things, but it is not absolutely necessary provided the standard itself is fully documented.

3.2.6 Filter by essential requirements

Since the future standard has to comply with all necessary requirements, these are used as a filter in the Develop phase to filter out all non-compliant Open Data Standards. If Research filters out such a non-compliant Open Data Standard, then any further engagement with that standard is stopped.

Those Open Data Standards which pass the filter enter as candidates into the final evaluation procedure. It is of course possible that no standard emerges from the filtering process. In this case, some of the essential requirements must be readjusted. Requirements which simply cannot be satisfied by any available Open Data Standard should be eliminated.

In the case of the COSPA project, four out of eight evaluated Open Data Standards met all essential requirements. Details of this analysis can be found at chapter 4.

3.2.7 Procedure selection

After an appropriate set of Open Data Standards and optional and/or conditional requirements have been discovered, the team involved now has to select the actual **selection procedure**. The procedure takes the optional and/or conditional requirements and the list of Open Data Standards as input, and outputs a target Open Data Standard. The development of this procedure is carried out by the people involved in of the planned migration.

To ensure that procedure is able to evaluate a target Open Data Standard, the given requirements are transformed into a model which can be used to process the list of Open Data Standards. This results in a rating for all Open Data Standards on the list.

Possible procedures range from a simple count of requirements met by an Open Data Standard right up to the use of elaborate computer programs which use existing formulae or techniques to rate the Open Data Standards. Different procedures can also be carried out consecutively, so that their outputs are merged into a single list. The range of possible procedures includes:

Ranking the Open Data Standard according to the number of requirements they fulfil or have a positive influence on.

A cost-benefit-analysis. This delivers a weighted averaging of the number of fulfilled requirements.

A questionnaire asking different people to evaluate whether a requirement is supported by an Open Data Standard or not, and then totalling the opinions.

Checking every Open Data Standard for each requirement separately and then merging the results into a single list, by either sorting on the basis of first, second, third places, or by giving ranking points for each Open Data Standard and then sorting by the sum of these points³.

As can be seen, there is a large range of possible procedures. The main task is to find and develop an appropriate procedure for the special case for which the selection process is used.

3.3 Review Phase:

The Review phase verifies that the selected procedure from section 3.2.7 is a suitable method for evaluating the target Open Data Standard.

3.3.1 Procedure Review

In this part of the method, it is possible to run another review of the developed procedure and

3 The first is used (at least in Germany) to generate a ranking of nations in the Olympic games. Countries are first sorted on the basis of the number of gold medals obtained, then by the number of silver medals and last by the number of bronze medals. The second is used in the Grand Prix d'Eurovision de la Chanson, where a first place within a participating nation is worth 10 points, the second 8, the third 7 and so on.

request changes so that the requirements-to-method mapping is as tight as possible.

Possible procedures for this review include:

- A presentation of the completed procedure, explaining how it works and seeking agreement on the approach.
- The procedure is applied to a set of requirements, and procedure parameters changed or fine-tuned until a previously defined ranking of requirements is obtained.

Once again, there are many ways of applying the method to adapt it to the specific situation the procedure has been developed for.

3.3.2 Quality Check

If in the course of the procedure review, the developed procedure is deemed unsuitable for the selection, the procedure will be either refined or reworked. When the suggested procedure is accepted, the procedure will be applied and the results evaluated.

3.3.3 Refinement Workshop

If the procedure is not accepted or there are additions or comments on the procedure, it has to be modified accordingly.

If there is a need for immediate feedback it may be helpful to arrange a workshop for relevant parties.

The procedure is then refined and reviewed iteratively until it is accepted.

3.4 Usage phase:

The aim of the **Usage** phase is to execute the chosen procedure. The outcome of this phase is a table of Open Data Standards and their corresponding “profit” values evaluated against the (possibly prioritised) optional requirements. Since one of these Open Data Standards will exhibit the highest profit value, it is most likely to become the target Open Data Standard.

3.4.1 ODS selection

The method is finalized by applying the developed procedure (section 3.2.7) to the input data. The result of that application is the target Open Data Standard to be used in the planned migration. Another output of the application of the method should be a set of documents describing the way the method has been applied and also its inputs and outputs. These documents can be used in reviews and for further applications of this method.

