Collaboration to Clarify the Cost of Curation

D4.4—Report on Risk, Benefit, Impact and Value

Deliverable Lead:	INESC-ID
Related Work package:	WP4
Author(s):	José Borbinha (INESC-ID) Raquel Bairrão (INESC-ID) Nuno Pradiante (INESC-ID) Ahmad Nadali (INESC-ID) Ricardo Vieira (INESC-ID) Diogo Proença (INESC-ID) Artur Caetano (INESC-ID) Raivo Ruusalepp (NLE) Neil Grindley (JISC)
Dissemination level:	Public
Submission date:	14 November 2014
Project Acronym:	4C
Website:	http://4cproject.eu
Call:	FP7-ICT-2011-9
Project Number	600471
Instrument:	Coordination and support actions (CSA)—ERA-NET
Start date of Project:	01 Feb 2013
Duration:	24 months

Project funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Version History

Version	Date	Changed pages / reason	Modified by
0.01	14 Apr 2014	Initial outline	JB
0.03	23 June 2014	New structure realigned with T4.5	JB
0.04	28 July 2014	New structure realigned with T4.1	JB
0.8	15 Sep 2014	New structure aligned with D3.2, D4.1 and D4.3.	DP
1.0	28 Sep 2014	Revised version for internal quality assurance.	JB
1.1	29 Sep 2014	Revised version for internal quality assurance	DP
1.2	30 Sep 2014	Last Revision prior to internal quality assurance.	DP
1.3	3 Oct 2014	Final Edit	PLSS
1.4	12 Nov 2014	Cases added, document revised for delivery	DP, RB
1.5	14 Nov 2014	Final Edit	PLSS

Acknowledgements

This report has been developed within the project “Collaboration to Clarify the Cost of Curation” (4cproject.eu). The project is an ERA-NET co-funded by the 7th Framework Programme of the European Commission.

The 4C participants are:

Participant organisation name	Short Name	Country
Jisc	JISC	UK
Det Kongelige Bibliotek, Nationalbibliotek Og Kobenhavns Universitetsbibliotek	KBDK	DK
Instituto de Engenharia de Sistemas e Computadores, Investigacao e Desenvolvimento em Lisboa	INESC-ID	PT
Statens Arkiver	DNA	DK
Deutsche Nationalbibliothek	DNB	DE
University of Glasgow	HATII-DCC	UK
University of Essex	UESSEX	UK
Keep Solutions LDA	KEEPS	PT
Digital Preservation Coalition Limited by Guarantee	DPC	UK
Verein Zur Forderung Der It-Sicherheit In Osterreich	SBA	AT
The University of Edinburgh	UEDIN-DCC	UK
Koninklijke Nederlandse Akademie van Wetenschappen -Knaw	KNAW-DANS	NL
Eesti Rahvusraamatukogu	NLE	EE

Author(s):	José Borbinha (INESC-ID) Raquel Bairrão (INESC-ID) Nuno Pradiante (INESC-ID) Ahmad Nadali (INESC-ID) Ricardo Vieira (INESC-ID) Diogo Proença (INESC-ID) Artur Caetano (INESC-ID) Raivo Ruusalepp (NLE) Neil Grindley (JISC)
Editor	Paul Stokes (Jisc)

Table of Contents

Acknowledgements 3
Table of Contents 5
Table of Figures 7
Executive Summary 8
1 Introduction 10
2 Fundamentals 11
2.1 Risk Management 11
2.1.1 Main Risk Management Definitions 12
2.1.2 Risk Management and Digital Curation 17
2.2 Business Model Canvas 17
3 A Method to Estimate Costs of Curation Focusing on Controls 21
4 Risk Identification and Analysis with the Indirect Economic Determinants 24
5 Risk Identification and Analysis with the Business Model Canvas 29
5.1 Business Model Canvas Details 33
5.1.1 Key Activities 33
5.1.2 Value Propositions 35
5.1.3 Customer Relationships 36
5.1.4 Key Resources 37
5.1.5 Channels 37
5.2 Business Model Canvas Risk Questions 37
5.2.1 Key Partnerships (KP) 38
5.2.2 Key Activities (KA) 38
5.2.3 Value Propositions (VP) 38
5.2.4 Customer Relationships (CR) 38
5.2.5 Customer Segments (CS) 38
5.2.6 Key Resources (KR) 38
5.2.7 Channels (CH) 38
5.2.8 Cost Structure (C$) 39
5.2.9 Revenue Streams (R$) 39
5.3 Generic Risks and Controls for the Generic BMC 39
6 Risk Identification and Analysis of the Case Studies 43
6.1 Case Study: LNEC 43
6.1.1 Identification of the relevant Indirect Economic Determinants 43
6.1.2 Analysis of the Indirect Economic Determinants 45
6.1.3 Analysis of the Business Model Canvas 46
6.2 Case Study: Portuguese Web Archive (PWA) 52
6.2.1 Identification of the relevant Indirect Economic Determinants 53
6.2.2 Analysis of the Indirect Economic Determinants 54
6.2.3 Analysis of the Business Model Canvas 56

7 Conclusions 61
8 References 62

Table of Figures

Figure 1—The Risk Management Process according to the ISO/FDIS 31000. 12

Figure 2—The UML domain model of a risk management scenario. 13

Figure 3—Conceptual map showing controls as the cost entities in a risk management perspective. 13

Figure 4—Typical Business Model Canvas 18

Figure 5—Business Model Risk Canvas 20

Figure 6—BPMN diagram of the pragmatic method to estimate costs of curation focusing on risks and controls 22

Figure 7—OAIS Functional Entities ([20], Page 4-1) 30

Figure 8—Generic BMC based on OAIS 32

Figure 9—Portuguese Civil Engineering National Laboratory BMC 46

Figure 10- Portuguese Web Archive BMC 56

Tables

Table 1—BPMN core elements used in this document 22
Table 2—A generic risk identification and analysis of the indirect economic determinants. 25
Table 3—Examples of relevant references for the generic controls. 28
Table 4—Key activities identification 34
Table 5—Value propositions identification 35
Table 6—Customer relationships identification 36
Table 7—Key resources identification 37
Table 8—Channels identification 37
Table 9—Generic controls identification 42
Table 10—Relevance of the Indirect Economic Determinants for the case of LNEC 44
Table 11—A risk identification and analysis of the indirect economic determinant for the case of the LNEC 45
Table 12—Key Partners Identification 47
Table 13—Value Propositions Identification 47
Table 14—Customer Relationships Identification 47
Table 15—Customer Segment Identification 48
Table 16—Key Resources Identification 48
Table 17—Channels Identification 48
Table 18—Cost Structure Identification 49
Table 19—Revenue Streams Identification 49
Table 20—Examples of relevant controls for the LNEC risks identified from the BMC. 51
Table 21—Relevance of the Indirect Economic Determinants for the case of PWA 54
Table 22—A risk identification and analysis of the indirect economic determinant for the case of the PWA 55
Table 23—Key Partners Identification 57
Table 24—Value Propositions Identification 57
Table 25—Customer Segment Identification 57
Table 26—Cost Structure Identification 58
Table 27—Revenue Streams Identification 58
Table 28—Examples of relevant controls for the PWA risks identified from the BMC 60

Executive Summary

This deliverable of the 4C project work package "WP4—Enhancement" proposes a pragmatic method for estimating costs of digital curation in two different scenarios:

a) **A "Current" scenario**, where the costs of controls already exist in the repository as a means to reduce the impact of the consequence of a risk/threat, change the likelihood of an event, or reduce the exposure to a vulnerability;
b) **A "Future" scenario**, where the costs of controls do not yet exist, but where repository managers are able to consider alternative scenarios of repository governance.

The foundations of this method draw from relevant sources, such as the ISO31000 standard and the Business Model Canvas (BMC)[1]; the deliverable D4.1 where the Indirect Economic Determinants are described and analysed; the deliverableD4.5 where the BMC is presented; the deliverable D4.3 where literature sources regarding trustworthiness and control are described and analysed; and the deliverable D3.2, which describes the cost concept model and gateway specification.

The purpose of this method is to make good use of established risk management concepts to raise awareness of repository costs. The technique behind this method analyses the repository with the support of a risk registry and is based on:

(1) A risk analysis of the indirect economic determinants (from "D4.1—A Prioritised assessment of the indirect economic determinants of digital curation"), complemented by a specialised analysis using as reference the "D4.3—Quality and trustworthiness as economic determinants in digital curation";
(2) A risk analysis of a BMC for digital curation (based on the results of task T4.5).

Examples of the method are presented in the form of two case studies.

This deliverable will be complemented in the future by the following extra mini-deliverables that will shape a toolset aiming to help guide the application of the method, namely:

- A generic BMC, with an associated generic registry of risk questions—a registry of typical risk questions that are relevant for the domain of digital curation and can help to define a specific set of risk questions for a specific BMC;
- A risk registry for digital curation—a registry of risks derived from the previous risk questions, and also common related controls, relevant for the domain of digital curation;
- A set of case studies that detail the application of the method to real repositories.

These mini-deliverables will be draw upon several strands of the project—some that are yet to be completed—and will reuse results from all the other activities. As such, they will be developed independently and presented as appendices to be integrated in the Curation Costs Exchange (CCEx) platform.

[1] The BMC is a model used in strategic management to document existing business models and develop new ones [10]. Generally BMCs are filled in at a brainstorming session. In this way a group of people can generate a relevant understanding of their business model. More details on the BMC technique in the 4C project will be reported in deliverable D4.5.

1 Introduction

In the Description of Work (DoW) this deliverable is described as follows:

> *"Working closely with T4.1 - this report will look at a range of inter-related issues from a perspective of Risk Management. The principal trade-off between costs is obviously benefit but if that can be measured with some objective relevance in some sectors of activity in the corporate world, in various contexts and for other different types of organisations, this can be a very complicated equation.* ***Using case studies, the role of risk and risk assessment will be considered in relation to curation as one of the principal drivers for governance****. In that sense, not only cost but also benefit, impact and value (and its relation to cost efficiency) will also be considered terminologically and by sector to try and characterise the influence of these factors as determinants. For example, one of the cost and risk factors that will be specifically looked into is the issue of loss and recovery from loss, as opposed to preventive curatorial action. In common with T4.3, attention will be paid to the criteria specified in the information dependency profile (T3.1) to optimise the value of these reports for use by the Assessment group."*

Risk can be defined as the "effect of uncertainty on objectives"[2]. With this in mind, and assuming that dealing with uncertainty is one of the important concerns of a digital curation stakeholder, the hypothesis that the body of knowledge on Risk Management (RM) can be valuable to the domain of digital curation sounds reasonable. RM involves establishing an appropriate infrastructure and organisational culture that allows for the application of a logical and systematic method of establishing the context; identifying; analysing; evaluating; treating; monitoring; and communicating risks associated with any activity, function, or process in a way that will enable organisations to minimize losses and maximise gains[4].

A core assumption of the work carried out in this task is that costs must be traded for controls. In this context a control is something used either to minimize negative impacts (in other words mitigate risks) or to take advantage of opportunities to produce value and thus bring gains. Helping repository managers to be aware of the controls being applied (or of the controls that should be applied) to avert risks is a contribution to understanding the costs of the business.

A formal risk assessment process in an organisation can be a complex business as it aims to identify all the necessary controls in a specific situation. Such a task often entails using experts to apply specialised knowledge about the multiple domains the organisation depends on. To address risk assessment within the curation domain, we propose to organise the generic common body of knowledge relating to digital curation as a framework. This generic framework can be later optimised as a means to address the specific needs of a specific sub-domain. Taken as a whole, the method aims to raise awareness of the costs associated with operating a repository and is grounded on:

- A generic risk analysis of the indirect economic determinants (reported in the deliverable "D4.1—A prioritised assessment of the indirect economic determinants of digital curation"), complemented by a specialised analysis from the deliverable "D4.3—Quality and trustworthiness as economic determinants in digital curation".
- A generic risk analysis based on a Business Model Canvas for the digital curation domain (based on the work carried out in the task T4.5)

Besides describing the approach, this report also shows the optimisation of the method when applied to the specific domains of real case studies.

2 Fundamentals

This section describes the fundamentals of risk management and the concepts behind the Business Model Canvas.

2.1 Risk Management

The main references on Risk Management (RM) from the International Organisation for Standardisation (ISO) are:

- ISO Guide 73: Vocabulary for risk management[4];
- ISO 31000: Risk management principles and guidelines[2];
- ISO 31004: Risk management—Guidance for the implementation of ISO 31000[5];
- IEC 31010: Risk assessment techniques[3].

According to those sources, organisations (that find RM relevant to their governance) should define an internal RM process taking as a starting point the generic method proposed in ISO 31000[2] (illustrated in Figure 1 below). IEC 31010 catalogues a set of techniques for risk assessment[3].

From ISO 31004 we learn that:

> "Organisations of all kinds face internal and external factors and influences that make it uncertain whether, when and the extent to which, they will achieve or exceed their objectives. The effect that this uncertainty has on the organisation's objectives is risk. (…)
>
> Likelihood is not just that of an event occurring, but the overall likelihood of experiencing the consequences that flow from the event, and the magnitude of the consequence in either positive or negative terms. Typically, there can be a range of possible consequences that can flow from an event, and each will have its own likelihood. The level of risk can be expressed as the likelihood that particular consequences will be experienced (including the magnitude). Consequences relate directly to objectives and they arise when something does or does not happen.
>
> Risk is the effect of uncertainty on objectives, regardless of the domain or circumstances, therefore an event or a hazard (or any other risk source) should not be described as a risk. Risk should be described as the combination of the likelihood of an event (or hazard or source of risk) and its consequence.
>
> The understanding that risk can have positive or negative consequences is a central and vital concept to be understood by management. Risk can expose the organisation to either an opportunity, a threat or both. (…)
>
> **Controls are measures implemented by organisations to modify risk that enable the achievement of objectives**. Controls can modify risk by changing any source of uncertainty (e.g. by making it more or less likely that something will occur) or by changing the range of possible consequences and where they may occur."

Even if we are not following a specific RM method as part of the governance framework of a repository, we cannot avoid having to deal with the identification of risks and controls. However, as a complete RM methodology can be complex and expensive to implement, we are here proposing a simplified method that can be used at least for a preliminary phase of costs estimation. If, after the application of this method, the stakeholders of a repository feel the RM principles are valuable for the governance in their

case and it is worthy to consider a proper and full RM method, then these preliminary results can be reused for that purpose.

