





# DELIVERABLE

Project Acronym:	PREFORMA
Grant Agreement number:	619568
Project Title:	PREservation FORMAts for culture information/e- archives

## **D8.1 Competitive Evaluation Strategy**

Revision: Final, 1.20, 21 April 2015

Authors:

Maristella Agosti (UNIPD) Nicola Ferro (UNIPD) Bert Lemmens (PACKED) Gianmaria Silvello (UNIPD)

**Reviewers:** 

**Claudio Prandoni (PROMOTER)** 

Project co-funded by the European Commission within the ICT Policy Support Programme				
	Dissemination Level			
Ρ	Public	X		
С	Confidential, only for members of the consortium and the Commission Services			



PREFORMA - Future Memory Standards PREservation FORMAts for culture information/e-archives EC Grant agreement no: 619568

# **Revision History**

Revision	Date	Author	Organisation	Description
0.10	2014-02-05	Nicola Ferro	UNIPD	First skeleton and initial draft
0.20	2014-02-21	Maristella Agosti, Nicola	UNIPD	Added main content to all
		Ferro, Gianmaria Silvello		sections
0.30	2014-03-18	Nicola Ferro	UNIPD	Consolidated draft
0.40	2014-02-21	Maristella Agosti, Nicola Ferro, Gianmaria Silvello	UNIPD	First consolidated version after input from the con- ference call to prepare the PREFORMA InfoDay
0.45	2014-11-06	Nicola Ferro, Gianmaria Silvello	UNIPD	Updated after input from conference call with other partners
0.50	2014-11-06	Bert Lemmens	PACKED	Input from challenge brief and invitation to tender
0.60	2014-11-14	Nicola Ferro, Gianmaria Silvello	UNIPD	Reorganization of the de- liverable according to part- ners input
0.61	2014-11-15	Gianmaria Silvello	UNIPD	Addition of categories and items
0.62	2014-11-16	Nicola Ferro	UNIPD	Update of evaluation model and its properties
0.70	2014-11-25	Nicola Ferro	UNIPD	Updated after input from in- ternal review, added nego- tiation protocols and check for any inconsistencies
1.00	2014-12-19	Nicola Ferro	UNIPD	Final revised version of the deliverable
1.10	2015-04-07	Nicola Ferro	UNIPD	Revised version according to reviewers comments at the 1st PREFORMA re- view meeting in Brussels on 26th March 2015
1.20	2015-04-21	Bert Lemmens	PACKED	Revised version after re- ceiving the first review re- port to take into considera- tion reviewers' comments



#### Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.



## Contents

E>	ecuti	ive Summary	7
1	Intro	oduction	9
2	Eva	luation Model	9
	2.1	Evaluation Process	9
		2.1.1 Score Computation	12
		2.1.2 Scoring Model	13
		2.1.3 Reviewer Type Aggregation	14
		2.1.4 Example of Use of the Evaluation Matrix	14
		2.1.5 Properties of the Evaluation Matrix	16
	2.2	Comparison Process	17
z	Πρε	ion of the PREFORMA Evaluation Matrix	18
5	2 1	Impact on the Challenge	20
	0.1	3.1.1 Description	20
		3.1.2 Itoms	20
	2.2		20
	5.2		25
			20
	<u> </u>	0.2.2 Items	20
	3.3		28
			28
	<b>•</b> •	3.3.2 Items	28
	3.4		29
			29
		3.4.2 Items	30
	3.5		30
4	Con	clusions	30
Re	eferer	nces	30
Α	Neg	otiation Protocol [Confidential]	32



PREFORMA - Future Memory Standards PREservation FORMAts for culture information/e-archives EC Grant agreement no: 619568



## **Executive Summary**

This deliverable defines the *competitive evaluation strategy* which will be used for the assessment of the results of the suppliers at the end of the design phase 1. The competitive evaluation strategy serves to choose those suppliers which will pass the design phase 1 and will continue with the prototyping and testing phases.

Evaluating and comparing suppliers requires us to identify two distinct processes:

• *evaluation process*: during this process each supplier is individually examined and it is scored according to its characteristics.

The evaluation process is formalized through the **evaluation matrix**, as described in Section 2.1.

The outcome of the evaluation process is the **supplier score**, that is a number representing the scoring achieved by the supplier;

• *comparison process*: once the suppliers have been scored, they are compared with each other on the basis of their supplier scores.

The comparison process is formalized through the **comparison matrix**, as described in Section 2.2.

The outcome of the comparison process is a **ranking of the suppliers**, based on their scorings.

In particular, as detailed in Section 3, the PREFORMA evaluation matrix consists of four categories: *Impact on the Challenge, Technical Approach, Quality of the Tender*, and *Price/Cost*. Each category contains several items which are scored using a likert scale ranging from 1 (bad) to 5 (excellent), according to well-established standards, like ITU-T P800. The score for each item is obtained with a weighted average of the scores assigned by different reviewers which belong to the three following reviewer types: *technical expert*, *domain expert*, and *external expert*. The score of a category is the weighted average of the scores of its items and the total score for a supplier is the weighted average of its category scores.



PREFORMA - Future Memory Standards PREservation FORMAts for culture information/e-archives EC Grant agreement no: 619568



## 1 Introduction

The aim of this document is to develop a painstaking method to *evaluate* and *compare* different suppliers. In particular, the developed method will be effectively employed in the evaluation and comparison of PREFORMA suppliers.

The proposed *competitive evaluation strategy* will be used for the assessment of the suppliers at the end of the design phase 1 in order to choose those who will continue with the prototyping and testing phases.

The proposed method is a general evaluation framework, which is applied to the evaluation and comparison of suppliers as a relevant use case. This means that the proposed method can be successfully employed in other projects, in addition to the PREFORMA project.

In the first part of the deliverable we present the general framework and in the second part we describe how it is applied to the PREFORMA case. In this way we can achieve two goals: first we obtain a general method for evaluation and comparison and we learn to adapt it to our needs; second we obtain the actual tool for evaluating and comparing PREFORMA suppliers at the end of design phase 1, which is the primary aim of this study.

The document is organized as follows: Section 2 describes the overall evaluation and comparison model; in particular Section 2.1 describes the design, the functioning and the properties of the evaluation matrix, while Section 2.2 illustrates the comparison matrix. Section 3 provides the design of the actual PREFORMA evaluation matrix. Appendix A contains confidential information agreed during the negotiation with the suppliers which will be checked at the end of design phase 1.