3.4.2 Applying the result

The target Open Data Standard has been defined and the migration can be planned, or if already under way, pursued. The final choice of Open Data Standard can also impact the

migration process. For example, in cases where the old software does not support this chosen Open Data Standard, migration primitives will need to be developed. And if the strategy allows the coexistence of old and new standards, bridges have to be developed.

4 Using the method in COSPA

On the evidence of the results from Deliverable D2.4/D2.5, most administrative business processes make use of either databases or office suite software together, together with their underlying standards. There is widespread use of word-processors to support these processes, both in terms of the total number of PAs, and the total range of administrative business processes using them. Processes claiming that a word-processor is used accounted for 112 out of a total list of 236 processes.

These administrative business processes need to share the same format for several reasons. One consistent format over all PAs enables vendor independence, increases vendor choice, decreases vendor costs, increases the return on investment, optimizes communication, increases available resources, is quicker and simpler to integrate, increases quality and flexibility, durability and, last but not least, reduces risks. A description and examples of how Open Standards help to achieve these benefits are available in the book *The Business Case for Open Standards* [1].

Since most administrative business processes use word-processing applications, it was decided to use the experience of word-processing among COSPA PA partners to exemplify the method.. This example should also help PAs not directly involved in COSPA to easily tailor the method to their own needs when selecting an Open Data Standard in other application spaces.

4.1 Research Phase

We began by listing data standards used in the most popular current word-processing software. At the time of writing, these included:

<i>Software</i>	<i>File Format</i>	<i>Document Format</i>	<i>Open</i>
Microsoft Word 2003 Microsoft Word 2002 and older versions	.xml / .doc	own implementations (XML and binary)	Y N
StarOffice Writer, OpenOffice.org Writer, KOffice KWord	.odt ⁴	OASIS OpenDocument	Y
AbiWord	.abw	own implementation (XML)	Y
LaTeX / TeX, LyX	.tex	own implementation	Y
TextMaker	.tmd	own implementation	N
WordPerfect	.wpd	own implementation	N
RagTime	.rdf	own implementation	N

Table 3 Word processing software standards

Table 3 form the basis of our work on defining the target Open Data Standard for use in PAs.

4 The OpenDocument format has replaced the Star Office XML schema.sxw.

To select an Open Data Standard for their word-processing-centric processes there must already be at least one Open Data Standard. Table 3 shows that six applications use an open format, and they will be evaluated.

If other Open Data Standards for administrative business processes are selected, it will also be helpful to check Open Data standardisation sites such as the Dutch OSOSS catalogue (<http://www.ososs.nl/index.jsp?page=191>) and other standardisation bodies (see <http://www.openstandards.net/>).

To ensure that the Open Data Standard meets PA needs, the second step was to collect requirements that point to an Open Data Standard. In the COSPA context, this took three rounds of data collection. The first and second rounds were carried out using questionnaires to PAs, and the results revealed communication problems that generated inadequate or insufficient data. For details on this, see deliverable D2.4/D2.5.

To produce more meaningful results, a new indirect strategy for collecting requirements was developed, based on collecting actual business processes rather than Open Data Standard requirements. This allowed partner PAs to provide information that was closer to their everyday usage in the workplace. The results were evaluated and subjected to detailed analysis.

The complete set of standards requirements taken from WP2 are shown in Table 4:

<i>Functional Requirements</i>
Friendly to transformations using XSLT or similar XML-based languages or tools
Small file size
Self describing
Human-readable
If XML based compatible with W3C name spaces in XML v1.0 specifications
Keeping content and layout information separate
Cross-platform
Standard can store meta data information
Extensible
Standard allows creation of user-defined data types
Support emerging requirements
Non-binary format
Modifiable
High fidelity
Support current word processor features
<i>Non-functional requirements</i>
Controlled by standardisation body
Conforms to national policies
Open format
Widely adopted
Royalty free
State-of-the art
Vendor independent
Standard can be reused by other applications

Table 4 Word processing, data standard requirements

4.2 Develop Phase

Checks were made over the requirements and Open Data Standard lists to ensure completeness and quality by circulating the data over the COSPA mailing-list. The output from D2.4/D2.5 was also extensively discussed in various internal workshops.