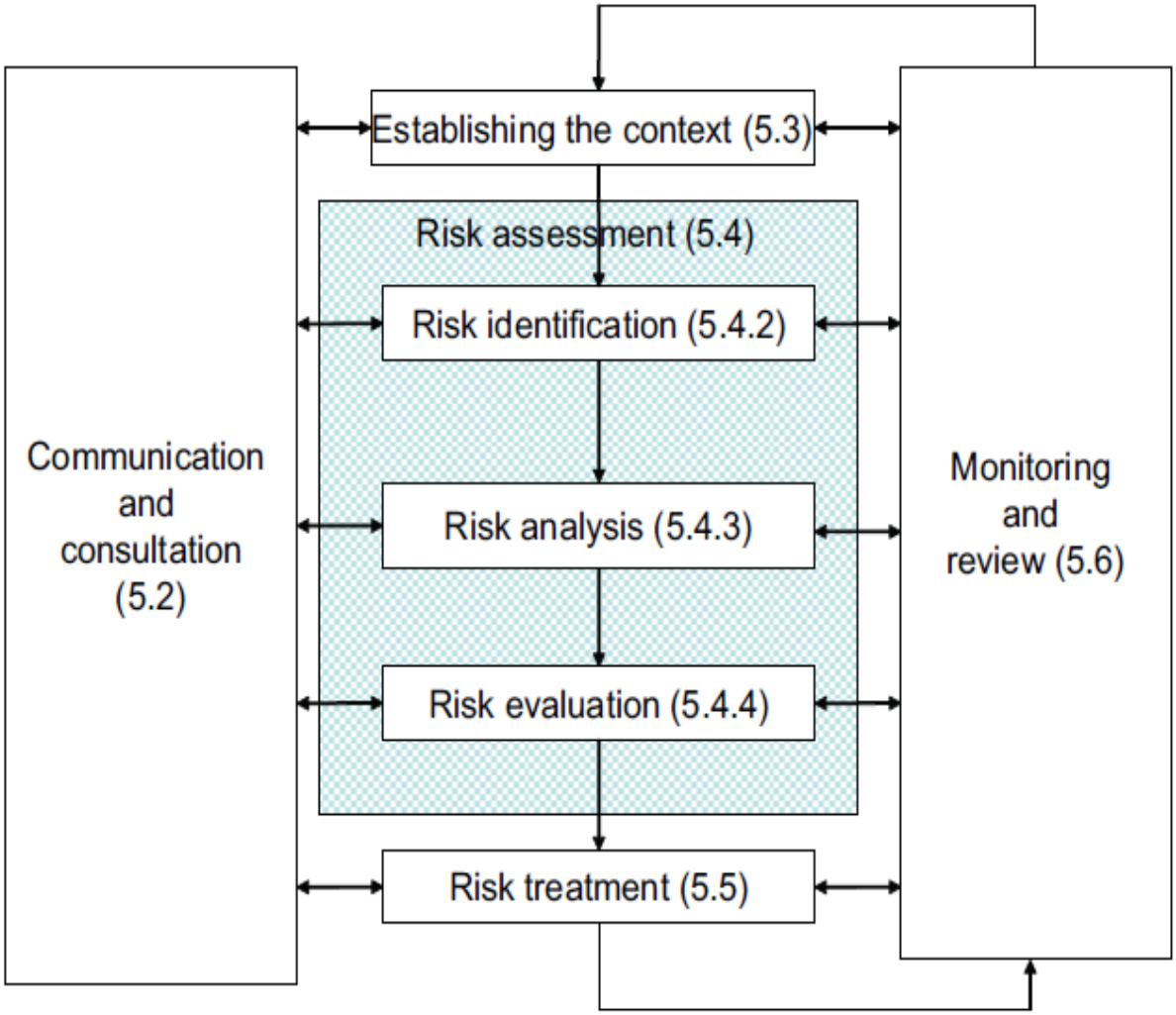


Figure 1—The Risk Management Process according to the ISO/FDIS 31000.

2.1.1 Main Risk Management Definitions

The definitions for risk management are defined in the ISO Guide 73 [4]. Figure 2 and Figure 3 provide two views of Risk Management, the first as a Unified Modelling Language (UML) domain model, and the second as a conceptual map.

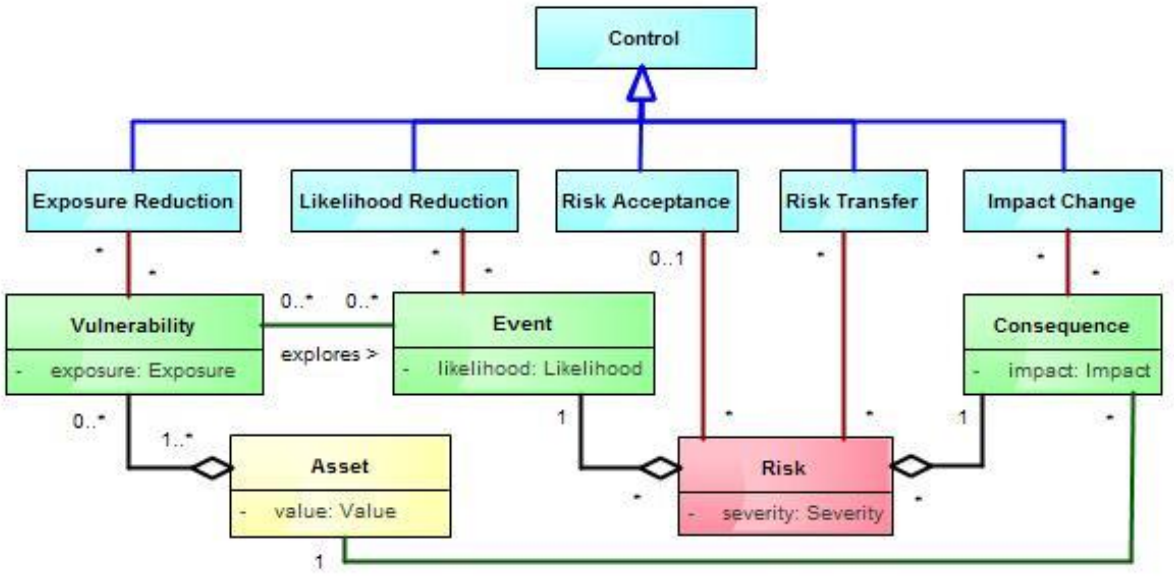


Figure 2—The UML domain model of a risk management scenario.

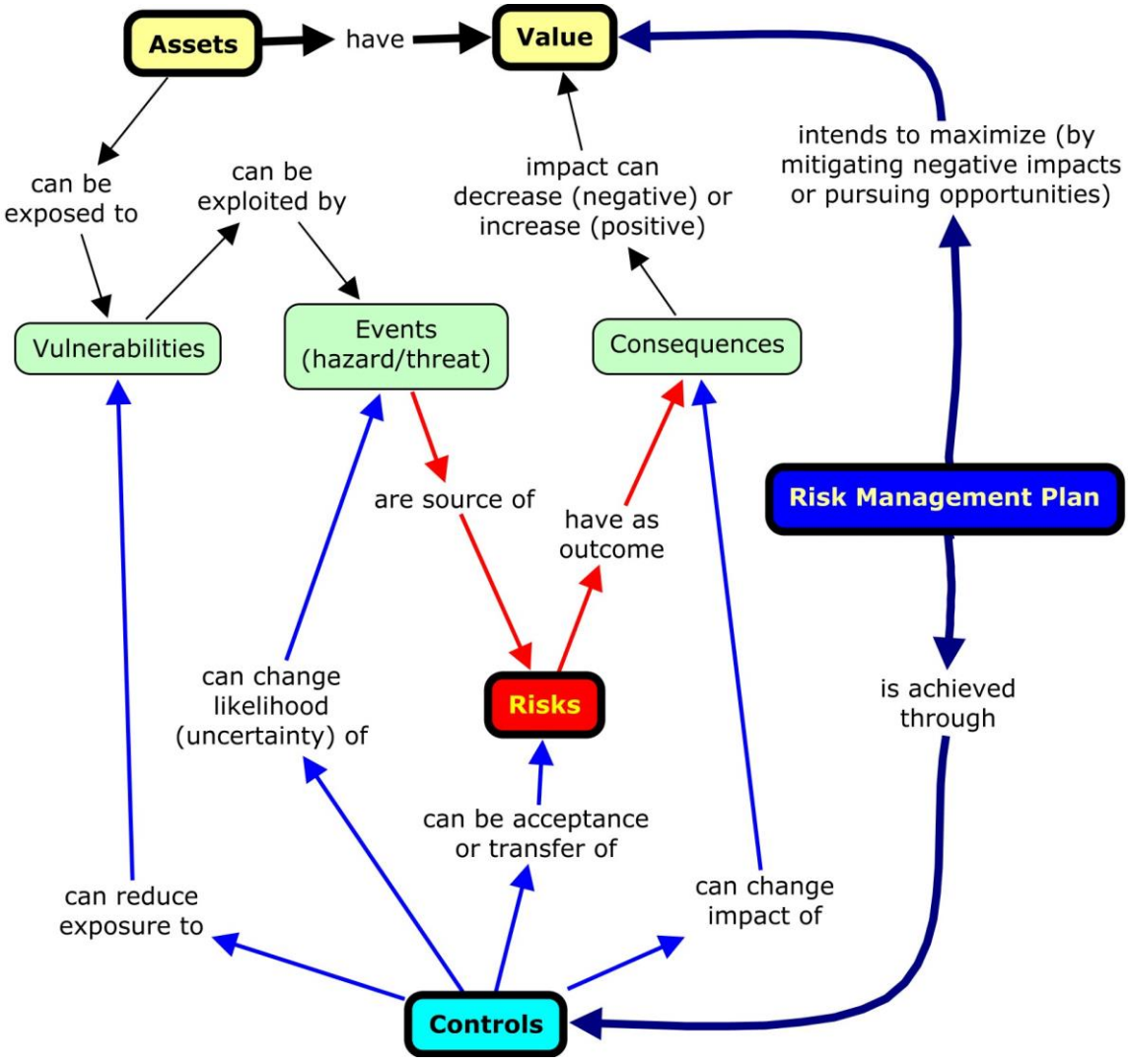


Figure 3—Conceptual map showing controls as the cost entities in a risk management perspective.

Definition 1.1: **Risk**

Risk—*effect of uncertainty on objectives*

NOTE 1—An effect is a deviation from the expected—positive and/or negative.

NOTE 2—Objectives can have different aspects (such as financial, health and safety, and environmental goals) and can apply at different levels (such as strategic, organisation-wide, project, product and process).

NOTE 3—Risk is often characterised by reference to potential **Events**(1.3) and **Consequences**(1.5), or a combination of these.

NOTE 4—Risk is often expressed in terms of a combination of the consequences of an event (including changes in circumstances) and the associated **Likelihood**(1.4) of occurrence.

NOTE 5—Uncertainty is the state, even partial, of deficiency of information related to, understanding or knowledge of, an event, its consequence, or likelihood.

Definition 1.2: **Control**

Control—*a measure that modifies* ***Risk****(1.1)*

NOTE 1—Controls include any process, policy, device, practice, or other actions which modify risk.

NOTE 2—Controls may not always exert the intended or assumed modifying effect.

Definition 1.3: **Event**

Event—*occurrence or change of a particular set of circumstances*

NOTE 1—An event can be one or more occurrences, and can have several causes.

NOTE 2—An event can consist of something not happening.

NOTE 3—An event can sometimes be referred to as an “incident” or “accident”.

NOTE 4—An event without consequences can also be referred to as a “near miss”, “incident”, “near hit” or “close call”.

Definition 1.4: **Likelihood**

Likelihood—*chance of something happening*

NOTE 1—In risk management terminology, the word “likelihood” is used to refer to the chance of something happening, whether defined, measured or determined objectively or subjectively, qualitatively or quantitatively, and described using general terms or mathematically (such as a probability or a frequency over a given time period).

NOTE 2—The English term “likelihood” does not have a direct equivalent in some languages; instead, the equivalent of the term “probability” is often used. However, in English, “probability” is often narrowly interpreted as a mathematical term. Therefore, in risk management terminology, “likelihood” is used with the intent that it should have the same broad interpretation as the term “probability” has in many languages other than English.

Definition 1.5: **Consequence**

Consequence—*outcome of an* ***Event****(1.3) affecting objectives*

NOTE 1—An event can lead to a range of consequences.

NOTE 2—A consequence can be certain or uncertain and can have positive or negative effects on objectives.

NOTE 3—Consequences can be expressed qualitatively or quantitatively.

NOTE 4—Initial consequences can escalate through knock-on effects

Definition 1.6: **Risk profile**

Risk profile *description of any set of* ***Risks****(1.1)*

NOTE—The set of risks can contain those that relate to the whole organisation, part of the organisation, or as otherwise defined.

Definition 1.7: **Risk treatment**

Risk treatment—*process to modify* ***Risk****(1.1)*

NOTE 1—Risk treatment can involve:

a) avoiding the risk by deciding not to start or continue with the activity that gives rise to the risk;

b) taking or increasing risk in order to pursue an opportunity;

c) removing the risk source;

d) changing the **Likelihood** (1.4);

e) changing the **Consequences** (1.5);

f) sharing the risk with another party or parties (including contracts and risk financing); and

g) retaining the risk by informed choice.

NOTE 2—Risk treatments that deal with negative consequences are sometimes referred to as “risk mitigation”, “risk elimination”, “risk prevention” and “risk reduction”.

NOTE 3—Risk treatment can create new risks or modify existing risks.

Definition 1.8: **Risk identification**

Risk identification—*process of finding, recognizing and describing* ***Risks****(1.1)*

NOTE 1—Risk identification involves the identification of risk sources, **Events** (1.3), their causes and their potential **Consequences** (1.5).

NOTE 2—Risk identification can involve historical data, theoretical analysis, informed and expert opinions, and **Stakeholder's** (1.10) needs.

Definition 1.9: **Risk assessment**

Risk assessment—*overall process of* ***Risk identification****(1.8), risk analysis and risk evaluation.*

Definition 1.10: Stakeholder

Stakeholder*—person or organisation that can affect, be affected by, or perceive themselves to be affected by a decision or activity*

NOTE—A decision maker can be a stakeholder.

Definition 1.11: Risk analysis

Risk analysis*—process to comprehend the nature of* ***Risk****(1.1) and to determine the level of risk*

NOTE 1—Risk analysis provides the basis for risk **Evaluation**(1.12) and decisions about **Risk treatment**(1.7).

NOTE 2—Risk analysis includes risk estimation.

Definition 1.12: Risk evaluation

Risk evaluation*—process of comparing the results of* ***Risk analysis****(1.11) with risk criteria to determine whether the* ***Risk****(1.1) and/or its magnitude is acceptable or tolerable*

NOTE—Risk evaluation assists in the decision about **Risk treatment**(1.7).

Definition 1.13: Risk management

Risk Management*—coordinated activities to direct and control an organisation with regard to* ***Risk****(1.1)*

Definition 1.14: Risk management plan

Risk Management Plan*—scheme within the* ***Risk management framework****(1.15) specifying the approach, the management components and resources to be applied to the management of* ***Risk****(1.1)*

NOTE 1—Management components typically include procedures, practices, assignment of responsibilities, sequence and timing of activities.