## 2 Evaluation Model

Evaluating and comparing suppliers requires us to identify two distinct processes:

• *evaluation process*: during this process each supplier is individually examined and it is scored according to its characteristics.

The evaluation process is formalized through the **evaluation matrix**, as described in Section 2.1.

The outcome of the evaluation process is the **supplier score**, that is a number representing the scoring achieved by the supplier;

• *comparison process*: once the suppliers have been scored, they are compared with each other on the basis of their supplier's score.

The comparison process is formalized through the **comparison matrix**, as described in Section 2.2.

The outcome of the comparison process is a **ranking of the suppliers**, based on their scorings.

#### 2.1 Evaluation Process

The evaluation process aims to grasp and model the principal aspects of the complex reality of interest.



The following entities are utilized to capture and formalize the different aspects of the evaluation process:

- **Category**: represents a main aspect of a supplier under examination. A category covers homogeneous properties of a supplier;
- Item: describes an elementary constituent of a category and it is used to detail a category;
- **Reviewer Type**: takes into account different angles according to which a category can be assessed, i.e. different facets and standpoints of a category. For example, within the same category we may consider the viewpoint of both a technical expert and a domain expert.

Figure 1 shows the relationship among category, item and reviewer type: a category is composed by different items which are evaluated from different point-of-views – i.e. reviewer types.

Category					
	Reviewer Type 1	Reviewer Type 2	Reviewer Type 3		
Item 1					
Item 2					

Figure 1: Relationship among category, item and reviewer type.

Note that reviewer types within a category are the same for all the items of the category. This is a direct consequence of the homogeneousness of the category: reviewer types – defined for the category – are naturally inherited by items.

Note that different categories are not required to be assessed by all the reviewer types, as shown in Figure 2, where "Category 1" is assessed by "Reviewer Type 1", "Reviewer Type 2" and "Reviewer Type 3" while "Category 2" is assessed just by "Reviewer Type 1" and "Reviewer Type 3".

This choice for representing the reality leads to a hierarchical way of scoring a supplier, since categories, reviewer types and items represent a different level of abstraction of the reality under examination. Thus it is opportune to exploit this intrinsic hierarchy using a weighted scoring scheme. This scoring scheme assigns different weights, i.e. different importance, to different levels in the hierarchy.

The scoring scheme is represented by means of the *evaluation matrix*, which is composed of the various categories of interest, which in turn are made up of different items, assessed by several reviewer types, as shown in Figure 3.

The final outcome of the evaluation matrix is the *supplier score*, as explained in section 3, which is computed as shown in Figure 3:



Category 1					
	Reviewer Type 1	Reviewer Type 2	Reviewer Type 3		
Item 1					
Item 2					
Categor	y 2				
	Reviewer Type 1		Reviewer Type 3		
Item 1					
Item 2					

Figure 2: Categories with different reviewer types.

• *item score*: is the weighted sum of the *item sub-scores*, i.e. the scoring by a reviewer type of that item.

In figure 3 the item sub-score is represented by a circle, and the item score by a hexagon;

- *category score*: is the weighted sum of the item scores for each item within that category. In figure 3 the category score is represented by a diamond;
- *supplier score*: is the weighted sum of the category scores for each category. In figure 3 the supplier score is represented with a sort of flag shaped figure.

The evaluation process is composed by the following steps:

- 1. definition of the reviewer types;
- 2. definition of the categories, their weights and the weights of their reviewer types;
- 3. definition of items and their respective weights;
- 4. scoring each item sub-score.

This procedure will be described in depth in Section 3 for the evaluation of the PREFORMA suppliers, which represents a relevant example of the application of the proposed method.



Category 1					
	Reviewer Type 1	Reviewer Type 2	Reviewer Type 3	Total	
Itom 1					
Item 2	+		+ =	+	
Total					
Catego	ry 2				
	Reviewer Type 1		Reviewer Type 3	Total	
ltone d					
Itom 2				+	
Total					
lotar					
Supplier	Score				

Figure 3: Basic functioning of the evaluation matrix.

#### 2.1.1 Score Computation

We can describe the detailed procedure for computing the supplier score. Let us consider Figure 3 in reverse order thus beginning from the explanation of the category score.

• The supplier score is a weighted sum of category scores: if we call  $C_i$  the category score of the *i*-th category and  $\alpha_i$  the coefficient expressing the weight of the *i*-th category, the supplier score is given by the equation (1),

supplier score = 
$$\alpha_1 C_1 + \alpha_2 C_2 + \ldots + \alpha_n C_n = \sum_{i=1}^n \alpha_i C_i$$
 (1)

where n is the total number of categories.

Note that from now on the subscript *i* in the following equations stands for the *i*-th category.



• The category score is a weighted sum of item scores: if we call  $I_{i,j}$  the score of the *j*-th item within the category  $C_i$  and  $\beta_{i,j}$  the coefficient expressing the weight of the *j*-th item, the category score is given by equation (2),

$$C_{i} = \beta_{i,1}I_{i,1} + \beta_{i,2}I_{i,2} + \ldots + \beta_{i,k_{i}}I_{i,k_{i}} = \sum_{j=1}^{k_{i}} \beta_{i,j}I_{i,j}$$
(2)

where  $k_i$  is the total number of items within the category  $C_i$ .

• The item score is a weighted sum of item sub-scores: In the context of the application of this evaluation method, that is the PREFORMA evaluation matrix presented in Section 3 we will need to consider categories with a maximum of three reviewer types. To simplify the notation, we tailor it to only three reviewer types; however, it could be easily extended to as many reviewer types as necessary.

