

An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS

Anna S. Palaiologk · Anastasios A. Economides ·
Heiko D. Tjalsma · Laurents B. Sesink

Published online: 20 May 2012

© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Financial sustainability is an important attribute of a trusted, reliable digital repository. The authors of this paper use the case study approach to develop an activity-based costing (ABC) model. This is used for estimating the costs of preserving digital research data and identifying options for improving and sustaining relevant activities. The model is designed in the environment of the Data Archiving and Networked Services (DANS) institute, a well-known trusted repository. The DANS–ABC model has been tested on empirical cost data from activities performed by 51 employees in frames of over 40 different national and international projects. Costs of resources are being assigned to cost objects through activities and cost drivers. The ‘euros per dataset’ unit of costs measurement is introduced to analyse the outputs of the model. Funders, managers and other decision-making stakeholders are being provided with understandable information connected to the strategic goals of the organisation. The latter is being achieved by linking the DANS–ABC model to another widely used managerial tool—the Balanced Scorecard (BSC). The DANS–ABC model supports costing of services provided by a data archive, while the combination of the DANS–ABC with a

BSC identifies areas in the digital preservation process where efficiency improvements are possible.

Keywords Digital preservation costs · Data archiving · Cost drivers · Activity-based costing · Cost model · Balanced scorecard · Cost estimation · Performance measurement · Strategic planning · Funding · Organisational structure · Total quality management · DANS · ABC · BSC

1 Introduction

Digital information has become a vital component of the knowledge economy in the twenty-first century. The public sector invests heavily in the creation and acquisition of digital information and has recognised the need to preserve essential information for use by future generations [37]. Digital archiving of data is an emerging issue. Universities and libraries are considering what to do with the datasets produced and used by their researchers. Governments and private organisations are wondering if it is worth preserving them. Costs are a problem. Long-term commitment (and its cost) is a problem too. The challenge lies within the fact that the core funding for institutions of digital archiving does not grow in line with information growth and new acquisitions, which makes it difficult to justify and make a case for digital preservation to institutional management [24]. On the other hand, financial sustainability is an important attribute of a trusted, reliable digital repository [5]. The information level provided to the management by annual cost figures (salaries, hardware depreciation, etc.) is not sufficient any more. Some of the reasons are the continuous, disproportionate increase in both the quantity of data and the complexity of datasets, the rapidly changing preservation strategies and a substantial amount of continuously changing research and development

Data Archiving and Networked Services (DANS).

A. S. Palaiologk
University of Macedonia, Amygdaleonas, Kavala 64012, Greece

A. S. Palaiologk (✉) · H. D. Tjalsma · L. B. Sesink
Data Archiving and Networked Services (DANS), Postbus 93067,
2509 AB Den Haag, The Netherlands
e-mail: palaiologk.anna@gmail.com
URL: www.dans.knaw.nl

A. A. Economides
University of Macedonia, Egnatia 158, Thessaloniki 54006, Greece

activities supporting and improving the archiving work. More detailed figures are required to enable better decisions and thus sustain the future of the archiving entity and its data.

Ideally, a cost model should enable accurate planning and ensure a realistic strategy by providing the management of an organisation with useful information. A good cost model eases forecasting and control and provides a basis for accountability and transparency. Of major importance are also characteristics like relevance, accuracy and usability. Relevance of the developed cost model to the reality of the archiving practices in the organisation ensures understanding of the model by employees and managers. Accuracy (low percentage of approximate estimations) and usability are key factors for efficiency.

Studies concerning costs of digital preservation started being published only recently. They all have added much to our understanding of the factors affecting a cost model in this field [18, 35, 36]. In general, the cost data of the existing studies are related to specific projects, institutions or materials, and therefore difficult to transfer to other contexts [6]. That is because it is advised that costs of preservation are determined for specific digital assets using specific technologies, at a specified level of reliability and in a specific institution [6]. The most relevant published studies investigating our problem are: the report of NASA Cost Estimation Toolkit (CET) [7, 19, 26, 28], the LIFE project [2, 34] and the Keeping Research Data Safe (KRDS) project [4]. All the above mentioned studies are based or mapped against the Open Archival Information System Reference Model (OAIS RM) [13] and are using in one way or another the activity-based costing (ABC) approach. The OAIS RM is an ISO standard (ISO 14721:2003). The purpose of this standard is to establish a system for archiving information, both digitalized and physical, with an organisational scheme composed of people who accept the responsibility to preserve information and make it available to a designated community. ABC is a cost allocation method pioneered by Harvard's Cooper and Kaplan [12], in the field of management accounting. ABC has been successfully applied to manufacturing and service industries [23, 32, 41] for improving tactical and strategic decision-making and for enhancing corporate cost control and customer profitability [8, 47].