NOTE 2—The risk management plan can be applied to a particular product, process and project, and part or whole of the organisation.

Definition 1.15: Risk management framework

Risk management framework*—set of components that provide the foundations and organisational arrangements for designing, implementing, monitoring, reviewing and continually improving* ***Risk management****(1.13) throughout the organisation*

NOTE 1—The foundations include the policy, objectives, mandate and commitment to manage **Risk**(1.1).

NOTE 2—The organisational arrangements include plans, relationships, accountabilities, resources, processes and activities.

NOTE 3—The risk management framework is embedded within the organisation's overall strategic and operational policies and practices.

2.1.2 Risk Management and Digital Curation

"Digital curation involves maintaining, preserving and adding value to digital research data throughout its lifecycle. The active management of research data reduces threats to their long-term research value and mitigates the risk of digital obsolescence"[19]. In simple terms, we can claim that the main objective of a repository is the curation of assets, which in their primary form are data sets but also can be services directly related to those data sets, in secondary or derived forms.[2]

According to the concepts outlined in section 2.1.1, one can argue that costs are what we have to give up for controls, which in turn are the measures that we have to put in practice to minimise loss (digital preservation) or to maximise gain (digital curation in a broad sense). In that sense, a control is anything we are considering to apply to either minimise negative impacts (mitigate risks by modifying threats, consequences or likelihood) or to take advantage of opportunities to produce value and thus bring gains.

However, we must also agree that, in most of the usual digital curation scenarios, it is usually very difficult to estimate the *absolute* value of an asset. For that reason, we are here ignoring the measurement of value[3], and focusing only in the identification of controls as the source of costs.

2.2 Business Model Canvas

The Business Model Canvas (BMC) is a model used in strategic management to document existing business models and develop new ones[10]. A BMC comprises nine building blocks that describe an organization, as illustrated in the Figure 4 below[4].

The BMC is designed to allow a group of people to fill it in through brainstorming sessions and thus create a relevant understanding of their business model. At the end of such a process each block must have at least one shared assumption about the business. It is even possible to develop more than one BMC in order to represent different, alternative views of the business. More information on the BMC in the 4C project will be found in deliverable D4.5—From Costs to Business Models.

The BMC was first proposed in Osterwalders thesis ("The Business Model Ontology—A Proposition in a Design Science Approach")[11]. After that, several authors developed or adopted this canvas approach for other purposes, such as, the Lean canvas[16]. In the meantime it has been suggested that doing a BMC exercise is already in some sense performing a risk assessment[17] [18].

[2] Note: More sophisticated business models can be considered, which might justify a link of this task to the task T4.5—From costs to business models, where these issues are expected to be addressed.

[3] This issue is to be explored (and validated) in the task T4.5.

[4] Image source: Alexander Osterwalder—Business Model Alchemist—
http://www.businessmodelgeneration.com/downloads/business_model_canvas_poster.pdf

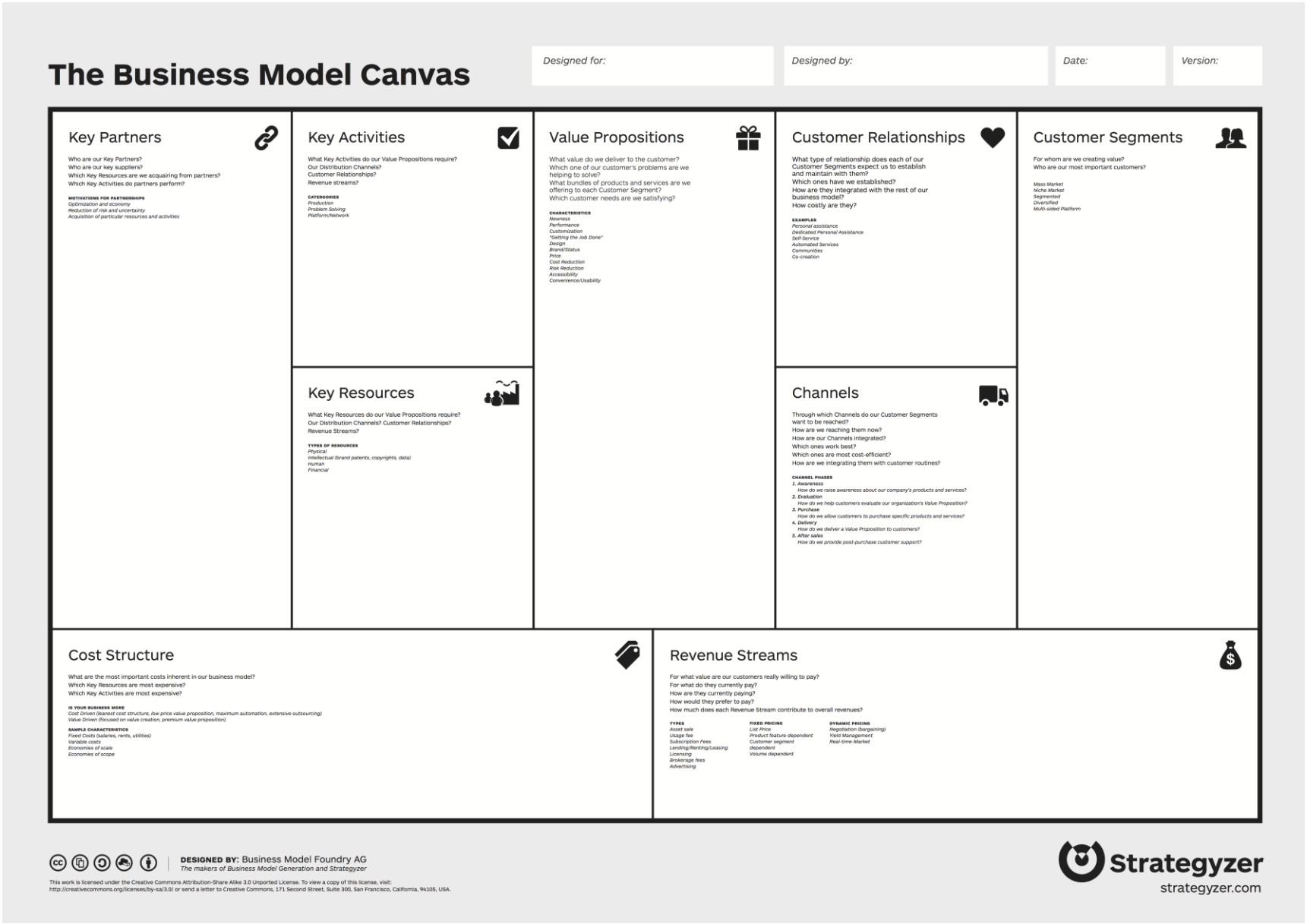


Figure 4—Typical Business Model Canvas

Some authors have gone even further and proposed the hypothesis that the BMC concept can even be extended to support a pragmatic risk analysis. This is illustrated in Figure 5 which is taken from Schliemann's article "BMI? Of course, but what about the Model Risks?"[12] in which he scopes the business model risk canvas. The motivation behind it is to understand both what can positively affect the value propositions of your business (opportunities) and what can negatively affect those same value propositions (risks).

The idea is to identify and understand the risks and their impact (positive and negative) on each of the nine building blocks of the BMC, as well as the risk appetite of the stakeholders upon which a business depends—stake holders in this context such as, regulators and investors. There is a huge body of knowledge from the risk management community on how to assess and measure risk through analytical tools but this new technique fills the need to introduce risk assessment at a higher level, scoping it visually in consideration for each of the building blocks of the BMC.

When applying this technique to identify the risks and their impact there should be a series of risk-related questions for each of the nine building blocks of BMC. Simple examples of these questions are proposed in the original business model risk canvas, but for real use these should be scoped for the business in question.

KPr

Who are our Key Partners?
Who are our key suppliers?
Which Key Resources are we acquiring from partners?
Which Key Activities do partners perform?

What are the effects of unavailability of a KP or a KP sourced asset?
What woud determine you might face a Partner loss?
What can affect your relationship with your KP? How and How much?
Would a talent-knowledge loss affect your viability or competence? What could trigger it?
How would liabilities or a reputational impact would affect KP support?
What impact can generate a legal or regulatory situation? How do you ensure compliance?

KAr

What Key Activities do our Value Propositions require?
Our Distribution Channels?
Customer Relationships?
Revenue streams?

What are the immediate and long term effects of unavailability of a KA?
What woud determine you might face a KA loss?
Can security vulnerabilities, bribey or fraud affect your KA? How and How much?
Would a talent-knowledge loss affect your viability or competence? What could trigger it?
How much would affect an erosion in internal and external quality? What value levers?
What impact can generate a legal or regulatory situation? How do you ensure compliance?

KRr

What Key Resources do our Value Propositions require?
Our Distribution Channels? Customer Relationships?
Revenue Streams?

How would a shortage on key reources impact your operation?
What woud determine you might face a key resource shortage?
Can security vulnerabilities, bribey or fraud affect your sourcing? How and How much?
Would a talent-knowledge loss affect your viability or competence? What could trigger it?

VPr

What value do we deliver to the customer?
Which one of our customer's problems are we helping to solve?
What bundles of products and services are we offering...?
Which customer needs are we satisfying?

How satisfaction is measured? What can cause degradation and how it would that impact?
Can security vulnerabilities, bribey or fraud affect your VP? How and How much?
Would a talent-knowledge loss affect your viability or competence? What could trigger it?
Are your internal policies and controls aligned to your VP? How do you measure it?
How is feedback routed back and internalized? Is the VP still valuable?

CRr

What type of relationship does each of our Customer Segments expect us to establish and maintain with them?
Which ones have we established?
How are they integrated with the rest of our business model?
How costly are they?

What events or factors would impact our relationship with them?
Do you have a clear understanding of their emotional experience with your brand?

CSr

For whom are we creating value?
Who are our most important customers?
What events or factors would erode their perceived value?

CHr

Through which Channels do our Customer Segments want to be reached?
How are we reaching them now?
How are our Channels integrated?
Which ones work best?
Which ones are most cost-efficient?
How are we integrating them with customer routines?

What are the potential impacts of outages in the service delivery?
What are the key factors or vulnerabilities ending in a potential service outage?
What is your dependency vs. authonomy of your channles?
What is your relative position (leverage) over the various channels?

C$r

What are the most important costs inherent in our business model?
Which Key Resources are most expensive?
Which Key Activities are most expensive?

What is the potential impact of variations (price and availbility) in the most important cost components?
What conditions in each of your markets can potentially impact your cost streams?
Can security vulnerabilities, bribey or fraud affect your cost? How and How much?
How would liabilities or a reputational impact affect your Costs Stream?

R$r

For what value are our customers really willing to pay?
For what do they currently pay?
How are they currently paying?
How would they prefer to pay?
How much does each Revenue Stream contribute to overall revenues?

What market events or competitor dynamics can impact what customers will be willing to pay?
What conditions in each of your markets can potentially impact your revenue streams?
Can security vulnerabilities or fraud affect your revenues? How and How much?

Figure 5—Business Model Risk Canvas

3 A Method to Estimate Costs of Curation Focusing on Controls

This section describes a method for estimating costs of curation in two different scenarios:

c) **"Current" scenario**, where the costs of controls already exist in the repository as a means to reduce the impact of a consequence of a risk, change the likelihood of an event, or reduce the exposure to a vulnerability;
d) **"Future" scenario**, where the costs of controls do not yet exist, but where repository managers are able to consider alternative scenarios of repository governance.

The foundations of this method draw from relevant sources, such as the ISO 31000 standard and the Business Model Canvas (BMC), as well as from the results presented in other deliverables, namely:

- Deliverable "D4.1—A Prioritised Assessment of Indirect Economic Determinants", which describes and analyses the Indirect Economic Determinants,
- Deliverable "D4.5—From Costs to Business Models", which describes the BMC,
- Deliverable "D4.3—Report on Trustworthiness and Quality", which summarizes and analyses sources on trustworthiness and control, and
- Deliverable "D3.2—A Cost Concept Model & Gateway Requirement Specification", which describes the Cost Concept Model and the Gateway Specification.

The core stages of the method are:

1. **Define the Context**: Define the requirements of the main elements—as proposed in D3.2: the organisation (mission, etc.); the assets (data and services), and the external stakeholders—and, for each of these elements:
   1.1. Identify the relevant economic determinants.
   Identify from the 4C catalogue the relevant economic determinants that apply (from D4.1).
   Identify eventual extra determinants that might also be relevant for the scenario
   1.2. Define the BMC for the scenario (as recommended by D4.5).

2. **Execute a Pragmatic Risk Assessment**: Use a risk repository, or consult experts, in order to:
   2.1. Identify relevant risks associated with the identified determinants.
   2.2. Identify relevant risks associated with the BMC.

3. **Recognise Actual Risk Treatment** (the "Current" scenario):
   3.1. Consolidate the risks identified (mainly, to detect repetitions and overlaps).
   Note: This is probably the best stage to identify potential positive impacts (if the identification of positive impacts is desired).
   3.2. Use internal information, and (if necessary) also consult a risk repository or experts, to identify the controls to apply for the consolidated risks.
   3.3. Estimate the costs for these controls (the ideal is to calculate these costs precisely, however, best estimates can also be useful).

4. **Simulate Alternative Risk Treatments** (an optional activity, to be executed as many times as needed, to explore possible alternative "Future" scenarios):
   4.1. Use internal information, eventually also consulting a risk repository or experts, and according to the businesses strategic view and governance rules, conceive alternative scenarios for controls of the identified risks.
   Note: This is probably the best stage to explore opportunities to exploit positive impacts (if the exploitation of positive impacts is desired).
   4.2. Make your best estimate for the costs of this new scenario.

Steps 1 to 4 are illustrated in Figure 6 in the form of a business process diagram (expressed in the BPMN – Business Process Modelling Notation language, using the set of core elements described in Table 1).