If we call  $I'_{i,j}$ ,  $I''_{i,j}$  and  $I'''_{i,j}$  the item sub-scores for the generic item  $I_{i,j}$  with respect to the first, second and third reviewer types, and we call  $\gamma'_i$ ,  $\gamma''_i$  and  $\gamma'''_i$  the coefficients expressing the weight of the first, second and third reviewer type within the category  $C_i$ , the item score is given by equation (3).

$$I_{i,j} = \gamma'_i I'_{i,j} + \gamma''_i I''_{i,j} + \gamma'''_i I''_{i,j}$$
(3)

It is generally understood that all the weighting coefficients  $\alpha$ ,  $\beta$  and  $\gamma$  are normalized to 1, as shown in equation (4). Besides being necessary for the correctness of computations, normalization allows us to easily compare scores of different items and different categories, gaining a quick overview of weak and strong points of a supplier.

$$\alpha_{1} + \alpha_{2} + \dots + \alpha_{n} = \sum_{i=1}^{n} \alpha_{i} = 1$$

$$\beta_{i,1} + \beta_{i,2} + \dots + \beta_{i,k_{i}} = \sum_{j=1}^{k_{i}} \beta_{i,j} = 1, \quad \forall i$$

$$\gamma'_{i} + \gamma''_{i} + \gamma''_{i} = 1, \quad \forall i$$

$$(4)$$

#### 2.1.2 Scoring Model

The following is the item sub-score grading system, which ranges from 1 to 5:

- 1 Bad
- 2 Poor
- 3 Fair



- 4 Good
- 5 Excellent.

Note that these values are consistent, for example, with the evaluation scale defined in Recommendation P.800 [ITU-T, 1996] by *Telecommunication Standardization Sector of ITU (ITU-T)*. The motivation of this choice is to render the evaluation matrix as close as possibile to international standardized evaluation methods, since the evaluation matrix should provide unbiased results taking into consideration different suppliers.

#### 2.1.3 Reviewer Type Aggregation

For each reviewer type, the item sub-score can be the result of the scoring of a single or multiple reviewers of that type. In the latter case, we consider as aggregate item sub-score the average of the scores assigned by the reviewers.

Let  $rt' \in [1, 2, ...]$  the number of reviewers of the first type,  $rt'' \in [1, 2, ...]$  the number of reviewers of the second type, and  $rt''' \in [1, 2, ...]$  the number of reviewers of the third type, then we have:

$$I'_{i,j} = \frac{1}{rt'} \sum_{k=1}^{rt'} I'^{(k)}_{i,j}$$

$$I''_{i,j} = \frac{1}{rt''} \sum_{k=1}^{rt''} I''^{(k)}_{i,j}$$

$$I'''_{i,j} = \frac{1}{rt''} \sum_{k=1}^{rt''} I''^{(k)}_{i,j}$$
(5)

where  $I_{i,j}^{\prime(k)}$ ,  $I_{i,j}^{\prime\prime\prime(k)}$ , and  $I_{i,j}^{\prime\prime\prime\prime(k)}$  are, respectively, the item sub–score assigned by the *k*-th reviewer of the first, second, and third type.

Note that, in general, rt', rt'', and rt''' can be different allowing for a further degree of flexibility. Indeed, you can decide to have the same number of reviewers for each reviewer type or you can decide to have different numbers of reviewers in each reviewer type if, for example, you need to make the assessment for a given reviewer type more stable.

#### 2.1.4 Example of Use of the Evaluation Matrix

Now we introduce an example of use of the evaluation matrix, which is shown in Table 1.

Note that "Reviewer Type 2" and "Reviewer Type 3" are not present within "Category 1", so their respective weights  $\gamma'_1$  and  $\gamma''_2$  are equal to 0.

Taking into consideration Table 1, when you move horizontally within the evaluation matrix, you have to use the weights  $\gamma'_i$ ,  $\gamma''_i$  and  $\gamma'''_i$  for adding scores. When you move vertically within a category, you have to use the weights  $\beta_{i,j}$  for adding scores. Finally, when you move vertically among categories, you have to use the weights  $\alpha_i$  for adding scores.



Supplier name: Vendor 1.					
	Reviewer Type 1	Reviewer Type 2	Reviewer Type 3	Total	
Category 1 ( $\alpha_1 = 0.3$ )	$\gamma'_1 = 1$	$\gamma_1'' = 0$	$\gamma_1''' = 0$		
ltem 1 ( $\beta_{1,1} = 0.4$ )	3.00	-	-	3.00	
ltem 2 ( $\beta_{1,2} = 0.6$ )	4.00	_	_	4.00	
Category 1 score	3.60	_	_	3.60	
Category 2 ( $\alpha_2 = 0.7$ )	$\gamma_2' = 0.4$	$\gamma_2'' = 0.3$	$\gamma_2^{\prime\prime\prime} = 0.3$		
ltem 1 ( $\beta_{2,1} = 0.3$ )	2.00	2.00	3.00	2.30	
ltem 2 ( $\beta_{2,2} = 0.1$ )	4.00	3.00	2.00	3.10	
ltem 3 ( $\beta_{2,3} = 0.4$ )	5.00	2.00	4.00	3.80	
ltem 4 ( $\beta_{2,4} = 0.2$ )	1.00	3.00	3.00	2.20	
Category 2 score	3.20	2.30	3.30	2.96	
Supplier Score 3					

Table 1: Example of evaluation matrix.

Now we describe in detail how to obtain Table 1, using the computations described in section 2.1, figure 3 and section 2.1.1:

• the item score of "Item 1" of the category "Category 1", using equation (3), is given by:

$$I_{1,1} = \gamma_1' I_{1,1}' + \gamma_1'' I_{1,1}'' + \gamma_1''' I_{1,1}''' = 1 \cdot 3.00 + 0 + 0 = 3.00$$

• the item score of "Item 2" of the category "Category 1", using equation (3), is given by:

$$I_{1,2} = \gamma_1' I_{1,2}' + \gamma_1'' I_{1,2}'' + \gamma_1''' I_{1,2}''' = 1 \cdot 4.00 + 0 + 0 = 4.00$$

• the category score of the category "Category 1", using equation (2), is given by:

$$C_1 = \beta_{1,1}I_{1,1} + \beta_{1,2}I_{1,2} = 0.4 \cdot 3.00 + 0.6 \cdot 4.00 = 3.60$$

• the item score of "Item 1" of the category "Category 2", using equation (3), is given by:

$$I_{2,1} = \gamma'_2 I'_{2,1} + \gamma''_2 I''_{2,1} + \gamma'''_2 I''_{2,1} = 0.4 \cdot 2.00 + 0.3 \cdot 2.00 + 0.3 \cdot 3.00 = 2.30$$

• the item score of "Item 2" of the category "Category 2", using equation (3), is given by:

$$I_{2,2} = \gamma'_2 I'_{2,2} + \gamma''_2 I''_{2,2} + \gamma''_2 I''_{2,2} = 0.4 \cdot 4.00 + 0.3 \cdot 3.00 + 0.3 \cdot 2.00 = 3.10$$