As shown in Table 4, the D2.4/D2.5 data was useful but did not contain many non-functional requirements. A proposal on how to gather this information was made to PAs and is included in 4.4.

The next step was to categorise the requirements according to the method explained in

section 3.2.5. In the COSPA case, there were two sources for these requirements. One was the set of requirements from D2.4/D2.5 not evaluated in D2.8/D2.9. These are already categorised requirements, and it is mandatory that the Open Data Standard fits them.

The second source is the conditional and/or optional prioritised requirements from D2.8/D2.9 as. These will be used in the next phase of the method if more than one Open Data Standard successfully passes the filter and matches the essential criteria.

The final categorised requirements collected from COSPA partner PAs are shown in Table 5:

Requirement	Category
Friendly to transformations using XSLT or other XML-based languages or tools	optional
State-of-the art	optional
Standard allows creation of user-defined data types	optional
Self describing	optional
Royalty free	conditional
Standard can store meta data information	conditional
Extensible	conditional
Support current word processor features	conditional
Keep content and layout information separate	conditional
High fidelity (preserve format, content and structure)	conditional
Controlled by standardisation body	conditional
Widely adopted	conditional
Open format	essential
Cross-platform	essential
Human-readable (Non-binary)	essential
Modifiable content	essential

Table 5 Word processing, data standard requirements and categories

The Open Data Standard list from Table 3 (Word processing software standards) is now filtered through the list of essential requirements.. A standard which does not meet one or more specific essential requirements is rejected from the list as a possible target Open Data Standard. Due to the demanding PA requirements in COSPA (i.e. open format) most of the currently used standards did not pass this phase.

The following table compares formats to requirements:

<i>Document Format</i>	<i>Modifiable content</i>	<i>Cross-platform</i>	<i>Open format</i>	<i>Human-readable</i>
OASIS OpenDocument	Y	Y	Y	Y
Abiword's implementation (XML)	Y	Y	Y	Y
(La)TeX implementation	Y	Y	Y	Y
Microsoft's implementation (XML)	Y	Y	Y	Y
TextMaker's implementation	Y	Y	N	N
WordPerfect's implementation	Y	Y	N	N
Microsoft's implementation (binary)	Y	N	N	N
Ragtime implementation	Y	N	N	N

Table 6 Available document formats compared with PA's essential requirements

Some of this information comes from a comparative assessment of open document formats [4, P. 47]. All other information comes from research on document formats.

The only Open Data Standards which satisfy all these essential requirements are:

- the OASIS standardised OpenDocument format (XML),
- Abiword's self-implementation (XML based),
- the (La)TeX format and
- Microsoft's XML reference schemas.

Important note: When referring to XML(-based) Open Data Standards it must be noted that even though they are XML based, typical document files are usually binary (ZIP) files containing a compressed version of the actual XML file format.

The next step involved examining the document formats closely to see if they complied with the conditional requirements, using the procedures suggested in Section 3.2.7. For the COSPA data, a simple count of requirements matched by an Open Data Standard was used. Standards are ranked according to the number of conditional requirements they meet, and the highest-ranked Open Data Standard that complies with the most requirements is the most likely candidate for the target Open Data Standard.

The resulting comparison of standards and conformity to conditional requirements is given in Table 7:

<i>Requirement</i>	<i>OASIS OpenDocument</i>	<i>Microsoft XML ref. schema</i>	<i>Abiword's impl.</i>	<i>La(TeX)</i>
Royalty free	Y	Y	Y	Y
Storable meta data information	Y	Y	Y	Y
Extensible	Y	Y	Y	Y
High fidelity	Y	Y	Y	Y
Support current word processor features	Y	Y	N	N
Keep content and layout separate	Y	Y	N	N
Widely adopted	Y	Y	N	N
Controlled by standardisation body	Y	N	N	N
Fulfilled conditional requirements	8/8	7/8	4/8	4/8

Table 7 Standards compared with PAs' conditional requirements

OASIS OpenDocument is the only Open Data Standard which complies with all of the conditional requirements and is therefore meets PA demands. Currently, OASIS OpenDocument is the only standard for office applications managed by a not-for-profit organisation and therefore the only standard that meets the Open Data Standard definition announced by the European Union. This is why the OASIS OpenDocument format is the Open Data Standard finally chosen for COSPA project office applications.