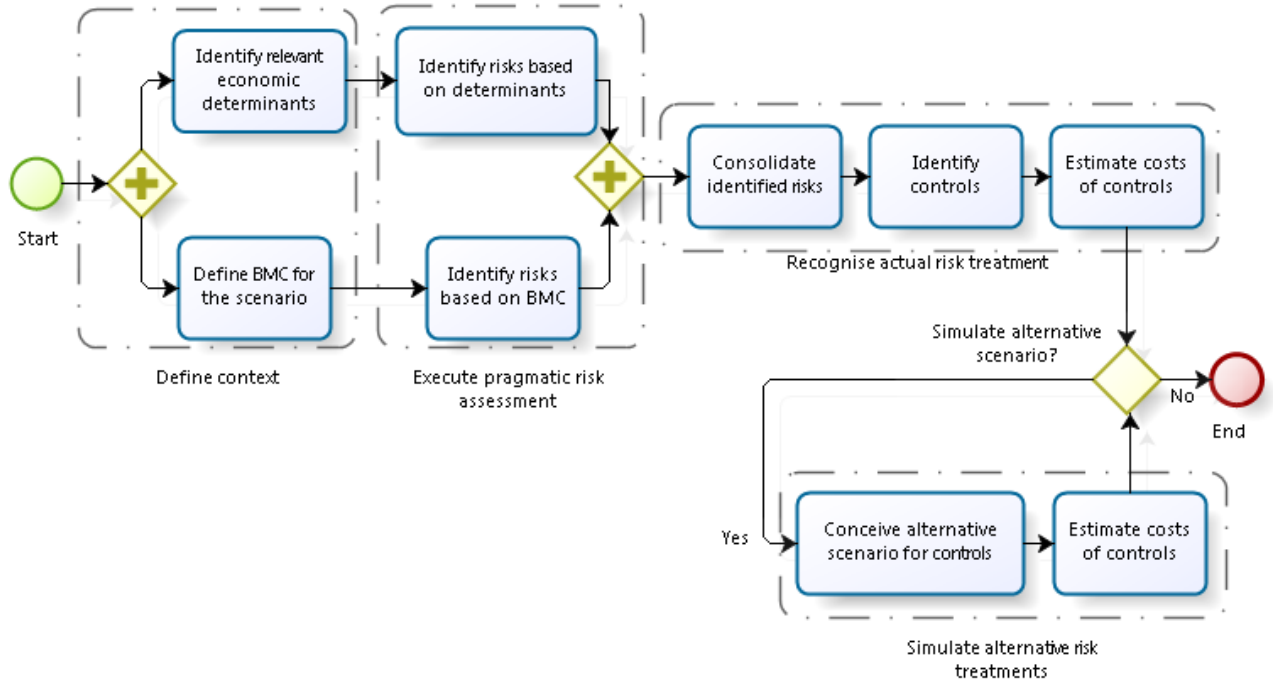


Figure 6—BPMN diagram of the pragmatic method to estimate costs of curation focusing on risks and controls

Name	Icon	Description
Start Event		An Event that indicates where a particular Process starts.
Sequence Flow		A connecting object that shows the order in which activities are performed in a Process and is represented with a solid graphical line
Task	Task 1	An atomic activity that is included within a Process.
Parallel Gateway		A parallel gateway is used to represent two concurrent tasks in a flow.
Group		An informal grouping of elements (only to help visualization).
Exclusive Gateway		Performs exclusive decisions.
End Event		An Event that indicates where a path in the process will end. In terms of Sequence Flows, the End Event ends the flow of the Process.

Table 1—BPMN core elements used in this document

There are two supporting techniques that can be used to practically identify the most relevant risks and related controls for digital curation:

- An analysis through the Indirect Economic Determinants (detailed below in Section 4)
- An analysis through a BMC with Risk (detailed below in Section 5).

4 Risk Identification and Analysis with the Indirect Economic Determinants

As reported in deliverable D4.1, "the initial list of indirect economic determinants was compiled through consultation with experts from the project consortium and has in the course of work been expanded through discussions with stakeholder groups and the project advisory board."

Table 2 shows the determinants that resulted from these activities, and also their perceived relation to the concepts of risk management, resulting from the analysis of these sources:

- D4.1—A prioritised assessment of the indirect economic determinants of digital curation;
- DRAMBORA—Digital Repository Audit Method Based on Risk Assessment[1].

Determinant	Related risks (possible consequence in case of hazard)	Generic controls (source of costs)
Authenticity	Loss of reputation and trust	Preservation plan
Benefit	Ability to deliver	Business plan
Confidentiality	Exposure to competitors	Security auditing Security certification
Efficiency	Exposure to financial uncertainty	Performance assessment Re-engineering / Change management Operations Maintenance Infrastructure
Flexibility	Inability to explore new opportunities	Re-engineering / Change management
Impact	Loss of reputation and trust	Marketing plan
Innovation	Exposure to obsolesce Uncertainty of early adopter Inability to explore new opportunities	Research and development Re-engineering / Change management
Interoperability	Ability to deliver	Operations
Quality	Quality of service	Quality auditing Quality certification
Reputation	Reputation	Marketing plan
Risk	Risk Management	Risk management plan
Sensitivity	Liability in sensitive data breaches Loss of reputation and trust	Security auditing Security certification

Determinant	Related risks (possible consequence in case of hazard)	Generic controls (source of costs)
Skills	Loss of efficiency if key staff leave	Staff assessment Staff training Staff salaries/benefits
Sustainability	Exposure to financial uncertainty	Business plan Marketing plan
Transparency	Loss of reputation and trust	Marketing plan
Trustworthiness	Loss of reputation and trust	Trustiness auditing Trustiness certification
Value	Ability to deliver	Business plan

Table 2—A generic risk identification and analysis of the indirect economic determinants.

This list only provides a generic view of the related risks and generic controls based on the indirect economic determinants from D4.1. However, as stated in Annex I of D3.2, these determinants are "context and stakeholder dependent," so each organisation must check for each of them if they are relevant and, in they are, also define each determinant in context. Based on the result of this exercise, the risk and controls must then be identified.

Table 3 presents examples of specific controls for the generic controls listed in Table 2. These references were taken from delivery "D4.3— Quality and trustworthiness as economic determinants in digital curation", and from the following sources (several domains were researched, such as IT Management and Governance, Quality Assessment, Digital Preservation and Information Management, etc.):

- **Fundamental sources:**
  - Business Model Canvas (BMC)[10];
  - ISO 9001, an international standard for quality management net systems that details the requirements for such systems[6];
  - ISO 25010, from systems and software engineering detailing the system quality models[7];
  - Control Objectives for Information and Related Technology (COBIT), which is a framework from the IT Governance Institute for governance and management of IT in organisations[13];
  - ISO 31000 family of standards for risk management[2];
  - Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)[1].
- **Sources of our main domain (digital curation and preservation), relevant for analysis or usage for controls for identified risks:**
  - PLATO[5], a preservation planning tool;
  - ISO 16363, a standard for the audit and certification of trustworthy digital repositories[9];

[5] http://www.ifs.tuwien.ac.at/dp/plato/intro/

  - Data Seal of Approval[6] (DSA), which offers an online tool that assess and certifies repositories on their data policies.
- **Other potentially relevant auxiliary references:**
  - ISO 27001, which describes the requirements for information security systems[8];
  - Strategic Marketing Plan audit from M. Baker[15], which describes the whole process of creating and assessing a marketing strategy.

This list should not be considered to be all encompassing, complete and final as there are other sources that can be used to identify controls. In addition it is important to stress that the Business Model Canvas itself can also be used as a control.

Generic controls	Examples of specific relevant references
Business plan	**BMC**—The Business Model Canvas allows organisations to fill their business model in a visual canvas that allows for easy understanding of their business, which can then be used to create a business plan of the organisation.
Infrastructure	**ISO 9001**—ISO 9001 contains requirements for the infrastructure and work environment of organisations that want to achieve a higher level of quality. It contains requirements for buildings, workspaces, utilities, equipment, and supporting services.
Maintenance	**ISO 25010**—ISO 25010 is a standard for systems and software quality requirements and evaluation. It takes into consideration maintainability as the "degree of effectiveness and efficiency with which a product or system can be modified by the intended Maintainers". This standard allows for measurement of maintainability and can help understanding the current state of a system regarding maintenance.
Marketing plan	**The Strategic Marketing Plan Audit**—Michael Baker provides in his book[15] a way to plan an organisation's marketing strategy. There are seven types of planning described in this book, and it guides the organisation in choosing which is the most appropriate. Then, having in mind the organisation's marketing objectives, it allows the organisation to measure how well the marketing strategy is aligned with the objectives and whether or not these objectives are being met. The audit considers eights steps: 1. Mission Statement 2. Marketing Appreciation 3. Conclusions and key assumptions 4. Strategic objectives 5. Core strategy 6. Key Policies 7. Administration and Control 8. Communication and timing.

[6] http://datasealofapproval.org/en/

Generic controls	Examples of specific relevant references
Operations	**COBIT**—The DSS01 process[14] provides guidelines and activities on how to coordinate and execute operations to deliver services. This process details how to perform operational procedures, manage outsource services, and manage facilities.
Performance assessment	**COBIT**—The APO07 process[14] of COBIT can be used to assess the performance of the organisation staff.
Preservation plan	**PLATO**—The preservation planning tool PLATO is a tool that provides support for preservation planning through the implementation of a preservation planning process and services for "content characterisation, preservation action and automatic object comparison".
Quality audit	**ISO 9001**—ISO9001 is the reference on quality certification. The goal of it is to provide assurance that an organisation's products and services meet the quality standards of their clients and also aims to continuously enhance quality.
Quality certification	**ISO 9001**—ISO 9001 provides a certification against the standard, which is recognised the world over as a seal of quality for an organisations products and services.
Re-engineering / Change management	**COBIT**—The BAI05 process[14] in COBIT delivers the activities needed by organisations to effectively manage organisational change. This process maximises the likelihood of implementing organisational change quickly and with reduced risk. It also enables organisations to manage change in a controlled manner; this includes standard changes in business processes, applications and infrastructure but also covers emergency maintenance.
Research and development	**ISO 9001**—Managing and implementing research and development can be difficult for organisations. Despite this the guidelines provided by ISO9001 can be used to this domain, in order to improve the quality of research and development.
Risk management plan	**ISO 31000 and ISO 31010**—This is a collection of standards for risk management. ISO 31000 provides the principles and guidelines for risk management and ISO 31010 provides the risk assessment techniques to be used in specific scenarios. These standards can be used to develop a risk management plan for an organisation. **DRAMBORA**—The Digital Repository Audit Method Based on Risk Assessment toolkit characterises digital curation as a risk-management activity, because it recognises the job of a digital curator as the rationalisation of the uncertainties and threats that inhibit efforts to maintain digital object authenticity and understanding, transforming these into manageable risks. There are six stages within the process. The first stages require that auditors develop an organisational profile, describing and documenting the repository's mandate, objectives, activities and assets. Then risks are derived from each of these, and assessed in terms of their likelihood and potential impact. In the end, auditors are encouraged to conceive of appropriate risk management responses to the identified risk.

Generic controls	Examples of specific relevant references
Security audit	**ISO 27001**—ISO 27001 is the reference on information security management. It provides a set of requirements, processes and controls that enable an organisation to mitigate and manage the risk regarding information management.
Security certification	**ISO 27001**—There is a certification for ISO 27001, which verifies that the organisation has good information management processes and controls and meets all the requirements. This certification is recognized throughout the world.
Staff assessment	**COBIT**—COBIT processes delivers procedures that can be used to assess staff and check whether staff are capable of dealing with arising issues. The APO07 process[14] allows for evaluation of staff performance, the maintenance of skills and competences, and maintaining adequate and appropriate staff to achieve the organisation business goals.
Staff salaries	**COBIT**—The APO07 process[14] aims to facilitate the managing of human resources in an organisation. It provides the activities necessary to manage contract of the staff, including salaries.
Staff training	**COBIT**—The BAI08 process[14] provides the guidelines to manage the knowledge existing in the organisation. This process allows for internal training of staff, as well as the identification of useful sources outside the organisation. The APO07 process also delivers guidelines on how to maintain staff skills.
Trust audit	**ISO 16363**—ISO 16363 provides guidelines and a checklist for repositories wishing to become trustworthy. **DSA**—The Data Seal of Approval is a self-assessment tool that can be used to check the current state of a repository. It contains 16 guidelines that evaluate the repository. There is an on-line tool that repositories can use to submit their assessment that is then reviewed by peers and, based on the answers and documentation provided, the Data Seal of Approval can then be awarded to the organisation.
Trust certification	**ISO 16363**—There is a certification for ISO 16363 that repositories can also use to certify their organisation and become recognized as a trustworthy repository. **DSA**—After submitting the answers and documentation for obtaining the Data Seal of Approval the repository is awarded the Seal. Amongst other things this includes the right to display a DSA logo on the organisation's website.

Table 3—Examples of relevant references for the generic controls.

5 Risk Identification and Analysis with the Business Model Canvas

The BMC technique can be used in the method outlined above to find risks and then controls for those risks. This in turn makes it possible to estimate the related costs as part of the overall costs of curation.

The purpose of this BMC is to represent a generic Business Model that can be applied to archives, serving as a template that can be instantiated to specific organizations.

To develop the OAIS BMC the recommended practice CCSDS 650.0-M-2 from the Consultative Committee for Space Data Systems (Magenta Book) was used. The first step involved the identification of the customer segments. In this analysis the Producers and Consumers were identified as being the customers of an archive. The designated community was also identified as a customer due to representing potential consumers.

Following the identification of the customer segments, the Value Propositions were identified. The value propositions building block "describes the bundle of products and services that create value for a specific Customer Segment".[10] In order to determine the Value Propositions, the following questions were posed and answered [10]:

1. What value do we deliver to the customer?
2. Which one of our customer's problems are we helping to solve?
3. What bundles of products and services are we offering to each Customer Segment?
4. Which customer needs are we satisfying?

The points of contact between the archive and the customers are encountered at the ingest and access stages. In access there are two types of interactions identified as *queries* and *orders.* The value proposition for the producers was identified as "Long-term preservation of AIP[7]" which is the value that producers take out of the archive and is directly linked to the ingest functional entity which allows the ingestion of data into the archive. For the consumers interactions two value propositions were identified, one for the *queries* depicted as "Resource Discovery" in the BMC and one for *orders* which is identified as "Access to Preserved Information" in the BMC.

[7] AIP—Archival Information Package

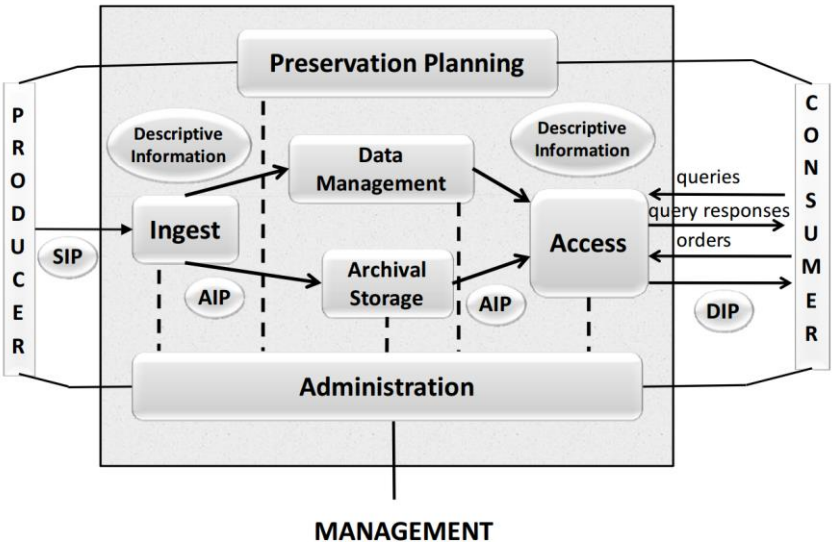


Figure 7—OAIS Functional Entities ([20], Page 4-1)

From these value propositions the channels were identified. The channels building block "describes how a company communicates with and reaches its customer segments to deliver a value proposition"[10]. In order to deliver the three value propositions identified earlier there is a need to create the appropriate channels. In this case one channel for each of the value propositions was identified. In order to enable long-term preservation of AIPs for Producers there is the need for a "Submission Service" to ingest data into the archive for long-term preservation. To enable the resource discovery value proposition, a query service is also needed to allow consumers to search for holdings of interest in the archive. Finally, to access the preserved information, there needs to be an order service which allows consumers the retrieve the identified holdings of interest from the archive.