• the item score of "Item 3" of the category "Category 2", using equation (3), is given by:

$$I_{2,3} = \gamma'_2 I'_{2,3} + \gamma''_2 I''_{2,3} + \gamma''_2 I''_{2,3} = 0.4 \cdot 5.00 + 0.3 \cdot 2.00 + 0.3 \cdot 4.00 = 3.80$$



• the item score of "Item 4" of the category "Category 2", using equation (3), is given by:

$$I_{2,4} = \gamma_2' I_{2,4}' + \gamma_2'' I_{2,4}'' + \gamma_2''' I_{2,4}''' = 0.4 \cdot 1.00 + 0.3 \cdot 3.00 + 0.3 \cdot 3.00 = 2.20$$

• the category score of the category "Category 2", using equation (2), is given by:

 $C_2 = \beta_{2,1}I_{2,1} + \beta_{2,2}I_{2,2} + \beta_{2,3}I_{2,3} + \beta_{2,4}I_{2,4} = 0.3 \cdot 2.30 + 0.1 \cdot 3.10 + 0.4 \cdot 3.80 + 0.2 \cdot 2.20 = 2.96$ 

• the supplier score, using equation (1), is given by:

supplier score = 
$$\alpha_1 C_1 + \alpha_2 C_2 = 0.3 \cdot 3.60 + 0.7 \cdot 2.96 = 3.15$$

Note that these computations are the same of the process shown graphically in Figure 3.

#### 2.1.5 Properties of the Evaluation Matrix

The evaluation matrix holds an useful property: it allows us to create groups of items within a category. This could be very useful in the case of categories with many items or in the case in which it is necessary to logically group items within a category.

For example table 1 could be reorganized grouping items in "Category 2", as shown in Table 2.

Supplier name: Vendor 1.					
Reviewer Type 1Reviewer Type 2Reviewer Type 3					
Category 1 ( $\alpha_1 = 0.3$ )	$\gamma'_1 = 1$	$\gamma_1'' = 0$	$\gamma_1^{\prime\prime\prime} = 0$		
Item 1 ( $\beta_{1,1} = 0.4$ )	3.00	_	_	3.00	
Item 2 ( $\beta_{1,2} = 0.6$ )	4.00	-	_	4.00	
Category 1 score	3.60	_	-	3.60	
Category 2 ( $\alpha_2 = 0.7$ )	$\gamma_2' = 0.4$	$\gamma_2'' = 0.3$	$\gamma_2^{\prime\prime\prime}=0.3$		
GROUP 1					
Item 1 ( $\beta_{2,1} = 0.3$ )	2.00	2.00	3.00	2.30	
Item 2 $(\beta_{2,2} = 0.1)$	4.00	3.00	2.00	3.10	
Total Group 1	1.00	0.90	1.10	1.00	
GROUP 2					
Item 3 $(\beta_{2,3} = 0.4)$	5.00	2.00	4.00	3.80	
Item 4 ( $\beta_{2,4} = 0.2$ )	1.00	3.00	3.00	2.20	
TOTAL GROUP 2	2.20	1.40	2.20	1.96	
Category 2 score	3.20	2.30	3.30	2.96	
Supplier Score3.15					

Table 2: Evaluation matrix of table 1 organized with groups.



According to Table 2 the computation of  $C_2$ , which has been shown above, can be organized as follows:

$$C_{2} = \underbrace{\beta_{2,1}I_{2,1} + \beta_{2,2}I_{2,2}}_{\text{Group 1}} + \underbrace{\beta_{2,3}I_{2,3} + \beta_{2,4}I_{2,4}}_{\text{Group 2}} = \underbrace{\left(0.3 \cdot 2.30 + 0.1 \cdot 3.10\right)}_{\text{Group 1} = 1.00} + \underbrace{\left(0.4 \cdot 3.80 + 0.2 \cdot 2.20\right)}_{\text{Group 2} = 1.96} = 1.00 + 1.96 = 2.96$$

This equation provides new information: the supplier "Vendor 1" is rated 1.00 in GROUP 1 of "Category 2" and 1.96 in GROUP 2 of "Category 2".

Thus this property of the evaluation matrix can be employed to summarize the content of a category into logical groups, hiding the detailed description of all the items within a category, as shown in Table 3.

Supplier name: Vendor 1.					
	Reviewer Type 1	Reviewer Type 2	Reviewer Type 3	Total	
Category 1 ( $\alpha_1 = 0.3$ )	$\gamma'_1 = 1$	$\gamma_1'' = 0$	$\gamma_1^{\prime\prime\prime}=0$		
ltem 1 ( $\beta_{1,1} = 0.4$ )	3.00	-	_	3.00	
ltem 2 ( $\beta_{1,2} = 0.6$ )	4.00	-	-	4.00	
Category 1 score	3.60	_	_	3.60	
Category 2 ( $\alpha_2 = 0.7$ )	$\gamma_2' = 0.4$	$\gamma_2'' = 0.3$	$\gamma_2^{\prime\prime\prime} = 0.3$		
TOTAL GROUP 1	1.00	0.90	1.10	1.00	
TOTAL GROUP 2	2.20	1.40	2.20	1.96	
Category 2 score	3.20	2.30	3.30	2.96	
Supplier Score					

Table 3: Evaluation matrix of table 2 summarized by groups.

Note that creating groups within a category is as simple as putting parentheses in the weighted sum of the category score, as shown in the equation above.

Furthermore grouping items within a category can lead to additional and useful information: we can determine the weight of a subset of the items of the category.

For example, considering Table 2, we can see that:

GROUP 1 WEIGHT =  $\beta_{2,1} + \beta_{2,2} = 0.3 + 0.1 = 0.4$ GROUP 2 WEIGHT =  $\beta_{2,3} + \beta_{2,3} = 0.4 + 0.2 = 0.6$ 

We know that 40% of the score of "Category 2" is determined by "Group 1" and 60% of the score of "Category 2" is determined by "Group 1".

#### 2.2 Comparison Process

The results of the evaluation matrix for each supplier is summarized by the **comparison matrix**, which offers a simple instrument for comparing different suppliers and choosing the best one.



In table 4 an example of comparison matrix is shown: the results are grouped according to category level.