The OASIS technical committee's remit was to create an open file format specification for office applications[3]. Members of the technical committee include Adobe Systems, Sun Microsystems, IBM and others. During the standardisation process, the standard was renamed as "OASIS Open Office XML". A complete history of the creation process and activities around OASIS OpenDocument format can be found in the Groklaw article at <http://www.groklaw.net/article.php?story=20050130002908154>. At the time of writing, the standard has been submitted to ISO for official standardisation.

It is now the PA's responsibility to decide which software to use. OASIS' OpenDocument format is supported by various office applications including:

- OpenOffice.org (native)
- KOffice (native)
- Sun Microsystems StarWriter
- IBM Workplace (native)
- AbiWord
- WordPerfect
- Kingsoft PowerWord
- NeoShine Office (China Standard Software)

- Red Office 2.0 (Beijing RedFlag)
- TextMaker

4.3 Usage Phase

The Usage phase is the final step in the process. The Open Data Standard has been selected and now has to be implemented. Implementation depends on which standard is currently being used. If a PA already uses the OpenDocument data standard, no further step is necessary since it is already using the most appropriate standard according to COSPA findings for word-processing, applications.

If, on the other hand, the PA is working with another data format, it must make the transition, using any resources, tools and tutorials that help convert proprietary file formats to the OpenDocument format

OpenOffice.org supports most standards by default, uses a variety of efficient filters, and also offers compatibility. These filters are vital to migration. There is no need to manually convert every single document to the new OpenDocument format, since batch jobs and analysis tools are available. The documents also have to be analysed in terms of their complexity, and this analysis can be automated with migration toolkits.

The most challenging part is the migration of Visual Basic macros, which are often used in PAs to automate processes. OpenOffice.org supports StarOffice Basic which is largely comparable to Microsoft's Visual Basic. But without adaptation, these macros are not executable. Sun's StarOffice Migration Kit can automate the conversion of macros.

4.4 Proposal: How to research requirement-matching for ODS (and OSS)

The starting point is the European Commission IDA programme (see IDA's Website at <http://europa.eu.int/idabc/>). IDA's mission is to support the implementation of community policies and activities by co-ordinating the establishment of pan-European telematics networks between administrations. For PAs, policies defined by such projects may be a (possibly essential) non-functional requirement for a planned migration. For example the "Free/Libre/Open Source Software: Policy Support" (FLOSSpols) project was launched on March 1, 2004 and is working on three specific tracks: government policy towards Open Source, gender issues in Open Source, and the efficiency of Open Source as a system for collaborative problem-solving.

Through this project, the European Commission's definition of open standards is part of the final version 1.0 of the European Interoperability Framework, which also encourages a favourable approach to Open Source software.

This definition can be taken as an essential requirement when filtering data standards to select a new one. Open Data Standards which do not meet this definition should be eliminated from the final evaluation process.

As described in section 3.1.3, potential adopters should examine their national policies and/or requirements. National or regional authorities may already have evaluated OSS and defined

best practices or policies when migrating to OSS.

For example, Germany's Federal Government Co-ordination and Advisory Agency for IT in the Federal Administration (KBSt), located within the Federal Ministry of the Interior, provides "Standards and Architectures for eGovernment Applications 2.0" (SAGA).

The KBSt is an inter-ministerial agency of the Federal Government whose mission is to ensure that the Federal administration optimizes information technology for specific fields in organizational, economic and technical terms. The released document contains a concise account of standards, processes, methods and the results of state-of-the-art IT development for eGovernment applications.

SAGA divides Open Data Standards into three categories. Competing standards which are not covered by SAGA should only be used in extremis. The first category contains mandatory standards which must be observed and preferred. Second, recommended standards are tried-and-tested, but are not mandatory and/or do not represent the preferred solution, or their mandatory status still requires wider agreement. Third, there is a section on standards under observation. These are in line with planned developments but have not yet reached maturity or have not yet proven their value on the market. German PAs are asked to follow and implement this guideline and take these requirements into consideration.