The next step was to identify the customer relationships that are described as "the types of relationships a company establishes with specific customer segments"[10]. To engage with Producers and ingest content, the archive needs to establish a "Submission Agreement" which describes the data model of the information to be ingested in the form of a Submission Information Package (SIP). Besides this, the archive allows the establishment of a "SIP Submission Session" that allow Producers to submit SIPs for ingest into the archive. For resource discovery, the relationship between the archive and consumers is established using a "Search Session" in which consumers perform queries on the archive holdings. Finally, consumers engage in an "Event or Adhoc DIP Dissemination Session" to retrieve holdings of interest in the form of a DIP. The DIP is created according to an "Order Agreement" which has been agreed upon by both the archive and Consumers.

The key resources describe "the most important assets required to make a business model work"[10]. In order to allow for long-term preservation of AIPs there is the need to have an "Archiving Infrastructure" which supports the ingest of AIPs and also the preserved objects described as "Preserved AIP" in the BMC. To allow consumers to find the relevant holdings "Descriptive Information" is used along with the "Archiving Infrastructure" to support the execution of queries. Finally, to allow the access to the preserved information, the "Archiving Infrastructure" is needed to support the generation of a DIP from an AIP according to the Order Agreement.

To finalise the rationale behind the creation of this generic BMC, the key activities have been identified. Key activities describe "the most important things a company must do to make its business model work"[10]. If we look again to the functional entities of OAIS (Figure 7) these were identified as key activities in the generic BMC. In order to allow for long-term preservation of AIPs, the archive must perform "SIP Ingestion and AIP generation" which makes part of the Ingest Functional Entity. It must also

perform “Preservation Planning”, “AIP Storage”—the Archival Storage Functional Entity—and must Perform “Archive Administration”—the Administration Functional Entity. To allow resource discovery the archive must perform “Data Management” which allows access to the descriptive information necessary to identify relevant holdings in the archive. Finally, for the archive to allow access to preserved information it must perform “DIP Dissemination” which is the Access Functional Entity and must also have “Archive Administration”.

The Key Partnerships, Cost Structure and Revenue Streams could not be properly identified from OAIS as these are context dependent. For example, an archive can have software providers as key partners if the software is developed outside of the archive. On the other hand, there are archives that have on-site development, and, as such, have a different relationship with software providers. The cost structure also changes depending on the legislation of the archive and established accounting laws. The revenue streams are also context dependent. For instance, an archive might have its main source of revenue from public funding if it is a public organization, it might also provide training that makes for another revenue stream. As such, these building blocks are empty in the generic OAIS BMC and are filled in the instantiations using real cases.

Figure 8 depicts the generic BMC based on OAIS. Table 4 through to Table 9 expands the BMC using definitions from OAIS[20]. Section 5.1 details the Business Model Canvas, section 5.2 the related generic risk questions and section 5.3 presents a generic example of an associated registry of common risks and controls.

The three different colours in the BMC indicate the links between the elements in the BMC, as shown below.

- Related to the first Value Preposition—**Long-term preservation of AIP**
- Related to the second Value Preposition—**Resource discovery**
- Related to the third Value Preposition—**Access to Preserved Information**

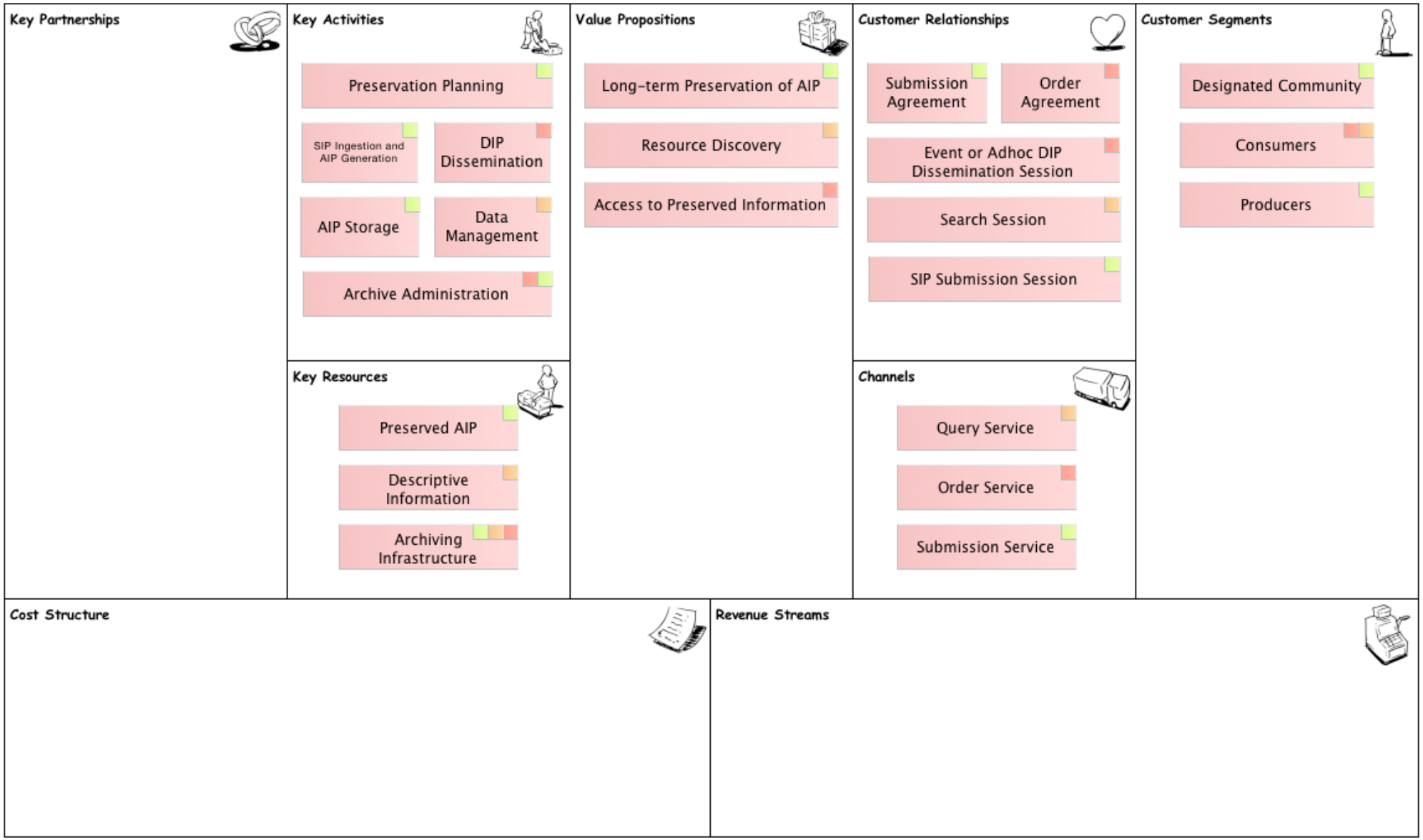


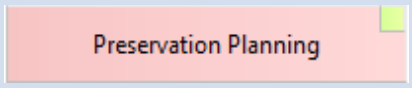
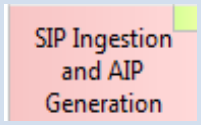
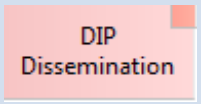
Figure 8—Generic BMC based on OAIS

5.1 Business Model Canvas Details

This section describes, for the purpose of our method, the entities of the generic BMC for digital curation. As part of the method, we propose that, after the definition of the BMC for a specific context, the related risk questions must be formulated, and after that, the risks must then be identified.

5.1.1 Key Activities

The key activities are the activities necessary to realise the business[10].

Key activities	Description
Preservation Planning 	"The OAIS functional entity which provides the services and functions for monitoring the environment of the OAIS and which provides recommendations and preservation plans to ensure that the information stored in the OAIS remains accessible to, and understandable by, and sufficiently usable by, the Designated Community over the Long Term, even if the original computing environment becomes obsolete." ([20], Page 1-14)
SIP Ingestion and AIP Generation 	The Submission Information Package (SIP) is "an Information Package that is delivered by the Producer to the OAIS for use in the construction or update of one or more AIPs and/or the associated Descriptive Information." ([20], Page 1-15) The Archival Information Package (AIP) is an Information Package, consisting of the Content Information and the associated Preservation Description Information, which is preserved within an OAIS. ([20], Page 1-9) The Ingest Functional Entity is "the OAIS functional entity that contains the services and functions that accept SIPs from Producers, prepares AIPs for storage, and ensures that AIPs and their supporting Descriptive Information become established within the OAIS." ([20], Page 1-12)
DIP Dissemination 	A Dissemination Information Package (DIP) is "an Information Package, derived from one or more AIPs, and sent by Archives to the Consumer in response to a request to the OAIS." ([20], Page 1-11)
Archive Administration 	The Administration Functional Entity "provides the services and functions for the overall operation of the Archive system. Administration functions include soliciting and negotiating submission agreements with Producers, auditing submissions to ensure that they meet Archive standards, and maintaining configuration management of system hardware and software. It also provides system engineering functions to monitor and improve Archive operations, and to inventory, report on, and migrate/update the contents of the

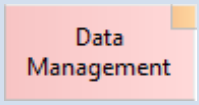
Key activities	Description
	Archive. It is also responsible for establishing and maintaining Archive." ([20], Page 4-2)
Data Management 	The Data Management Functional Entity "provides the services and functions for populating, maintaining, and accessing both Descriptive Information which identifies and documents Archive holdings and administrative data used to manage the Archive. Data Management functions include administering the Archive database functions (maintaining schema and view definitions, and referential integrity), performing database updates (loading new descriptive information or Archive administrative data), performing queries on the data management data to generate query responses, and producing reports from these query responses." ([20], Page 4-2)
AIP Storage 	The Archival Storage Functional Entity "provides the services and functions for the storage, maintenance and retrieval of AIPs. Archival Storage functions include receiving AIPs from Ingest and adding them to permanent storage, managing the storage hierarchy, refreshing the media on which Archive holdings are stored, performing routine and special error checking, providing disaster recovery capabilities, and providing AIPs to Access to fulfil orders." ([20], Page 4-2)

Table 4—Key activities identification

5.1.2 Value Propositions

The business seeks to solve customer problems and satisfy customer needs with value propositions—providing something that the customers perceive as having value and hence worth paying for[10].

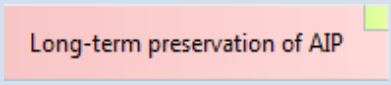
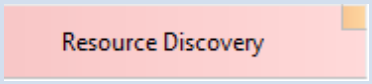
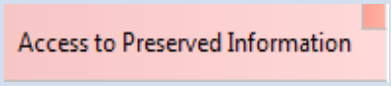
Value propositions	Description
Long-term preservation of AIP 	**"**Long Term Preservation is the act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term." ([20], Page 1-13) "Long Term may extend indefinitely. In the OAIS reference model there is a particular focus on digital information, both as the primary forms of information held and as supporting information for both digitally and physically archived materials." ([20], Page 1-1)
Resource Discovery 	"The access functional entity contains the services and functions which make the archival information holdings and related services visible to Consumers." ([20], Page 1-8) It also provides the services and functions that support Consumers in determining the existence, description, location and availability of information stored in the OAIS. ([20], Page 4-2)
Access to Preserved Information 	"Allows Consumers to request and receive information from the archive. Access functions include communicating with Consumers to receive requests, applying controls to limit access to specially protected information, coordinating the execution of requests to successful completion, generating responses (Dissemination Information Packages, query responses, reports) and delivering the responses to Consumers." ([20], Page 4-3)

Table 5—Value propositions identification

5.1.3 Customer Relationships

Customer relationships are established and maintained with each Customer Segment[10].

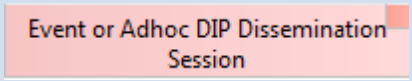
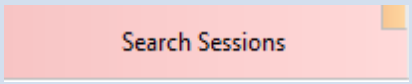
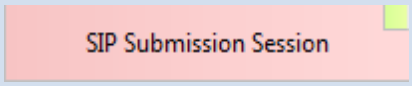
Customer relationships	Description
Submission agreement 	"The agreement reached between an OAIS and the Producer that specifies a data model, and any other arrangements needed, for the Data Submission Session. This data model identifies format/contents and the logical constructs used by the Producer and how they are represented on each media delivery or in a telecommunication session." ([20], Page 1-15)
Order agreement 	"An agreement between the Archive and the Consumer in which the physical details of the delivery, such as media type and format of Data, are specified." ([20], Page 1-13)
Event or Adhoc DIP Dissemination Session 	"A delivery of media or a single telecommunications session that provides Data to a Consumer. The Data Dissemination Session format/contents is based on a data model negotiated between the OAIS and the Consumer in the request agreement. This data model identifies the logical constructs used by the OAIS and how they are represented on each media delivery or in the telecommunication session." ([20], Page 1-10) A DIP Dissemination Session can either be Event based or Adhoc. In case it is event based it is "a request that is generated by a Consumer for information that is to be delivered periodically on the basis of some event or events." ([20], Page 1-11) If it is adhoc it means that there is "a request that is generated by a Consumer for information the OAIS has indicated is currently available." ([20], Page 1-9)
Search sessions 	"A session initiated by the Consumer with the Archive during which the Consumer will use the Archive Finding Aids to identify and investigate potential holdings of interest." ([20], Page 1-15)
SIP Submission Session 	"A delivery of media or a single telecommunications session that provides Data to an OAIS. The Data Submission Session format/contents is based on a data model negotiated between the OAIS and the Producer in the Submission Agreement. This data model identifies the logical constructs used by the Producer and how they are represented on each media delivery or in the telecommunication session." ([20], Page 1-11)

Table 6—Customer relationships identification

5.1.4 Key Resources

The key resources are the resources needed to allow the business to function[10].