Our project started with a narrow problem focusing on creating and testing a model which would estimate the costs of research datasets preserved by a trusted digital repository [42]. In the process, the scope of the project expanded looking for a way to express the true value of the organisation to its stakeholders. In this paper, we present both the cost model generated and the spin-offs of the project which enable the achievement of financial sustainability.

The level of difficulty of applying an ABC model in an organisation depends on how standardised the processes of the industry it belongs to are. The activities of the archiving

and preservation industry have not been standardised enough yet; thus, we use the case study approach to induct our model. The generated model can be used in different environments; however, the results it produces might differ, reflecting each time the philosophy and problems of that particular organisation. We present the results for our case organisation as an example of how the model can be exploited.

This paper is organised in 6 more sections. Section 2 contains the methodology which was followed and a step-by-step description of the design of the DANS-ABC model. Some previous studies and their usefulness for our project are examined in Sect. 3. In Sect. 4, the components of the DANS-ABC model are described and its use is explained. In Sect. 5, the results of testing our model on empirical cost data on long-term digital archiving activities are presented. Section 6 illustrates how costs can be linked to strategy with the help of a Balanced Scorecard (BSC). Interesting conclusions and future recommendations are stated at the end of the paper, in Sect. 7.

2 Methodology

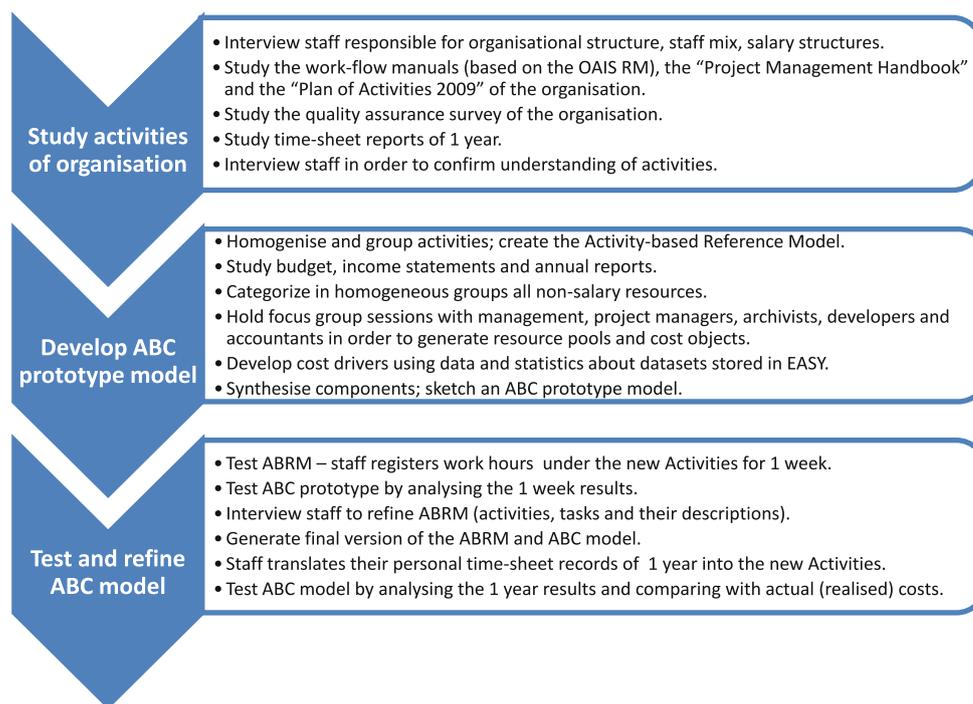
The estimation of costs of preserving digital data archives is a rather new and quite specific research area with only a few experiences. Case studies are seen as the best possible method for acquiring inside information about a new area of interest [17]. Working with a concrete case study allows us to both apply and test techniques of existing generic models while at the same time introduce a number of new features to solve problems that have not been addressed so far. That is why the case study research approach is chosen for this paper. More precisely, we follow the roadmap for building grounded theory from case study research proposed by Eisenhardt [17].

As for the costing method, as already stated in the introduction to this paper, it was decided to follow the ABC approach. The decision is supported by following three facts:

- (1) Staff is the major resource pool in digital archiving, up to 65–70 % of total expenses [4].
- (2) The work processes of DANS are modelled according to the OAIS Reference Model. This reference model consists of a list of functional components (activities) of a system dedicated to the long-term preservation of digital information; an initial set of activities.
- (3) ABC is a widely applied and tested managerial tool with many successful examples of use in both private and public sectors; a safe path to follow.