	Supplier: Vendor 1	Supplier: Vendor 2
Category 1 Score ( $\alpha_1 = 0.3$ )	3.60	4.20
Category 2 Score $(\alpha_2 = 0.7)$	2.96	2.10
Total Supplier Score	3.15	2.73

Table 4: Example of comparison matrix with results grouped at category level.

If you need a greater level of detail, you can group results at group level, as shown in table 5.

	Supplier: Vendor 1	Supplier: Vendor 2
Category 1 Score ( $\alpha_1 = 0.3$ )	3.60	4.20
Category 2 ( $\alpha_2 = 0.7$ )		
TOTAL GROUP 1	1.00	0.90
TOTAL GROUP 2	1.96	1.20
Category 2 Score	2.96	2.10
Total Supplier Score	3.15	2.73

Table 5: Example of comparison matrix with results grouped at group level.

Obviously you can choose an even further level of detail, for example, showing the items of a category.

## **3** Design of the PREFORMA Evaluation Matrix

As explained in Section 2.1, designing the evaluation matrix involves the following steps:

- 1. defining reviewer types;
- 2. defining categories, their respective weights and the weights of reviewer types within each category;
- 3. defining items and their respective weights.

Figure 4 shows the main standpoints that need to be covered by the evaluation process.

We need to consider three reviewer types, which correspond to the three main stakeholders involved in the PREFORMA project, as shown in Figure 4. They are:

- **Technical Expert**: the proposed solution is analyzed by a technical expert, who is evaluating the solution from the technical point-of-view;
- **Domain Expert**: the proposed solution is analyzed by a domain expert, who verifies if the solution well fits the requirement of the domain where it will be used;





Figure 4: Angles covered by the evaluation process.

• External Expert: the proposed solution is analyzed by an expert external to the PREFORMA consortium, to compensate for any possible biases.

In particular, each supplier solution will be reviewed by the following number of reviewers for each reviewer type:

- **Technical Expert**: 2 reviewers for each proposal, i.e. rt' = 2;
- **Domain Expert**: 3 reviewers for each proposal, i.e. rt'' = 3;
- External Expert: 1 reviewer for each proposal, i.e. rt''' = 1;

The following is a detailed description of all the categories and all the items for each category is presented. Each category is specified by:

- name;
- brief description;
- weight of the category;
- weights of the different reviewer types within the category;



- complete list of all the items of the category. Each item is specified by:
  - name;
  - brief description;
  - meaning of the scores for that item;
  - weight of the item.

#### 3.1 Impact on the Challenge

#### 3.1.1 Description

The category "Impact On the Challenge" concerns the extent of how well the proposed idea, solution or technology meets the challenge as detailed in the Brief, and whether it will have the desired impact.

This category pertains all three reviewer types.

The following parameters are valid for the category:

- Category  $C_1$  IMPACT ON THE CHALLENGE
- Category weight:  $\alpha_1 = 35\%$ This category evaluates the extent to which the proposed solution meets the challenge.
- Reviewer Type weights:
  - "Technical Expert" reviewer type weight:  $\gamma_1' = 30\%$
  - "Domain Expert" reviewer type weight:  $\gamma_1'' = 40\%$
  - "External Expert" reviewer type weight:  $\gamma_1'' = 30\%$

#### 3.1.2 Items

- Item  $I_{1,1}$  Basic research questions
  - description: This item concerns: (i) how to interpret and implement standard specifications; (ii) how to determine whether a file is what it claims to be; and (iii) how to make OS project sustainable.

In order to evaluate the aspect (i) the reviewer has to consider if the project establishes a methodology or an objective frame of reference to interpret and implement the standard specifications against the background of the current variations of interpretations and implementations by software vendors and if there is a need to consolidate the diverse implementations or if it is a better to centralize the interpretation to a specific implementation (i.e. promote one interpretation and implementation as the standard).

In order to evaluate the aspect (ii) the reviewer has to determine whether a file is what it claims to be, i.e., in this context, what makes a file a valid file, or is it conform to the



#### "standard"?

In order to evaluate the aspect (iii) the reviewer has to consider how the open source project can be developed and sustained in the short and long run and if an open source community can operate as the normative source for the answer to the first and second question.

- *item weight*:  $\beta_{1,1} = 10\%$
- Item  $I_{1,2}$  Conformance Checker
  - description: This item concerns the conformance checker and it contains some general aspects and some aspects which are specific of the file type the checker is operating with.

For evaluating the general aspects the reviewer has to consider: (i) if the project develops an open-source conformance checker; (ii) if the conformance checker enables implementation of the OAIS Quality assurance function at Ingest, validating (QA results) the successful transfer of the SIP to the temporary storage area; (iii) if the conformance checker enables implementation of the OAIS Generate AIP function at ingest, transforming one or more SIPs into one or more AIPs that conform to the archive's data formatting standards and documentation standards; (vi) if the conformance checker enables implementation of the OAIS Archival Information Update function at ingest, providing a mechanism for updating (repackaging, transformation) the contents of the archive.

For evaluating the file type specific aspects the reviewer has to consider: (PDF/A) if basic research activities involve checking for the existence of PDF/A functionalities and if they are implemented in accordance with the specifications for PDF/A; (TIFF) if basic research activities involve checking for the existence of TIFF functionalities and if they are implemented in accordance with the specifications for TIFF; (AV) if basic research activities involve the definition of a profile for an audiovisual preservation file.

The conformance checker would be used to evaluate and if possible fix a SIP and convert it into an AIP. Practically this would mean: Your repository receives an exotic TIFF or PDF file and convert it into a TIFF/A or PDF/A file, not by transcoding the file, but by stripping/adding/editing information in the header and the structure of the file. This was how we conceived the Generate AIP function in the OAIS model.

As for evaluating the use of OAIS as a reference framework, "enable" refers meeting the technical requirements for integrating the conformance checker in existing workflows and designing a shell that allows for performing essential tasks that fit the OAIS framework. So for example Generate OAIS requires a machine readable report that can be used by a transcoder to convert the file. Or when small errors are concerned, the fixer module should be able to perform the conversion.