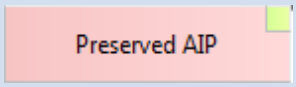
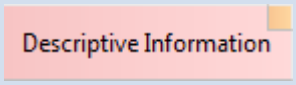
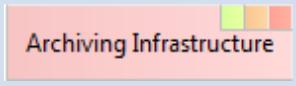
Key resources	Description
Preserved AIP 	An Archival Information Package is “an Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS.” ([20], Page 1-9)
Descriptive Information 	“The set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding, ordering, and retrieving of OAIS information holdings by Consumers”. ([20], Page 1-11)
Archiving Infrastructure 	The Archiving infrastructure contains the services and functions for the ingestion, storage and retrieval of AIPs. ([20], Page 1-9)

Table 7—Key resources identification

5.1.5 Channels

The communication, sales and distribution channels allow the business to reach clients and offer them the value proposition[10].

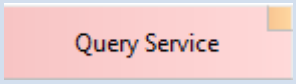
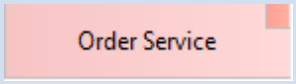
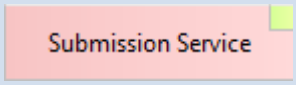
Channels	Description
Query Service 	The Query service allows consumers to perform queries on the holdings of the archive, to locate, analyse, order or retrieve potential information of interest. ([20], Page 1-15 and Page 1-8)
Order Service 	A service (Ordering Aid) that assists the Consumer in discovering the cost of, and in ordering, AIPs of interest. ([20], Page 1-13)
Submission Service 	The submission service supports the SIP ingestion. Producers submit SIPs through this service and receive receipt confirmations when the SIP is correctly ingested into the Archive. ([20], Page 4-5)

Table 8—Channels identification

5.2 Business Model Canvas Risk Questions

This section exemplifies risk questions for each of the sections of the Business Model Canvas. These questions were adopted from “BMI? Of course, but what about the Model Risks?”[12]. The questions are provided as examples which can then be adapted to specific scenarios. Their purpose is to facilitate the identification of risks for each of the sections of the BMC.

5.2.1 Key Partnerships (KP)

- "What are the effects of unavailability of a KP or a KP sourced asset?"
- "How would you determine that you might face a Partner loss?"
- "What can affect your relationship with your KP? How and how much?"
- "Would a talent/knowledge loss affect your viability or competence? What could trigger it?"
- "How would liabilities or a reputational impact affect KP support?"
- "What impact would be generated by a legal or regulatory situation? How do you ensure compliance?"

5.2.2 Key Activities (KA)

- "What are the immediate and long term effects of unavailability of a KA?"
- "How would you determine that you might face a KA loss?"
- "Can security vulnerabilities, bribery or fraud affect your KA? How and how much?"
- "Would a talent/knowledge loss affect your viability or competence? What could trigger it?"
- "How much of an effect would there be if there was an erosion in internal and external quality?"
- "What can generate a legal or regulatory situation and what would its impact be? How do you ensure compliance?"

5.2.3 Value Propositions (VP)

- "How is satisfaction measured? What can cause a degradation in satisfaction and how would that impact?"
- "Can security vulnerabilities, bribery or fraud affect your VP? How and how much?"
- "Would a talent/knowledge loss affect your viability or competence? What could trigger it?"
- "Are your internal policies and controls aligned to your VP? How do you measure it?"
- "How is feedback routed back?"
- Is the VP still valuable?"

5.2.4 Customer Relationships (CR)

- "What events or factors would impact our relationship with our customers?"
- "Do you have a clear understanding of their emotional experience with your brand?"

5.2.5 Customer Segments (CS)

- "What events or factors would erode their perception of value?"

5.2.6 Key Resources (KR)

- "How would a shortage of key resources impact your operation?"
- "What might cause you to face a key resource shortage?"
- "Can security vulnerabilities, bribery or fraud affect your sourcing? How and how much?"
- "Would a talent/knowledge loss affect your viability or competence? What could trigger it?"

5.2.7 Channels (CH)

- "What are the potential impacts of outages in the service delivery?"
- "What are the key factors or vulnerabilities that could end in a potential service outage?"

- "What is the dependency versus autonomy of your channels?"
- "What is your relative position (your control) over the various channels?"

5.2.8 Cost Structure (C$)

- "What is the potential impact of variations (price and availability) in the most important cost components?"
- "What conditions in each of your markets can potentially impact your cost streams?"
- "Can security vulnerabilities, bribery or fraud affect your cost? How and how much?"
- "How would liabilities or a reputational impact affect your Costs Stream?"

5.2.9 Revenue Streams (R$)

- "What market events or competitor dynamics could impact what customers will be willing to pay?"
- "What conditions in each of your markets could potentially impact your revenue streams?"
- "Can security vulnerabilities or fraud affect your revenues? How and how much?"

5.3 Generic Risks and Controls for the Generic BMC

Use of the generic BMC ideally allows for the extraction of a set of risk related questions that can lead to the identification of actual risks. The questions can be extracted from each of the BMC sections and/or users can use the questions provided in Section 5.2. For example, for the key activities, we can formulate questions aiming at verifying whether the organisation is able to cope with change or not, such as "Are the current repository activities flexible enough to comply with preservation requirements changes?" or "Is the repository able to comply with an emerging technology change?"

For the customer relationships we can formulate questions like "If there is a case of an unauthorised access to the contents of the repository, will that affect the trustworthiness or reputation of the repository?" This particular question aims to verify whether a) there is sensitive information in the repository, and b) whether any unauthorized access will discourage customers from continuing to use the repository.

For the key resources examples related questions could be "Is there a training plan for staff?" or "Are digital assets normalized to comply with manageable formats that can be used for digital preservation?" The first question aims at verifying whether there is a risk of not having the necessary knowledge in the organisation to continue fulfilling its mandate. It also estalishes whether there is a risk in having obsolete skills. The second question aims at verifying whether the contents of the repository are available in an appropriate format.

After the formulation of the risk questions, the next step is to identify the related risks, and then their respective controls (either those already in place, for an understanding of the actual costs of the "Current" scenario, or those that could potentially be applied, to explore possible "Future" scenarios).

Generic risks and controls were identified after analysing the results of the DRAMBORA [1] report. The Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) represents an effort to conceive

criteria, means and methodologies for risk assessment of digital repositories. The risks and controls that better align with the generic BMC model were selected and are presented in Table 9 below[8].

Generic risks from the generic BMC (The possible consequence in case of an occurrence)	**Generic possible controls** (the source of costs)
Loss of reputation	Repository assessment: external and internal audit and risk analysis
Activity allocates insufficient resources	Use mechanisms to measure activity efficiency in terms of allocated resources, procedures and policies
Community requirements change substantially	Identify, monitor and review the understanding of the community requirements and of the repository objectives
Enforced cessation of repository operations	Plan for continuation of preservation activities beyond repository's lifetime
Community feedback not received	Use mechanisms (e.g. email, surveys) for soliciting feedback from repository users community
Community feedback not acted upon	Define policies to acknowledge community's feedback
Business objectives not met	Define mission statement and repository's objectives accordingly to business requirements
Business fails to preserve essential characteristics of digital information	Define main characteristics of digital content for information preservation
Business policies and procedures are inefficient	Document and make available business policies and procedures
Perception of the extent of repository's success	Document, monitor and review audit feedback, internal risk assessment
Loss of key member(s) of staff	Appoint a sufficient number of appropriately qualified personnel
Personnel suffer skill loss	Implement mechanisms to identify ongoing personnel training requirements
Budget reduction	Define a financial preservation plan to assure self-sustainability of repository
Software failure or incompatibility	Install software updates
Hardware failure or incompatibility	Monitor hardware performance

[8] In most cases the suggested controls are single examples off possible controls. In a real world situation there will often be two or more appropriate controls that could be used to mitigate the risk.

Generic risks from the generic BMC (The possible consequence in case of an occurrence)	Generic possible controls (the source of costs)
Obsolescence of hardware or software	Maintain hardware/software up to date to meet repository objectives
Media degradation or obsolescence	Allocate resources to monitor media storage lifetime and assess potential value of emerging technologies
Physical intrusion of hardware storage space	Implement security infrastructure (e.g. passwords, encryption software)
Local destructive or disruptive environmental phenomenon	Implement physical security measures (e.g. video-record)
Non availability of core utilities (e.g. electricity, gas)	Define internal means to nullify disruption of service, monitor and review contract agreements of provider's services
Loss of other third-party services	Document and review service level contracts or service commitments with utility provider
Loss of availability of information and/or service	Use mechanisms to evaluate efficiency of software/hardware systems service levels to meet preservation strategies
Loss of authenticity/integrity of information	Monitor, record and validate integrity of received content
Loss or non-suitability of backups	Define a backup strategy plan: store content in remote location, multiple copies of information, test backup systems
Preservation strategies result in information loss	Define, review and implement preservation plans
Inability to validate effectiveness of preservation	Implement and review strategies for physical archival storage and migration
Non-traceability of received, archived or disseminated package	Monitor and record information provenance during its preservation lifecycle
Metadata to information referential integrity is compromised	Maintain referential integrity between metadata and archived content
Documented change history incomplete or incorrect	Document and monitor semantic and technical context
Authentication subsystem fails	Implement authentication subsystems to reflect agreed access rights and restrictions
Authorisation subsystem fails	Implement authorisation subsystems to reflect agreed

Generic risks from the generic BMC (The possible consequence in case of an occurrence)	Generic possible controls (the source of costs)
	access rights and restrictions
Inability to validate effectiveness of dissemination mechanism	Disseminate a complete and authentic object as originally submitted

Table 9—Generic controls identification

6 Risk Identification and Analysis of the Case Studies

This section describes the application of the pragmatic method to estimate costs of curation focusing on risks and controls to two case studies, (1) The National Laboratory of Civil Engineering (LNEC) from Portugal and (2) the Portuguese Web Archive (PWA). Part of the analysis for these cases was performed in the scope of task 4.5.

6.1 Case Study: LNEC

The LNEC[9], established in November 1946, is a public Science and Technology institution, which is subject to Government supervision through the Ministry of Public Works, Transports and Communications. Its competences, as referred to in the definition of strategic guidelines, are performed in conjunction with the Ministry of Science, Technology and Higher Education. Its activity is developed in the various fields of civil engineering and its main assignments are the execution, supervision and promotion of scientific research and technological developments to achieve progress, innovation and good practices in civil engineering. The institution is also responsible for providing an unbiased and suitable scientific and technical support to the executive power in its governing and regulatory activities. At present, it has about 680 staff, of which approximately 42% have a University degree and about 22% are researchers holding a PhD or an equivalent degree[10]. It also has about 80 scientific research fellows with a grant awarded by LNEC. From the annual LNEC budget, about 50% of LNEC's income results from the generation of private revenues, namely referring to the provision of science and technology services, the remaining 50% deriving from the National Budget and from other sources[10].

LNEC undertakes research in the following areas:

- Usage of monitoring technologies to gather observation data and automatic communication systems;
- Development of "smart systems" for just-in-time dam safety control;
- Risk analysis of dam construction and operation;
- Characterisation and modelling of future deterioration of dams and their foundations.

The Portuguese Dam Safety Legislation that regulates the dam safety control of big Portuguese dams (dams where the maximum height from the foundation is more than 15 meters or with a reservoir with more than 100 000 m^3 of capacity) gives LNEC the responsibility of surveying the behaviour and the structural safety of approximately 150 concrete and masonry dams. This responsibility includes the generating observation plans, the periodic inspection of the dam structure and potential anomalies, the generation of analysis and interpretation reports of the observed behaviour and also,the management and operation of an electronic archive of data concerning the dam safety.

6.1.1 Identification of the relevant Indirect Economic Determinants

The identification of the relevant Indirect Economic Determinants was performed with the help of a stakeholder from LNEC. In this process the list of Indirect Economic Determinants from D4.1 was analysed. Each of the Indirect Economic Determinants it was checked to see if it was relevant for LNEC.

[9] http://www.lnec.pt

[10] 2006 Social Report

There were no additional Indirect Economic Determinants beyond those outlined in D4.1 found to be relevant for this case.

Determinant	Relevant?	Rationale
Authenticity	Yes	The observed data, which cannot be reproduced, is important for decision-making
Benefit	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Confidentiality	Yes	Important mostly because the observed data is not public. Collected data is technical and therefore difficult to interpret—the data is cyphered—and it's necessary to obtain additional information in order to create value from it.
Efficiency	Yes	It's mainly relevant in terms of obtaining better results with less resources and costs, not in terms of long-time preservation of the observed data.
Flexibility	Yes	LNEC only uses processes to collect the observed data, it doesn't offers this service to others (as in a scientific repository).
Impact	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Innovation	Yes	Relevant because of the scientific domain of LNEC business—maintenance and preservation of the observed data.
Interoperability	Yes	The observed data is already normalized when collected in the needed reading formats, so it's not so relevant. There is also documentation and technical reports.
Quality	Yes	The observed data, which cannot be reproduced, is important for decision-making
Reputation	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Risk	Yes	Related to all the other indirect economic determinants.
Sensitivity	No	Not Applicable to LNEC
Skills	Yes	Training and documentation.
Sustainability	Yes	To guarantee value to the organizations that work with LNEC regarding the analysis of the observed data.
Transparency	Yes	To guarantee the trustworthiness and quality of the observed data.
Trustworthiness	Yes	The observed data, which cannot be reproduced, is important for decision-making
Value	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.

Table 10—Relevance of the Indirect Economic Determinants for the case of LNEC

6.1.2 Analysis of the Indirect Economic Determinants

From the analysis of the relevant indirect economic determinants for the case of LNEC we identified the risks for this case based on the generic analysis provided by Table 2. For examples of specific, relevant references for the controls identified, refer to Table 3.