Another project from the United Kingdom also addresses Open Data Standards policies: the eGovernment Interoperability Framework at <http://www.ogc.gov.uk/index.asp?docid=2190>. The key decision of this policy are as follows:

- the UK Government will consider OSS solutions alongside proprietary ones in IT procurements. Contracts will be awarded on a value-for-money basis.
- The UK Government will only use products for interoperability that support open standards and specifications in all future IT developments.
- The UK Government will seek to avoid lock-in to proprietary IT products and services.
- The UK Government will consider obtaining full rights to bespoke software code or customisations of COTS (Commercial Off The Shelf) software it procures wherever this achieves best value for money.
- The UK Government will explore further the possibilities of using OSS as the default exploitation route for Government funded R&D software.

In general, standards and architectures for eGovernment applications (like SAGA and e-GIF) are obligatory for applications used multilaterally and this also applies to IDA eGovernment standards.

There may be other recommendations and drafts for ensuring interoperability and accessibility in addition to those given above. However, the question of whether they have to be taken into consideration depends on both the size of a PA and its own interoperability and accessibility requirements. These need to be examined and discussed when collecting a list of requirements, since they have a direct impact on the Open Data Standard selection process.

5 Results

The target Open Data Standard selected for word-processing, among COSPA PAs is the OASIS OpenDocument format. The following gives a deeper look at the OpenOffice.org file format guidelines and technical details.

5.1 The (OASIS Open Office-compliant) OpenOffice.org File Format

The OpenOffice.org (OOo) file format is growing very popular in the office automation field. The idea behind the creation of the format is to give to users an open format that is interoperable and flexible, avoiding the need to rely on a single company and its proprietary formats.

Proprietary data formats cause problems when using files no longer supported by the format creator, since access to the specifications is usually unavailable and users will have serious problems accessing these files.

The OOo file format is released under the GNU license and is largely based on the XML specification.

5.1.1 Guidelines

There are several guiding principles for underlying this standard.[5]

- **Separation of Content, Layout, and Meta Information**
The actual content of the files is always separated from the layout and the general information about the document. This allows greater flexibility in editing a document.
- **Standards based.** The format is based on different, widely-used standards, such as XML and the ZIP format to reduce the size of documents.
- **Uniform Representation of Formatting and Layout Information**
As OOo enables the use of formatting embedded inside XML documents or formatting through styles, this representation must be uniform inside the generated XML files.
- **Structured Format**
All structured information inside a document is stored as XML structures, allowing easy access using standard XML tools.
- **Idealized Format**
The file format represents documents in an abstract way, not binding them to the application that generated them. This boosts system interoperability.
- **Common Format Across All Office Applications**
The same format is used for all office applications, making it easier and more flexible to generate and process OOo files.
- **Open For Extensions and Supplemental Information**
Attributes can be added to the format, while maintaining format compatibility.

5.1.2 Technical Details about the OpenOffice.org Format

Every file saved by OOo is a package containing different XML files. All files are packaged and compressed using the ZIP format. By using an archiving program that supports the ZIP format, it is possible to open such a file and see in detail the contents.

The typical content of an OOo file is the following:

- content.xml
- styles.xml
- meta.xml
- settings.xml
- META-INF/manifest.xml
- A *Picture* folder, where all the images embedded in the document are stored (this folder is only inserted if the document contains images).