- *item weight*:  $\beta_{1,2} = 20\%$
- Item  $I_{1,3}$  Reference Implementation



- *description*: This item considers many aspects of the reference architecture of the proposed solutions. The aspects that need to be evaluated by the reviewer are:

(i) healthy ecosystem: the project establish a healthy ecosystem around an open source 'reference' implementation for specific file formats;

(ii) demonstration files: technology providers contribute demonstration files with good and bad samples of the corresponding reference implementation;

(iii) documentation of the source code: technology providers contribute comprehensive documentation of the source code, which allows for automated generation of the internal API of the application;

(iv) documentation of the software: technology providers contribute comprehensive documentation of the conformance checker for developers, such as quick start guide, cookbooks and other tutorials;

(v) online technical support: technology providers ensure online availability at the development platform for technical support to other developers deploying the conformance checker;

(vi) marketing at conferences: technology providers market the reference implementation and conformance checker at conference for professional networks of developers and digital preservationists;

(vii) propose changes and additions: technology providers draft proposals for changes and additions to the standard specifications;

(viii) participate in work-groups: technology providers participate in technical workgroups that maintain a standard specification;

(ix) facilitate OAIS Monitor Designated Communities: the network of common interest enables implementation of the OAIS Monitor Designated Communities function for Preservation Planning, interacting with Archive Consumers and Producers to track changes in their service requirements and available product technologies;

(x) facilitate OAIS Develop Preservation Strategies and Standards: the network of common interest enables implementation of the OAIS Develop Preservation Strategies and Standards function for preservation planning, developing and recommending strategies and standards, and for assessing risks, to enable the Archive to make informed trade-offs as it establishes standards, sets policies, and manages its system infrastructure;

(xi) facilitate OAIS Establishing Standards and Policies: the network of common interest enables implementation of the OAIS Establishing Standards and Policies function by the Administration of the Archive system and maintain them.

- *item weight*:  $\beta_{1,3} = 20\%$ 

#### • Item $I_{1,4}$ – Future/wider challenges/future proof

- description: The reviewer has to evaluate the potential of the proposal to address future/wider challenges in the area in an innovative way (e.g. by developing or employing novel concepts, approaches, methodologies, tools, or technologies).
- item weight:  $\beta_{1,4} = 5\%$



• Item  $I_{1,5}$  – Commercial feasibility

- description: The reviewer has to evaluate the extent to which the approach demonstrates commercial feasibility, and whether it is a realistic commercialization plan or route to market; in particular it has to be considered the following aspects: (i) integration with text, image, moving image editors; (ii) integration with digital repositories; (iii) integration with transcoding software; (iv) integration of additional conformance checkers; (v) integration of additional reporters; (vi) providing consulting services; (vii) providing customization services; (viii) providing support services.
- *item weight*:  $\beta_{1,5} = 15\%$

#### • Item $I_{1,6}$ – Open source work practices

 description: The reviewer has to evaluate this item considering that the development of software in open source projects in PREFORMA must utilise effective open source work practices.

In particular it has to be considered the following aspects: (i) nigthly builds; (ii) open development platform; (iii) issue/bug trackers; (iv) developer communication channels (e.g. use of forums, use of mailing lists for different stakeholder groups (users, developers, etc.) and use of IRC, provision of roadmaps, provision of documentation, provision of easy hacks, etc.).

- *item weight*:  $\beta_{1,6} = 5\%$
- Item  ${\it I}_{1,7}$  Open Source release practice / Delivery and installation
  - description: This category concerns aspects regarding the delivery and installation of a system.

In particular it has to be considered the following aspects: (i) executable source code: for each executable of developed software that is provided in an open source project, the source code must always be provided for that executable.;

(ii) instructions for making executables: for each executable of developed software that are provided in an open source project, instructions for how to create the executable from the source code must always be provided;

(iii) open source tools for making executables: for each executable of developed software that are provided in an open source project at the PREFORMA open source portal, open source tools (provided under any license approved by Open Source Initiative) for creation of the executable from the source code must be provided;

(iv) executables for multiple platforms: there must always be executables for several different platforms (at least for: MS Windows 7, Mac OSX, common Linux distributions such as Ubuntu, Fedora, Debian, and Suse).



- *item weight*:  $\beta_{1,7} = 10\%$ 

#### • Item I<sub>1.8</sub> – Open Source interaction practice

- description: Individuals in companies contracted by PREFORMA will adopt a workpractice which promote a diverse long-term sustainable Open Source community (which have active participants and contributors from several different organisations). In particular it has to be considered the following aspects: (i) engage in timely fashion:

companies contracted by PREFORMA for development and provision of software and associated digital assets in Open Source projects must be responsive with respect to contributions to the project and are expected to engage in activities in a timely fashion;

(ii) open collaboration: companies contracted by PREFORMA for development and provision of software and associated digital assets in Open Source projects must be responsive with respect to contributions to the project and are expected to promote an open collaboration and become active community members which adhere to established community values and work-practices;

(iii) promote external contribution: companies contracted by PREFORMA for development and provision of software and associated digital assets in Open Source projects must be responsive with respect to contributions to the project and are expected to promote external contributions to each Open Source project;

(iv) contribute to other projects: companies contracted by PREFORMA for development and provision of software and associated digital assets in Open Source projects must be responsive with respect to contributions to the project and are expected to be active contributors in other relevant Open Source projects that are related to the Open Source project for which they are contracted;

(v) interact with standardisation organisations: the open source projects conducting development of software for PREFORMA must actively engage in interacting with relevant organisations that maintain the standard specifications used by the open project. The aim is to provide feedback, resolve technical issues;

(vi) interact with software providers: the open source projects conducting development of software for PREFORMA must actively engage in interacting with relevant software providers (i.e. those providers which have developed software used for creation of files in the specific file format checked by the PREFORMA software) for provision of feedback, resolving technical issues, and contribute in a dialogue for improvement of their interpretation of the technical specifications of standards implemented in their software; (vii) respect of the negotiation protocol.

- *item weight*:  $\beta_{1.8} = 10\%$

#### • Item *I*<sub>1,9</sub> – Open Source IPR distribution

- description: This category has to be evaluated by considering the following aspects: (i) software and source code: "GPLv3 or later" and "MPLv2 or later": all software de-



veloped during the PREFORMA project must be provided under the two specific open source licenses: "GPLv3 or later" and "MPLv2 or later";

(ii) open formats EIFv1.0/open standards: All digital assets developed during the PRE-FORMA project must be provided in open file formats, i.e. an open standard as defined in the European Interoperability Framework for Pan-European eGovernment Service (version 1.0 2004). This item concerns the degree of proprietary solution, i.e. if the system uses an open standard solution or a proprietary solutions;

(iii) CC-BYv4.0: all digital assets developed during the PREFORMA project must be provided under the open access license: Creative Commons CC-BY v4.0; (vii) respect of the negotiation protocol.