In developing and applying the ABC approach, we borrowed some standards and practices for developing successful estimation models used in software engineering. These included:

Fig. 1 Process followed to develop the DANS–ABC model



transparency of the model, usefulness of the model, developer participation, feedback, user training and a goal-oriented approach [15, 20, 22, 40]. The stages involved in developing our ABC model for digital archiving and the methods we used for our data analysis are summarized in Fig. 1.

There are three broad stages:

- (1) Study of the current archiving process and all other activities performed in DANS,
- (2) Development of an ABC prototype model and
- (3) Refinement of the ABC prototype model and generation of the final version of the model.

To refine the ABC prototype model (stage 3), we followed three guidelines proposed by Horngren et al. [25]: we classified as many of the total costs as direct costs as is economically feasible, we expanded the number of indirect cost pools until each of these pools is more homogenous and we used the cause-and-effect criterion, when possible, to identify the cost-allocation base for each indirect-cost pool.

2.1 Site selection

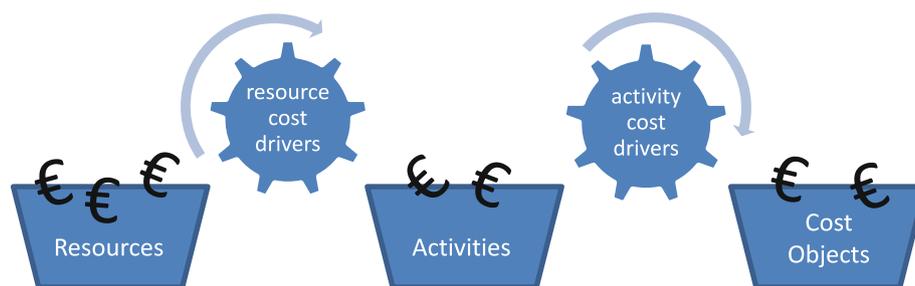
DANS—Data Archiving and Networked Services—is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW), and is also supported by the Netherlands Organization for Scientific Research (NWO). Since its establishment in 2005, DANS has been providing storage of and continuous access to research data mainly in the social sciences and humanities. It is a national disciplinary trusted repository for long-term preservation of research data.

The on-line access and support services it provides are free of charge. The collection of its repository comprises of a respectable amount of heterogeneous datasets: 14.000 datasets with a size of 1,5 Terabytes available to public and 10 datasets not available to the public with a size of 20 Terabytes. Data depositors are individual academic researchers, research centres and institutions, government departments and the national statistical agency. The first dataset of the collection (P0001) contains data of a survey from 1962. It was archived in 1964. To support its workflow and to provide access to its data collection DANS has developed its own platform—the DANS Electronic Archiving System (EASY). In general, the institute is using innovative techniques and has a large number of ongoing frameworks, standards and software development projects on both national and international levels. All activities are performed within a mixed budget of approximately 3,8 million euro a year. Since 2007, DANS has been using the web-based time-sheet software of Replicon, Inc. and has collected reasonably detailed records on how its staff has spent their working hours. The many and diverse activities of DANS are reflected in the time-sheet records. Both these records, as well as statistics on characteristics of datasets retrieved from EASY were essential information sources for the development of the ABC model.

2.2 Approach chosen

ABC is an integrated method of estimating costs by means of individual activities as the basis for assigning costs to other cost objects such as products or services, since the latter ones use the resources of an organisation in a non-uniform

Fig. 2 Components of an ABC model and their role



way [25]. It provides the management team with information to understand the use of scarce organisational resources in various business activities. Managers can then focus on areas of high cost, identify the factors that influence these costs, benchmark performance, and quantify improvements in the time and cost of the activities performed [39]. ABC is a complement to total quality management (TQM). It provides quantitative data that can track the financial impact of improvements implemented as part of the TQM initiative [38].

In the ABC approach, resources are first traced to activities and activity costs are then traced to products/services. This principle is fundamentally different from the traditional costing system that assumes that products/services utilize resources directly and inextricably.

More precisely, according to the ABC method, the organisational resources are logically grouped into resource pools first. Then, various tasks performed in the organisation are grouped into major and functionally homogeneous activity pools. These two components are termed cost pools. The component used to assign resource costs to activities is called a resource cost driver. Costs of the resources are assigned to activities based on their use of those resources, and activity costs are reassigned to cost objects (outputs) based on the proportional use of those activities by the cost objects. This mechanism is illustrated below in Fig. 2.

2.3 Data collection

One major advantage of this study is the reliance on both quantitative and qualitative data. Three types of data are gathered: archival records, interview data and direct observations.