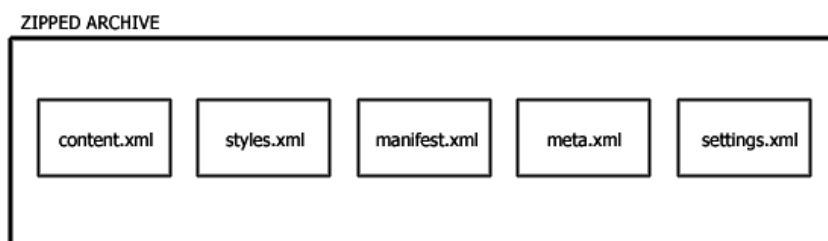


Figure 2 Structure of the OpenOffice.org file format

Manifest.xml contains a list of all files present in the archive. It is the central point of the archive structure, listing all other files present. The content shown in Table 8 can differ between different files, and some sections such as *Picture* can also never appear.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE manifest:manifest PUBLIC "-//OpenOffice.org//DTD
Manifest 1.0//EN" "Manifest.dtd">
<manifest:manifest
xmlns:manifest="http://openoffice.org/2001/manifest">
<manifest:file-entry manifest:media-
type="application/vnd.sun.xml.writer" manifest:full-path="/" />
<manifest:file-entry manifest:media-type="" manifest:full-
path="Pictures/" />
<manifest:file-entry manifest:media-type="application/binary"
manifest:full-path="layout-cache" />
<manifest:file-entry manifest:media-type="text/xml" manifest:full-
path="content.xml" />
<manifest:file-entry manifest:media-type="text/xml" manifest:full-
path="styles.xml" />
<manifest:file-entry manifest:media-type="text/xml" manifest:full-
path="meta.xml" />
<manifest:file-entry manifest:media-type="text/xml" manifest:full-
path="settings.xml" />
</manifest:manifest>
```

Table 8 Sample OpenOffice.org manifest.xml file

Meta.xml stores all information about the document, such as the number of pages of content, creation date, and language details.

Styles.xml contains all styles available for the current document, using a concept similar to Cascading Style Sheets (CSS), which is popular in web-development.

Settings.xml stores all settings for the OpenOffice user interface used in document creation.

Content.xml is the content of the document inserted in an XML file.

Each XML file follows a *Document Type Definition* (DTD) schema that allows easier interoperability. All DTDs used by Oo are stored inside the program's folder and are publicly accessible.

Four file extensions are available, to distinguish the different types of document:

- SXW to store Oo Writer documents.
- SXC to store spreadsheet documents
- SXI for presentations
- SXD for drawings

To encourage simplification, the structure of the file remains the same for all file extensions.

5.2 The OpenDocument format

OpenDocument is the new XML-based file format designed for the new OpenOffice.org 2.0. It is based on the version we have been describing, but is not compatible with OpenOffice 1.0 formats. This decision was taken to ensure that the format is not tied to a particular office application, and that the data contained in documents is not associated to a particular application and version. This meant losing compatibility with the old format to acquire greater flexibility.

The advantage is that two office automation suites, OpenOffice.org 2.0 and KWord 1.4 (part of KOffice the KDE Office suite) can support the same format. Given that any developer willing to use the format is free to do so, this can help disseminate the format.

5.2.1 Technical details

As in the case of the OpenOffice.org format, an OpenDocument file is a Zipped archive containing the following files:

- content.xml
- meta.xml
- settings.xml
- styles.xml
- mimetype
- layout-cache

Different folders may also be present or not depending on the specific document. For example, the Pictures folder contains the pictures that have been inserted in the document and that will be inserted as a link in the content.xml file.

- Thumbnails
- Pictures
- Configurations2
- META-INF

The specifications of the format are freely available for all developers willing to adopt it [8]. An overview of all the details of the format and the technical details of the implementation is available on-line [7].

5.2.2 Extensions supported

The main extensions for the format are

- ODT to store text documents
- ODS to store spreadsheets documents
- ODP for presentations

More extensions are available, a summary can be seen in table 9, where the document templates are listed.

Extension	File type
ODT	Text
ODS	Spreadsheet
ODP	Presentation
ODG	Drawing
ODC	Chart
ODF	Formula
ODB	Database
ODI	Image
ODM	Master Document
OTT	Text Template
OTS	Spreadsheet Template
OTP	Presentation Template
OTG	Drawing Template

Table 9 Different extensions available for the OpenDocument data format

Once again, new extensions to the format had to be established to avoid conflicts with older extensions.

5.3 A final overview

The Organization for the Advancement of Structured Information Standards (OASIS) has approved an open format that covers the requirements of word-processing, spreadsheets, charts and graphical applications.

We have already mentioned the OpenOffice.org file format created in 2002 to produce a standardised format for office automation applications. This format has been rendered even more abstract, and in 2005 was renamed as OpenDocument, to emphasise the interoperability of the format across different software suites. It had to sacrifice compatibility with the old format to reach this higher level of abstraction.

The format is based on the following features :

- Open format.