- *item weight*:  $\beta_{1,9} = 5\%$ 

#### 3.2 Technical Approach

#### 3.2.1 Description

This category regards all the technical aspects concerning the proposal. The following parameters are valid for the category:

- Category  $C_2$  TECHNICAL APPROACH
- Category weight:  $\alpha_2 = 35\%$ This category evaluates the technical quality of the proposal.
- Reviewer Type weights:
  - "Technical Expert" reviewer type weight:  $\gamma_2' = 50\%$
  - "Domain Expert" reviewer type weight:  $\gamma_2'' = 20\%$
  - "External Expert" reviewer type weight:  $\gamma_2^{\prime\prime\prime}=30\%$

#### 3.2.2 Items

- Item  $I_{2,1}$  Architecture
  - description: This item concerns infrastructural aspects, technical specifications and system features of a system. In particular it has to be evaluated by considering:

(i) Interoperability: this item concerns the degree at which the solution can interoperate with other components and solutions;

(ii) Scalability: this item concerns the degree at which the solution is scalable and expansible;

(iii) Portability: source code must be built for portability between technical deployment platforms (platform independent);

(iv) Modularity: source code must be built in a modular fashion for improved maintainability;



(v) Deployment: the Conformance Checker must allow for deployment in the five infrastructures/ environments defined in the Challenge Brief, i.c. PREFORMA website, stand alone, networked, in legacy system and in test environment;

(vi) Interface: the Conformance Checker must interface with other software systems via APIs;

- *item weight*:  $\beta_{2,1} = 30\%$
- Item *I*<sub>2,2</sub> Performances and Quality
  - *description*: The goal of this item is to evaluate the general performances and the quality, which are measured from both an objective and a subjective point of view.
  - *item weight*:  $\beta_{2,2} = 10\%$

#### • Item I<sub>2,3</sub> – Shell Services and features

- *description*: This item concerns functionalities and services offered by a system. It has to be evaluated by considering:

(i) checking at creation time: the Shell component of the Conformance Checker must facilitate conformance checking of files at four moment in the life cycle of a digital document, identified in the use cases of the challenge brief, including conformance checking at creation time, transfer time, digitisation time and migration time;

(ii) checking at transfer time: The Shell component of the Conformance Checker must facilitate conformance checking of files at four moment in the life cycle of a digital document, identified in the use cases of the challenge brief, including conformance checking at creation time, transfer time, digitisation time and migration time; (iii) checking at digitisation time: The Shell component of the Conformance Checker must facilitate conformance checking of files at four moment in the life cycle of a digital document, identified in the use cases of the challenge brief, including conformance checking at creation time, transfer time, digitisation time and migration time; (iii) checking at creation time, transfer time, digitisation time and migration time; transfer time, digitisation time and migration time;

(iv) checking at migration time: The Shell component of the Conformance Checker must facilitate conformance checking of files at four moment in the life cycle of a digital document, identified in the use cases of the challenge brief, including conformance checking at creation time, transfer time, digitisation time and migration time;

(v) automated checks: The Shell component of the Conformance Checker must allow for automating the procedures for checking, reporting and fixing preservation file;

(vi) periodical checks: The Shell component of the Conformance Checker must allow for configuring fully automated, periodical checks;

(vii) batch processing: The Shell component of the Conformance Checker must allow for batch processing of extensive file sets;



(viii) addotopma: The Shell component of the Conformance Checker must allow for configuration of additional components in particular implementation checkers, policy checkers and reporters for other preservation file formats that are developed in the PRE-FORMA ecosystem;

(ix) use by non-expert users: The Shell component of the Conformance Checker must allow for use by non-expert users;

(x) operate without Internet: The Shell component of the Conformance Checker must be operational in a closed zone with no Internet access.

- *item weight*:  $\beta_{2,3} = 30\%$
- Item *I*<sub>2,4</sub> Implementation Checker Services and features
  - *description*: This item has to be evaluated following different criteria on the basis of the file type the checker is designed for.

For text the checker has to test the compliancy of: PDF 1.4 (PDF/A-1) [ISO 19005-1, 2005], PDF 1.7 [ISO 32000-1, 2008], PDF/A-2 [ISO 19005-2, 2011] and PDF/A-3 [ISO 19005-3, 2012].

For images the checker has to test the compliancy of: TIFF/EP [ISO 12234-2, 2001] and TIFF/IT [ISO 12639, 2004].

For Audio/video the checker has to test the compliancy of: MKV, OGG, Lossless JPEG2000 [ISO/IEC 15444, 2004], Lossless FFV1 and LPCM [IEC 60958, 2014].

– item weight:  $\beta_{2,4} = 10\%$ 

#### • Item I<sub>2.5</sub> – Policy Checker Services and features

- *description*: This item has to be evaluated by considering technical metadata for text, technical metadata for image and technical metadata for av.
- *item weight*:  $\beta_{2,5} = 5\%$
- Item I<sub>2,6</sub> Reporter Services and features
  - *description*: This item has to be evaluated by considering if the checker produces both machine readable report and human readable report.

Machine readable report must provide preservation metadata for each file checked and allowing external software agents to further process the file. The machine readable report will be produced using a standard XML format, implemented by all conformance checkers in the PREFORMA ecosystem, which allows the reported module to combine output from multiple checker components in one report.

Human readable report must provide a human readable report, assessing the preservation status of a batch of files as a whole, reporting to a non-expert audience whether a file is compliant with the standard specifications, and addressing improvements in the creation/digitisation process.



- *item weight*:  $\beta_{2,6} = 10\%$ 

#### • Item I<sub>2.7</sub> – Metadata fixer Services and features

- *description*: This item has to be evaluated by considering if the checker:

(i) aligns embedded metadata: The Metadata fixer component of the Conformance Checker must allow for performing fully automated fixes of incongruities in the metadata embedded in the file, based on the report of the implementation checker. Such automated fixes may include making embedded technical metadata conform with the properties of video and audio essence contained by the preservation file;

(ii) essences normalising metadata: The Metadata fixer component of the Conformance Checker must allow for performing fully automated fixes of incongruities in the metadata embedded in the file, based on the report of the implementation checker. Such automated fixes may include normalising embedded administrative metadata about the preservation file.