Archival records, including time sheet reports, financial accounts, project manuals, workflow descriptions and policy handbooks, cover the period 2005–2009. Interviews with staff and external partners and focus group discussions on activities, drivers and explanation of variances provided a better understanding of the data archiving process in the organisation. These were conducted between October 2008 and May 2009. During the same period the author was permitted to observe meetings of different departments and projects and be present at conferences and gatherings of the international projects the organisation participated in. Finally, driver

counts were collected from statistics kept in EASY. These were used to test the ABC model.

We based the creation of the DANS–Activity-based Reference Model (DANS–ABRM), one of the most important components of the model, on information extracted from time-sheets of employees. The final shaping occurred after personal interviews with the 38 employees working in DANS during the study conduction. Quantitative data on staff costs were retrieved from their time-sheet records by mapping them against the new names and meanings of activities (as described in the ABRM). The employees did the mapping themselves. In order to avoid misunderstandings and mistakes, a ‘help-desk’ was available during the whole process for questions and clarifications and ‘clarifications’ group meetings took place. For people who had been hired for a limited period of time and were not working in DANS anymore the responsibility of transferring activities of the past in the new template was allocated to their supervisors. The final result represents the annual workload of 51 employees—48,461.40 working hours in total. The time frame we selected to work with spans from August 2007 till July 2008—a sample of one year. Earlier time-sheet records were not available.

3 Previous studies

As we have already mentioned above, we found out that the most relevant studies investigating our topic are the NASA CET, the LIFE project results and the KRDS. NASA CET estimates the lifecycle costs from implementation through a time-limited period of operations for NASA science projects [7, 19, 26–28]. LIFE is inspired by a lifecycle costing methodology originally developed for paper-based library collections [48, 49] and further refined for digital materials [46]. KRDS is oriented towards the preservation of research data in Higher Education Institutions (including national data centres hosted within them in the UK) [4].

A common observation after reviewing existing studies would be that the level of detail and the developed components were not suited to cover the needs of our case study organisation. Another specificity of previously related work is the development of a cost model for entities where digital preservation and dissemination are only parts of a wider

Table 1 Comparison of existing cost models

Study	Advantages	Disadvantages
CET [7, 19, 26–28]	<ul style="list-style-type: none"> + Good, detailed structure + Most of the ABC elements present + A prototype suite of Excel-based tools and a database of comparable costs 	<ul style="list-style-type: none"> – Internal design (based and adapted to the NASA needs only) – No clear distinction between costs of staff activities and other costs – Big differences in context of activities, nature of datasets and organisational aims – Inapplicable unit of measurement and cost objects – Does not address long-term archiving requirements
LIFE [2, 32]	<ul style="list-style-type: none"> + Based on a comprehensive review of existing lifecycle models + Methodology to predict costs of preserving digital information + Good, more suitable categorization of activities 	<ul style="list-style-type: none"> – Difficult to apply cost objects – Developed primarily for library use – Treatment of some direct and most indirect costs is poorly developed – Activities do not cover full range of work done in an independent archiving entity – Costs in each stage are treated as independent elements; not connected – Inflexible nor easily adjustable
Keeping Research Data Safe (KRDS), [4]	<ul style="list-style-type: none"> + Key cost variables and potential units of record are clearly listed + Does not concentrate only on archive phase + Model is tested with four case studies + Includes a benefits model potentially useful for strategic decisions 	<ul style="list-style-type: none"> – Mainly a resources and activities framework not strictly developed as an ABC model – Categorisation of activities (in phrases like “Pre-archive”) is confusing/unsuitable for DANS – Amount of activities was not sufficient to cover full work aspect in DANS – Unsuitable level and type of detail of activities for DANS (a need for a middle-ground solution was apparent) – A generic model leaving a lot of space for different interpretations and adaptations, e.g. no reference to stakeholder performing each activity (data provider, data manager, etc.); a very time-consuming activity – No specific unit of measurement proposed

range of activities. Finally, all existing studies estimating data preservation costs either focus on very specific digital objects such as microfilms, e-mail batches, e-journals and websites or they estimate costs in a very arbitrary way, e.g. per Gigabyte. None of these cost objects or units of costs measurement are optimal for research data.

Table 1 presents an analysis of advantages and disadvantages of these studies from the perspective of DANS; from the perspective of considering them as alternative solutions for estimating costs of digital preservation of research data of humanities and social sciences.

4 Proposed model

Here, we attempt to give an insight to how the developed ABC model is structured, what it consists of and how it works. It is not, though, a thorough presentation of the model and its

components since such an analysis would exceed the limits of a journal paper.