- Compression.

- Multi platform support.
- Extendibility.
- Strong separation of content, layout and metadata
- Abstraction from the source application

The plan to create an open and extensible format seems to have been completed. The original mission was to “.. *create, as a community, the leading international office suite that will run on all major platforms and provide access to all functionality and data through open-component based APIs and an XML-based file format*” [4].

6 Conclusions

The method we have presented acts as a tool for selecting an Open Data Standard that can serve as a basis for making decisions about software purchasing or updating in a migration scenario.

This deliverable provides the theory and a number of examples that support this method, which has been applied in an illustrative way to COSPA data gathered from Work Package 2 deliverables D2.4/D2.5. For word-processing-centric processes, this method selected the OpenDocument format as the best, most relevant Open Standard to ensure that functional requirements were met.

Using this OpenDocument file format, PAs will benefit from greater interoperability and flexibility. They no longer need depend on a vendor or a platform, and the option of using another Operating System is now a matter for PAs to decide for themselves. The Linux Desktop has become increasingly popular among organisations seeking to migrate from other Operating Systems. Licence costs, portability, security and Open Standards make Linux a natural choice for European PAs.

Many European PAs have already migrated to Linux Desktop or plan to do so in the near future. News about OSS-related government activities in Europe and elsewhere can be found at the IDABC Open Source Observatory.[6]

The selected Open Data Standard enables PAs to migrate to such Open Source Operating Systems and/or Desktops.

Because of the demanding essential requirements (i.e. openness and cross-platform-compatibility), the OOo file format will almost certainly offer a preferred solution for PA's spreadsheet and presentation applications.

Appendix A - Abbreviations

FUB	Free University of Bolzano
USFD	University of Sheffield
ODS	Open Data Standard
DS	Data Standard
OSS	Open Source Software
OASIS	Organisation for the Advancement of Structured Information Standards
IETF	Internet Engineering Task Force
W3C	World Wide Web Consortium
TCP/IP	Transmission Control Protocol/Internet Protocol
IDA	Interchange of Data between Administrations
PA	Public Administration
EC	European Commission
e-GIF	e-Government Interoperability Framework (UK)
FLOSSpols	Free/Libre/Open Source Software: Policy Support
SAGA	Standards and Architectures for eGovernment Applications
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations
OOo	OpenOffice.org
DTD	Document Type Definition

Appendix B - Glossary

A list of words used with a special meaning throughout this document:

Method	A technique of solving a general task by a structured approach.
Procedure	The (mathematical) technique for making a selection.
Open Data Standard	File format based on an Open Standard, normally named identically to the Open Standard they implement. XML: .xhtml, .odt, .xml, HTML: .xhtml, .html, .htm OASIS OpenOffice Format: .od*
Requirement	A term describing a desired feature in a software application.
Openness	The minimum requirements for “openness” are: the standard is completely described in publicly accessible documents, this description may be distributed freely, and the standard may be implemented without restrictions, i.e. royalty-free, with no legal bindings.
Interchange of Data between Administrations	IDA is a European Commission driven strategic initiative using advances in information and communications technology to support rapid electronic exchange of information between Member State administrations.
Open Source Software	Any computer software distributed under a license which allows users to change and/or share the software freely. Many programs use a specific license agreement satisfying the Open Source Definition.

References

- 1 Erik Sliman, *Business Case for Open Standards*. Retrieved on 04/03/2005 from <http://www.openstandards.net/viewOSnet1C.jsp?showModuleName=businessCaseForOpenStandards>
- 2 EUROPA – IDABC. *European Interoperability Framework for pan-European e Government Services*.
- 3 OASIS (2004). *OASIS Open Office XML Format TC*. Retrieved on 01/03/2005 from http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office
- 4 A. Rizk (Valoris). *Comparative assessment of Open Document Formats Market Overview*. ENTR/02/21-IDA/MIDDLEWARE-XML v1.6
- 5 OpenOffice.org. *xml*:. Retrieved on 04/03/2005 from <http://xml.openoffice.org/>
- 6 EUROPA - IDABC. *Open Source Observatory News*. Retrieved on 04/03/2005 from <http://europa.eu.int/idabc/en/chapter/469>
- 7 J. David Eisenberg, *OpenDocument Essentials*, O'Reilly Associates, retrieved on 06/09/2005 from <http://book.evc-cit.info>
- 8 OpenDocument specification document, retrieved on 06/09/2005 from <http://www.oasis-open.org/committees/download.php/12572/OpenDocument-v1.0-os.pdf>