- *item weight*:  $\beta_{2,7} = 5\%$ 

#### 3.3 Quality of the Tender

#### 3.3.1 Description

This category deal with all the aspects related to project management and how the negotiation protocol has been taken into account.

The following parameters are valid for the category:

- Category  $C_3$  QUALITY OF THE TENDER
- Category weight:  $\alpha_3 = 15\%$ This category deal with all the aspects related to project management and how the negotiation protocol has been taken into account.
- Reviewer Type weights:
  - "Technical Expert" reviewer type weight:  $\gamma_3' = 35\%$
  - "Domain Expert" reviewer type weight:  $\gamma_3'' = 35\%$
  - "External Expert" reviewer type weight:  $\gamma_3^{\prime\prime\prime}=30\%$

#### 3.3.2 Items

- Item  $I_{3,1}$  **Project plan** 
  - description: This item evaluates the extent to which the tender shows a clear plan for the development of a working solution, and whether it is a reasonable plan to finish phase 3 in time. It must be verified if the proposal respects of the negotiation protocol.



- *item weight*:  $\beta_{3,1} = 15\%$ 

#### • Item I<sub>3,2</sub> – Effectiveness management

- description: This item evaluates the effectiveness of the management.
- *item weight*:  $\beta_{3,2} = 15\%$
- Item I<sub>3,3</sub> Resource allocation
  - description: The extent to which the tenderer and/or subcontractor appear to have dedicated the resources (e.g. human capital, equipment etc.) necessary to perform the scope of the tender. It must be verified if the proposal respects of the negotiation protocol.
  - *item weight*:  $\beta_{3,3} = 15\%$
- Item I<sub>3,4</sub> Risk Assessment / Risk factors
  - description: The extent to which crucial risks (technical, commercial and other) to project success appear to be identified, and how effectively these will be managed. this item concerns the riskiness of a system and the acceptance of these risks.
  - *item weight*:  $\beta_{3,4} = 15\%$
- Item *I*<sub>3,5</sub> Negotiation Protocol
  - description: The extent to which the recommendations expressed in the negotiation protocol (see Appendix A) have been implemented in the project.
  - *item weight*:  $\beta_{3,5} = 40\%$

#### 3.4 Costs

#### 3.4.1 Description

This category concerns the financial aspects of a supplier.

- Category  $C_4 \text{COSTS}$
- Category weight:  $\alpha_4 = 15\%$ This category concerns the financial aspects of a supplier.
- *Reviewer Type weights*:
  - "Technical Expert" reviewer type weight:  $\gamma'_3 = 35\%$
  - "Domain Expert" reviewer type weight:  $\gamma_3'' = 35\%$
  - "External Expert" reviewer type weight:  $\gamma_3^{\prime\prime\prime}=30\%$



#### 3.4.2 Items

- Item  $I_{4,1}$  **Price/cost** 
  - *item weight*:  $\beta_{4,1} = 100\%$

#### 3.5 Overall Considerations

In figure 5, the distribution of category weights  $\alpha_i$  is shown.

You can notice that:

- the first two categories, i.e. "Impact on the Challenge" and "Technical Approach" determine 70% of the supplier score. In fact these categories together constitute the core aspects of a solution;
- both the categories "Quality of the Tender" and "Costs" determine 30% of the supplier score, in order to give the possibility of trading-off the impact on the challenge and the technical approach with other important factors.

In figure 6, the distribution of reviewer types weights  $\gamma_i$  is shown. Figure 6 gives an overview of the reviewer type weights, and you can notice that:

- *external experts* always account for a 30% of the score in order to compensate any possible bias and ensure a fair evaluation;
- *technical experts* determine 50% of the score for the "Technical Approach" category, where they are the most entitled to evaluate while they have the same weight of the domain experts when it comes to the "Quality of the Tender" and "Costs" categories;
- *domain experts* determine 40% of the score for the "Impact on the Challenge" category since they are the actual stakeholders of the solutions proposed by the suppliers.

## 4 Conclusions

This deliverable described the competitive evaluation strategy and its scoring model.

As far as the overall evaluation procedure of the first design phase is concerned (art. 1.2 of Framework Agreement), it will follow the same lines adopted for the evaluation of the PREFORMA tenders as described in [Lemmens et al., 2014] and here summarized:

- 1. individual evaluation phase by each type of expert;
- 2. consensus meeting among reviewers;
- 3. final ranking by the whole evaluation committee;
- 4. negotiation meeting with top ranked suppliers in order clarify any leftover doubts and/or adjust costs to the available budget.





Figure 5: Category weights comparison.



Figure 6: Dimensions weights comparison among categories.



### References

- IEC 60958 (2014). Digital audio interface Part 1: General. Standard IEC 60958-1 Ed. 3.1 b:2014.
- ISO 12234-2 (2001). Electronic still-picture imaging Removable memory Part 2: TIFF/EP image data format. Recommendation ISO 12234-2:2001.
- ISO 12639 (2004). Graphic technology Prepress digital data exchange Tag image file format for image technology (TIFF/IT). Recommendation ISO 12639:2004.
- ISO 19005-1 (2005). Document management Electronic document file format for long-term preservation Part 1: Use of PDF 1.4 (PDF/A-1). Recommendation ISO 19005-1:2005.
- ISO 19005-2 (2011). Document management Electronic document file format for long-term preservation Part 2: Use of ISO 32000-1 (PDF/A-2). Recommendation ISO 19005-2:2011.
- ISO 19005-3 (2012). Document management Electronic document file format for long-term preservation Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3). Recommendation ISO 19005-3:2012.
- ISO 32000-1 (2008). Document management Portable document format Part 1: PDF 1.7. Recommendation ISO 32000-1:2008.
- ISO/IEC 15444 (2004). Information technology JPEG 2000 image coding system: Core coding system. Recommendation ISO/IEC 15444-1:2004.
- ITU-T (1996). Methods for Subjective Determination of Transmission Quality. *ITU-T Reccomendation P.800*.
- Lemmens, B., Elfner, P., Lundell, B., Prandoni, C., and Fresa, A. (2014). Deliverable D2.2 – Tender Specifications. PREFORMA PCP Project, EU 7FP, Contract N. 619568. http://www.digitalmeetsculture.net/wp-content/uploads/2014/05/PREFORMA\_ D2.2\_Tender-Specifications\_v2.1.pdf.

## A Negotiation Protocol [Confidential]

Confidential.