Determinant	Related risks (possible consequence in case of hazard)	Typical controls (source of costs)
Authenticity	Loss of reputation and trust	Preservation plan
Benefit	Ability to deliver	Business plan
Confidentiality	Exposure to competitors	Security auditing Security certification
Efficiency	Exposure to financial uncertainty	Performance assessment Re-engineering / Change management Operations Maintenance Infrastructure
Flexibility	Inability to explore new opportunities	Re-engineering / Change management
Impact	Loss of reputation and trust	Marketing plan
Innovation	Exposure to obsolescence Uncertainty of early adopter Inability to explore new opportunities	Research and development Re-engineering / Change management
Interoperability	Ability to deliver	Operations
Quality	Quality of service	Quality auditing Quality certification
Reputation	Reputation	Marketing plan
Risk	Risk Management	Risk management plan
Skills	Loss of efficiency if key staff leave	Staff assessment Staff training Staff salaries/benefits
Sustainability	Exposure to financial uncertainty	Business plan Marketing plan
Transparency	Loss of reputation and trust	Marketing plan
Trustworthiness	Loss of reputation and trust	Trustiness auditing Trustiness certification
Value	Ability to deliver	Business plan

Table 11—A risk identification and analysis of the indirect economic determinant for the case of the LNEC

6.1.3 Analysis of the Business Model Canvas

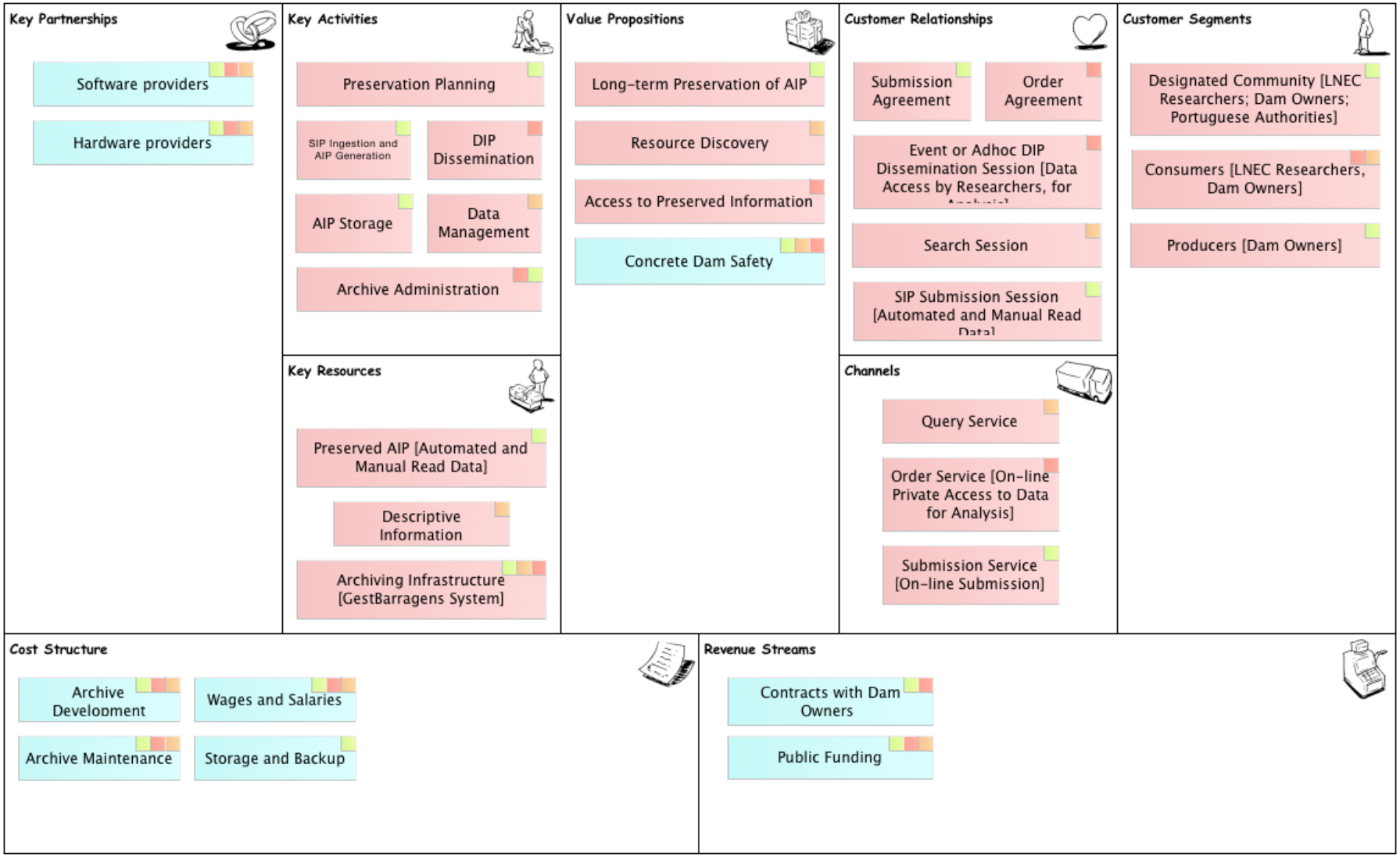


Figure 9—Portuguese Civil Engineering National Laboratory BMC

6.1.3.1 Business Model Canvas Details

The objects in the LNEC BMC that are particular to this case study are shown in the tables below. The BMC sections that are not detailed in this section use the same description in Section 5.1.

1) Key Partners

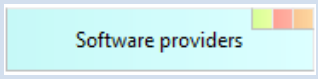
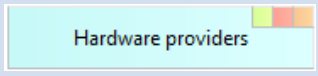
Key Partners	Description
Software Providers 	Software Providers develop and maintain specialised software products to suit the organisations requirements. These software products are used to support the value propositions of the organisation.
Hardware Providers 	Hardware providers sell and maintain the hardware deployed in the organisation to support the value propositions of the organisation.

Table 12—Key Partners Identification

2) Value propositions

Value propositions	Description
Concrete Dam Safety 	Another value proposition of LNEC's archive besides the ones in the generic OAIS model is to improve concrete dam safety through the collection of data from dams and further analysis by LNEC researchers.

Table 13—Value Propositions Identification

3) Customer relationships

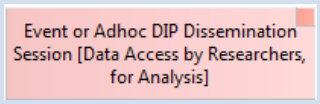
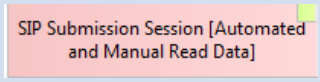
Customer relationships	Description
Event or Adhoc DIP Dissemination Session 	LNEC's researchers initiate DIP Dissemination sessions with the archive to collect relevant information on dam structural behaviour.
SIP Submission Session 	There are sensors installed in dams that have automated procedures to read data and send it to the archive for ingest. However, there are some measurements that have to be performed by people on a periodic basis and this data is submitted manually to the archive for ingest.

Table 14—Customer Relationships Identification

4) *Customer segments*

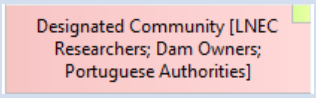
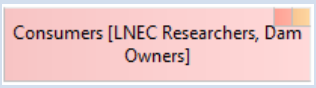
Customer segments	Description
Designated Community 	The LNEC Researchers use the information in the Archive to analyse dam structures behaviour over time. As such there are both a consumer and part of the designated community as the archive must fulfil their requirements. Dam Owners also have specific requirements as they are both consumers and producers. The Portuguese authorities also have specific requirements which are detailed in law, as there is a mandate to collect this information by law.
Consumers 	The LNEC Researchers use the information in the archive to analyse the dam structures behaviour and also to predict future behaviour. The Dam Owners also use the information in the archive to check on their dam's structural integrity.
Producers 	The data collected through sensors and by staff from the dam owners is ingested to the archive for long-term preservation and analysis.

Table 15—Customer Segment Identification

5) *Key resources*

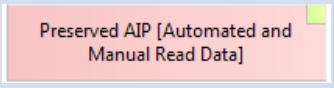
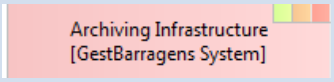
Key resources	Description
Preserved AIP 	The AIPs in the archive consist of the data collected from dams and ingested automatically and manually.
Archiving Infrastructure 	The GestBarragens system is an information system that provides the whole infrastructure for the Archive.

Table 16—Key Resources Identification

6) *Channels*

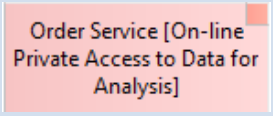
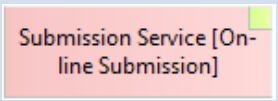
Channels	Description
Order Service 	The order service enables on-line private access to data for analysis for the LNEC researchers. Due to the sensitive aspect of the information in the archive there is the need to maintain private access.
Submission Service 	Dam Owners submit their data, collected automatically and manually, through an on-line submission service made available by LNEC's archive.

Table 17—Channels Identification

7) Cost structure

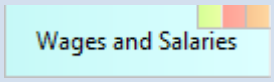
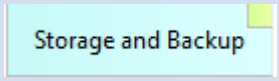
Cost structure	Description
Archive Development 	One of the costs of the archive is development. This cost comes from the initial development of the archive and also new functions and requirements that arise from the designated community and when new technology is needed.
Archive Maintenance 	Another of the costs of the archive is maintenance. Maintenance of the archive is performed both by the software and hardware providers and also by the archive staff. There is the need to perform periodic procedures to guarantee that the archive is running smoothly and the information in the archive remains relevant for the designated community.
Wages and Salaries 	One of the costs of running the archive is the wages and salaries of the staff that supports the operation of the archive. There are staff with different qualification in the team and wages and salaries differences are also taken in consideration.
Storage and Backup 	One of the main costs in the archive is related to storage and backups. As storage is not outsourced there is the need to check and maintain periodically the hardware that deals with the information storage in the archive.

Table 18—Cost Structure Identification

8) Revenue streams

Revenue streams	Description
Contracts with Dam Owners 	The contracts that LNEC have with the dam owners so that they can ingest and store data in the archive are one of the revenue streams of the archive.
Public Funding 	The main portion of the revenue of the archive comes from public funding. As LNEC is a public organization there is budget allocated for LNEC as part of the annual government budget.

Table 19—Revenue Streams Identification

6.1.3.2 BMC Risk Identification and Controls for LNEC

Table 20 shows the analysis of controls for the risk identification and analysis of the BMC for the LNEC case study. For further details relating to the controls see Table 3.

BMC Section	Risk	Rationale
Key Partnerships	Obsolescence of hardware or software	Selected risks regarding the outsourcing of services the repository may depend on to deliver the preservation business
	Physical intrusion of hardware storage space	
	Loss of other third-party services	
	Loss of availability of information and/or service	
	Loss or non-suitability of backups	
	Preservation strategies result in information loss	
	Inability to validate effectiveness of preservation	
Key Activities	Non-traceability of received, archived or disseminated package	Risks associated with the overall business of the repository, mainly the preservation workflow
	Metadata to information referential integrity is compromised	
	Documented change history incomplete or incorrect	
	Authentication subsystem fails	
	Authorisation subsystem fails	
	Activity allocates insufficient resources	
	Inability to validate effectiveness of dissemination mechanism	
	Loss of authenticity/integrity of information	
Key Resources	Local destructive or disruptive environmental phenomenon	Risks related to the resources, infrastructure and personnel which sustain the repository business
	Enforced cessation of repository operations	
	Loss of key member(s) of staff	
	Personnel suffer skill loss	
	Software failure or incompatibility	
	Hardware failure or incompatibility	

BMC Section	Risk	Rationale
Value Proposition	Business objectives not met	Risks regarding the vision and value of a repository
	Business fails to preserve essential characteristics of digital information	
	Business policies and procedures are inefficient	
Customer Segments	Community requirements change substantially	Risk that relate to what the repository should deliver within the community vision
Customer Channels	Community feedback not received	Risks related to the communication and dissemination of the business provided by a repository
	Perception of the extent of repository's success	
Customer Relationships	Community feedback not acted upon	Risks associated with the community that makes use of the repository for their research work
	Loss of availability of information and/or service	
Cost Structure	Loss of key member(s) of staff	Risks regarding the cost to support the repository business
	Obsolescence of hardware or software	
	Loss or non-suitability of backups	
	Loss of authenticity/integrity of information	
	Non availability of core utilities (e.g. electricity, gas)	
Revenue Stream	Budgetary reduction	Risks related to the worth of a repository business and the value it offers to the community
	Perception of the extent of repository's success	

Table 20—Examples of relevant controls for the LNEC risks identified from the BMC.

6.2 Case Study: Portuguese Web Archive (PWA)

The Portuguese Web Archive[11] preserves the information published on the web of interest for the Portuguese community for future access. It also provides research resources—for instance in the fields of History, Sociology or Linguistics—and preserves information from the past that is no longer available on the Internet. With the creation of a system that supports regular crawls of the Portuguese web and its long term storage and access, it is intended to provide the following services:

- **Term search over the archived contents**: it enables the identification of archived contents over the years that contain certain terms;
- **URL search over the archived contents**: it allows users to identify several versions of a content gathered from a given URL;
- **New search engine over the Portuguese web**: the archive enables searching over several archived Portuguese web collections. Providing a search service over the most recent collection, as current web search engines do, is attainable in a relatively small additional effort and it is an interesting service for the Portuguese community;
- **Historical collections of web contents for research purposes**: the web has information about the many and varied subjects reflecting society changes across time. Researchers from different fields use the web as a source of information for their studies. Providing archived web collections will enable these researchers to store and process web data locally on their computers without having to crawl the web themselves;
- **Characterisation reports of the Portuguese web**: a web archive system must be tuned according to the characteristics of the archived data. For this reason, Portuguese web characterisations are periodically generated. As these studies are interesting to a broader audience, they will be published. Characterising national webs is interesting to measure the spread of information technologies in different societies and the evolution of the web across time;
- **Backup system of the archived information**: it is a distributed system that enables Internet users to provide disk space to store backup copies of the archived contents through the installation of a small application on their computers. If a failure happens on the central repository, the archived collection will be recovered from the backup copies stored on the users' computers. Any individual or institution can contribute to preserve the web by providing some disk space on their computers;
- **Archived data parallel processing system**: it allows researchers to execute their programs over the archived web data using several computers in parallel.

As a bounus, the Portuguese Web Archive also strives to achieve the following goals:

- **Train human resources** in web archiving to enable the maintenance of the system in the future;
- **Export know-how, experience and technology** in web archiving to other countries, especially the Portuguese language ones;
- **Contribute to the increase of the number of domains registered under ".PT"**, the free historical archiving of the information published under this domain could be an additional motivation for registrars;

[11] http://sobre.arquivo.pt/

- **Publish scientific and technical papers** that enable the sharing of the acquired knowledge and receiving feedback from the community regarding the work performed

6.2.1 Identification of the relevant Indirect Economic Determinants

The identification of the relevant Indirect Economic Determinants was performed with the help of a stakeholder from PWA. In this process the list of Indirect Economic Determinants from D4.1 was analysed. Each of the Indirect Economic Determinants it was checked to see if it was relevant for PWA. There were no additional Indirect Economic Determinants beyond those outlined in D4.1 found to be relevant for this case.

Determinant	Relevant?	Rationale
Authenticity	Yes	The websites that are ingested to the archive must remain authentic to allow to be used as evidence in court.
Benefit	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Confidentiality	No	The data in the Archive is publicly available.
Efficiency	Yes	It's mainly relevant in terms of obtaining better results with less resources and costs, not in terms of long-time preservation of the observed data.
Flexibility	Yes	PWA uses processes to collect the websites data, and offers this service to others.
Impact	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Innovation	Yes	Relevant because of the domain of PWA business—maintenance and preservation of the websites data.
Interoperability	No	As the archive just stores web sites in their original form, there is not a substantial degree on interoperability.
Quality	Yes	The web sites, which are constantly changing, must maintain the same quality standards as when they were online.
Reputation	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.
Risk	Yes	Related to all the other indirect economic determinants.
Sensitivity	No	Not Applicable to PWA.
Skills	Yes	Training and documentation.
Sustainability	Yes	To guarantee value to the organizations that work with PWA.
Transparency	Yes	To guarantee the trustworthiness and quality of the web sites collected.

Determinant	Relevant?	Rationale
Trustworthiness	Yes	In order for the collected data to be used as evidence in court there must be a high degree of trust from courts and other public services.
Value	Yes	Seen as a result of the other indirect economic determinants, such as authenticity, efficiency, impact, trustworthiness.

Table 21—Relevance of the Indirect Economic Determinants for the case of PWA

6.2.2 Analysis of the Indirect Economic Determinants

From the analysis of the relevant indirect economic determinants as applied to the case of PWA we identified the risks for this case based on the generic analysis provided by Table 2. For examples of specific, relevant references for the controls identified, refer to Table 3.

Determinant	Related risks (possible consequence in case of hazard)	Typical controls (source of costs)
Authenticity	Loss of reputation and trust	Preservation plan
Benefit	Ability to deliver	Business plan
Efficiency	Exposure to financial uncertainty	Performance assessment Re-engineering / Change management Operations Maintenance Infrastructure
Flexibility	Inability to explore new opportunities	Re-engineering / Change management
Impact	Loss of reputation and trust	Marketing plan
Innovation	Exposure to obsolescence Uncertainty of early adopter Inability to explore new opportunities	Research and development Re-engineering / Change management
Quality	Quality of service	Quality auditing Quality certification
Reputation	Reputation	Marketing plan
Risk	Risk Management	Risk management plan
Skills	Loss of efficiency if key staff leave	Staff assessment Staff training Staff salaries/benefits
Sustainability	Exposure to financial uncertainty	Business plan Marketing plan
Transparency	Loss of reputation and trust	Marketing plan

Determinant	Related risks (possible consequence in case of hazard)	Typical controls (source of costs)
Trustworthiness	Loss of reputation and trust	Trustiness auditing Trustiness certification
Value	Ability to deliver	Business plan

Table 22—A risk identification and analysis of the indirect economic determinant for the case of the PWA

6.2.3 Analysis of the Business Model Canvas

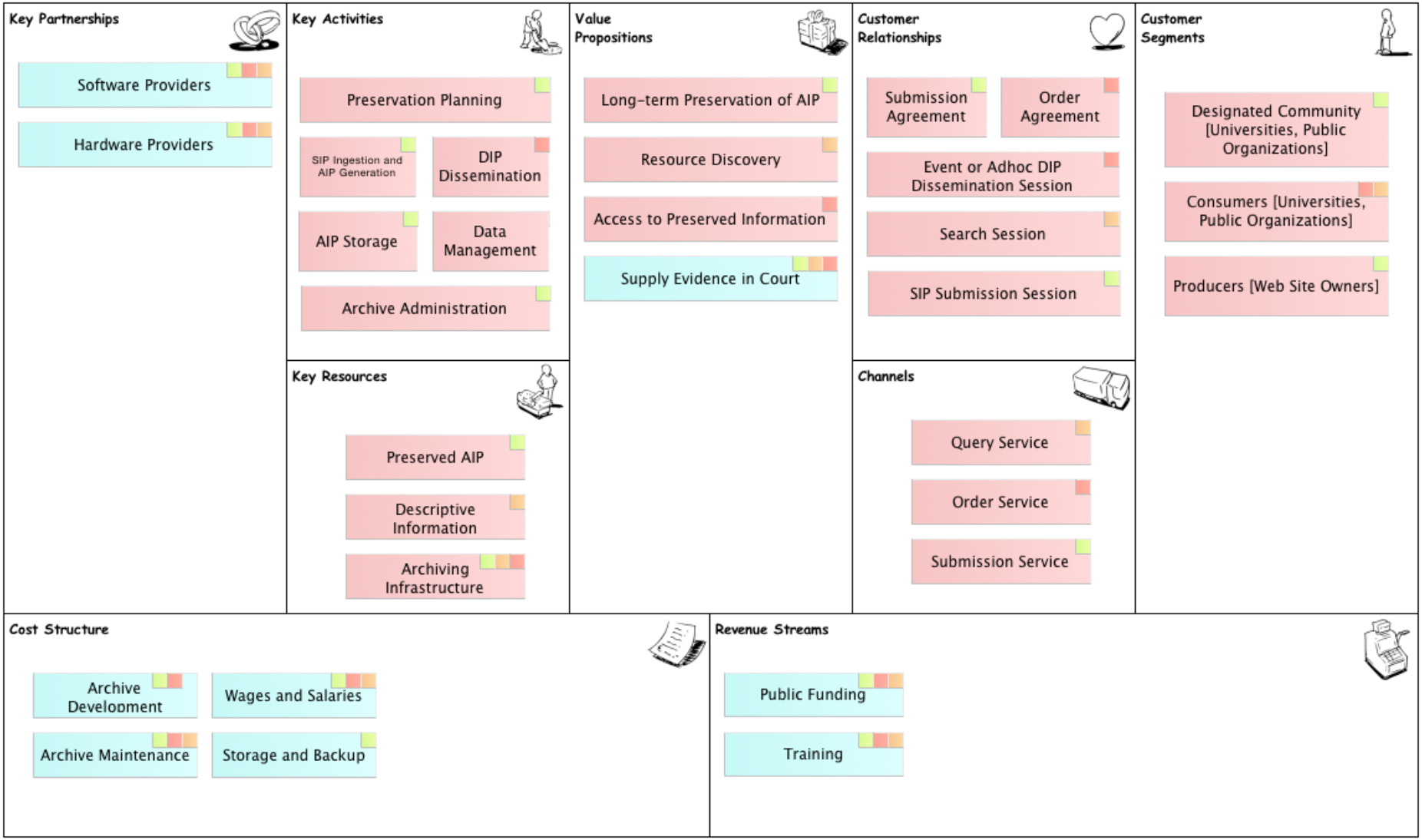


Figure 10- Portuguese Web Archive BMC

6.2.3.1 Business Model Canvas Details

The objects in the PWA BMC that are particular to this case study are shown in the tables below. The BMC sections that are not detailed in this section use the same description in Section 5.1.

Key Partnerships

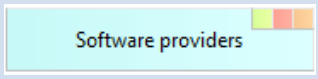
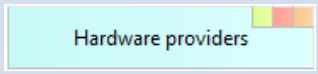
Key Partnerships	Description
Software Providers 	Software Providers develop and maintain specialised software products to suit the organisations requirements. These software products are used to support the value propositions of the organisation.
Hardware Providers 	Hardware providers sell and maintain the hardware deployed in the organisation to support the value propositions of the organisation.

Table 23—Key Partners Identification

9) Value propositions

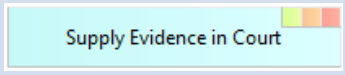
Value propositions	Description
Supply Evidence in Court 	One specific value proposition of PWA is the provision of evidence in court, as the archive contains a collection of Portuguese web sites at different points in time that might be used as evidence in court.

Table 24—Value Propositions Identification

10) Customer segments

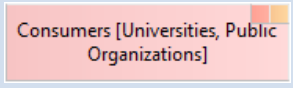
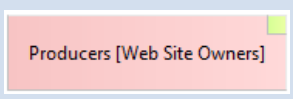
Customer segments	Description
Designated Community 	Universities and Public organisations are the main consumers of the archive with specific requirements.
Consumers 	Universities and Public organisations use the information in the archive for various objectives. Public organisations, such as courts, can use the information to provide evidence in court cases and prosecution. Universities can use the information for research (for example in the fields of sociology or information technology).
Producers 	The web site owners are the producers of the information that is ingested in archive.

Table 25—Customer Segment Identification

11) Cost structure

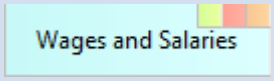
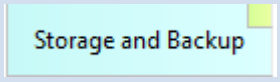
Cost structure	Description
Archive Development 	One of the costs of the archive is development. This cost comes from the initial development of the archive and also new functions and requirements that arise from the designated community and also when new technology is needed.
Archive Maintenance 	Another of the costs of the archive is maintenance. Maintenance of the archive is performed both by the software and hardware providers and also by the archive staff. There is the need to perform periodic procedures to guarantee that the archive is running smoothly and the information in the archive remains relevant for the designated community.
Wages and Salaries 	One of the costs of running the archive is the wages and salaries of the staff that supports the operation of the archive. There are staff with different qualification in the team and wages and salaries differences are also taken in consideration.
Storage and Backup 	One of the main costs in the archive is related to storage and backups. As storage is not outsourced there is the need to check and maintain periodically the hardware that deals with the information storage in the archive.

Table 26—Cost Structure Identification

12) Revenue streams

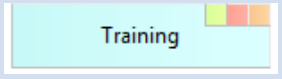
Revenue streams	Description
Public Funding 	The main portion of the revenue of the archive comes from public funding. As the Foundation for National Scientific Computing (FCCN) where PWA is incorporated is a public organization there is budget allocated for PWA as part of the annual government budget.
Training 	PWA offers training in the field of web archiving.

Table 27—Revenue Streams Identification

6.2.3.2 BMC Risk Identification and Controls for PWA

Table 28 shows the analysis of controls for the risk identification and analysis of the BMC for the PWA case study. For further details relating to the controls see Table 3.

BMC Section	Risk	Rationale
Key Partnerships	Obsolescence of hardware or software	Selected risks regarding the outsourcing services the repository may depend on to deliver the preservation business
	Physical intrusion of hardware storage space	
	Loss of other third-party services	
	Loss of availability of information and/or service	
	Loss or non-suitability of backups	
	Preservation strategies result in information loss	
	Inability to validate effectiveness of preservation	
Key Activities	Non-traceability of received, archived or disseminated package	Risks associated with the overall business of the repository, mainly the preservation workflow
	Metadata to information referential integrity is compromised	
	Documented change history incomplete or incorrect	
	Authentication subsystem fails	
	Authorisation subsystem fails	
	Activity allocates insufficient resources	
	Inability to validate effectiveness of dissemination mechanism	
	Loss of authenticity/integrity of information	
Key Resources	Local destructive or disruptive environmental phenomenon	Risks related to the resources of infrastructure and personnel which sustain the repository business
	Enforced cessation of repository operations	
	Loss of key member(s) of staff	
	Personnel suffer skill loss	
	Software failure or incompatibility	
	Hardware failure or incompatibility	

BMC Section	Risk	Rationale
Value Proposition	Business objectives not met	Risks regarding the vision and value of a repository
	Business fails to preserve essential characteristics of digital information	
	Business policies and procedures are inefficient	
Customer Segments	Community requirements change substantially	Risk that relates with what the repository should deliver within the community vision
Customer Channels	Community feedback not received	Risks related to the communication and dissemination of the business provided by a repository
	Perception of the extent of repository's success	
Customer Relationships	Community feedback not acted upon	Risks associated with the community that makes use of the repository for their research work
	Loss of availability of information and/or service	
Cost Structure	Loss of key member(s) of staff	Risks regarding the cost to support the repository business
	Obsolescence of hardware or software	
	Loss or non-suitability of backups	
	Loss of authenticity/integrity of information	
	Non availability of core utilities (e.g. electricity, gas)	
Revenue Stream	Budgetary reduction	Risks related to the worth of a repository business and the value it offers to the community
	Perception of the extent of repository's success	
	Personnel suffer skill loss	

Table 28—Examples of relevant controls for the PWA risks identified from the BMC

7 Conclusions

This deliverable proposed a pragmatic method for estimating costs of curation that can be associated with the following:

a) **"Current" scenario**, where the costs of controls already exist in the repository as a means to reduce the impact of a consequence of a risk, change the likelihood of an event, or reduce the exposure to a vulnerability;
b) **"Future" scenario**, where the costs of controls do not yet exist, but where repository managers are able to consider alternative scenarios of repository governance.

The foundations of this method make use of relevant sources of literature, such as the ISO 31000 and the Business Model Canvas (BMC), and the results from D4.1, where the indirect economic determinants are described and analysed; from D4.5, where the BMC is presented; from D4.3, where literature sources regarding trustworthiness and control are described and analysed; and from D3.2, which describes the cost concept model and gateway specification.

The focus of this deliverable was to present the method as a pragmatic technique, and provide some examples through case studies. The detailed guide of the application of the method will be provided by a toolset, integrated in the CCEx, made of the following:

- A generic BMC, with an associated generic registry of risk questions: a registry of typical risk questions that are relevant for the domain of digital curation and that can guide to define a specific set of risk questions for a specific BMC;
- A risk registry for digital curation: a registry of risks derived from the previous risk questions, and also common related controls, relevant for the domain of digital curation;
- Case studies: a library of detailed applications of the method to real repositories.

8 References

[1] Digital Curation Centre and Digital Preservation Europe, (February 2007), “DCC and DPE Digital Repository Audit Method Based on Risk Assessment, v1.0.”, retrieved 22.03.2014, from http://www.repositoryaudit.eu/download ISO (2009)

[2] ISO/FDIS 31000:2009(E): Risk Management—Principles and guidelines

[3] IEC/FDIS 31010:2009(E): Risk management—Risk assessment techniques

[4] ISO GUIDE 73:2009(E/F): Risk management—Vocabulary

[5] ISO/TR 31004:2013(E): Risk management—Guidance for the implementation of ISO 31000

[6] ISO 9001:2008: Quality management systems—Requirements

[7] ISO/IEC 25010:2011: Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models

[8] ISO/IEC 27001:2013: Information technology—Security techniques—Information security management systems—Requirements

[9] ISO 16363:2012: Space data and information transfer systems—Audit and certification of trustworthy digital repositories

[10] A. Osterwalder, “Business Model Generation”, Alexander Osterwalder & Yves Pigneur, 2009

[11] A. Osterwalder, “The Business Model Ontology—A Proposition in a Design Science Approach”, PhD thesis. University of Lausanne. 2004

[12] M. Schliemann, “BMI? Of course, but what about the Model Risks?” [Online]. Available: http://beyondeconomics.es/bmi-of-course-but-what-about-the-model-risks/

[13] IT Governance Institute. COBIT 5 – A business Framework for the Governance and Management of Enterprise IT, 2012

[14] IT Governance Institute. COBIT 5—Enabling Processes, 2012

[15] M. Baker, “The Strategic Marketing Plan Audit”, Cambridge Strategy Publications, 2008

[16] LeanStack, “Lean Canvas—1 Page Business Model”. [Online]. Available: http://leanstack.com/

[17] J. Parrisius, “(Talk:) Business Modeling to Reduce Risk”. [Online]. Available: http://juliusparrisius.wordpress.com/2013/03/25/business-modeling-to-reduce-risk/

[18] S. McAfee, “Why Do For-Profits Get All The Best Resources?! 2 Tools Every Nonprofit Can Use to Manage Risk”. [Online]. Available: http://www.leanimpact.org/nonprofit-program-development-risk-management-tools/

[19] Digital Curation Centre, “What is digital curation?”. [Online]. Available: http://www.dcc.ac.uk/digital-curation/what-digital-curation

[20] The Consultative Committee for Space Data Systems, Space data and information transfer systems – Open archival information system – Reference model – Magenta Book. June 2012. CCSDS 650.0-M-2.