The proposed model is presented in four steps. First, the components of the DANS–ABC model are stated. A full activity analysis follows, where identification of activities takes place and their categorisation is explained. After that, activities are connected to the cost objects and cost drivers are identified and described. Finally, we explain the way the cost model is supposed to work, step by step.

4.1 Components of the DANS–ABC model

The DANS–ABC model is comprised of five building blocks:

- Resource pools,
- Resource cost drivers,
- Activities (grouped),

- Activity cost drivers, and,
- Cost objects.

These building blocks are linked to each other and their role is to facilitate our understanding of how the business of the organization is structured. Below, we describe each of the above components in more detail and we analyse them in their elements.

Resource pools are economic elements that are the sources of cost. In our study, they emerged from grouping the different budget items (costs) mentioned in the income statements and annual reports of the organisation. The logic behind this is that after money is allocated to a specific budget item, this item becomes a resource that the organisation can use to pursue its goals—to preserve and disseminate datasets. As a result, all costs of the organisation can be tracked down to the following resource pools:

Non-salary resource pools

- Office
- Data acquisition
- IT Services and Equipment

Salary resource pool

- Staff

Costs associated with the *Office* resource pool can be distinguished to ‘business-sustaining costs’ and ‘costs related to activities’. Business-sustaining costs include rent of an office, furniture, consumables, food, banking fees and communication (telephone, mail, etc.).

Given the nature of the organisation, there is no direct relationship of these costs to current cost objects. Under costs related to activities we list conference travel costs, representation expenses, trainings, subscriptions to magazines/literature, membership fees for scientific organisations, conferences/workshops expenses and printing & promotion. The relationship of these costs to cost objects can be established through the activities that they support using cost drivers.

Costs associated with the *Data Acquisition* resource pool are costs of buying data from Statistics Netherlands (CBS) and digitization services. They can be directly linked to cost objects.

Costs associated with the *IT Services and Equipment* resource pool consist of hardware expenses for example for internal storage and of software expenses which are included in the IT services provided by a third party organisation.

Costs associated with the *Staff* resource pool represent payment of all employees of the organisation including outsourcing.

Resource cost drivers are the best single quantitative measures of the frequency and intensity of demands placed on a

resource pool by activities or cost objects. It is used to assign costs from the resource pools to activities. Concrete allocation of non-salary costs to cost objects exceeds the purposes of this paper and is a matter for further research. Below, we analyse the resource cost drivers created for the resource pool ‘Staff’.

There are four resource cost drivers for the resource pool ‘Staff’. They determine both the content of the work of employees (including outsourced people) and the level of their payment. The resource cost drivers for the ‘Staff’ of DANS are: Archive, General, ICTa and ICTb. ICT is the group containing most employees and having the biggest variance between the lowest and highest payment rate. ICTa represents the more permanent, lower payment rate employees, while ICTb contains mainly outsourced experts with a much higher payment rate.

Activities of DANS are listed and described in the Activity-based Reference Model of DANS. The DANS-ABRM was formed based on the activities taking place in the institute and in combination with the available literature. It is important to note that the DANS-ABRM is not a lifecycle. It is a reflection of the operational functions taking place in a modern archive.

Activity cost drivers measure the frequency and intensity of the demands placed on activities by cost objects [39]. Along with the resource cost drivers, they are the key components of this study. Activity cost drivers are the elements that enable the calculation of future costs. A cost driver can be any situation or event that causes a change in the consumption of an activity or influences its quality. An activity may have multiple cost drivers. Cost drivers do not necessarily need to be quantified.

As *cost objects*, datasets of the three most common disciplines in DANS (history, social sciences and archaeology) were selected for our testing. In the future, with several adjustments, the DANS-ABC model can effectively be used for other-than-datasets cost objects too, such as services or customers.

The overall DANS-ABC model resulting from the analysis is shown in Fig. 3. Note that presentation of non-salary resource cost drivers exceeds the purposes of this paper and that due to the lack of space only some examples of activity cost drivers have been depicted. You can contact the authors for the full lists of the DANS-ABC elements.

It is important to note here that the ABC employment does make the lump of unallocated overhead costs much smaller, but it does not eliminate it completely. It is not possible to assign all costs to cost objects because many times there is no meaningful method to do so. Assigning overhead costs to cost objects in an arbitrary manner is not recommended since it would distort the real picture—destroy the ability to make meaningful relative judgements of product costing. Arbitrary allocation of overhead costs to cost objects would prevent the organisation from seeing the real relationship between cost drivers and those costs that are influenced by the drivers.

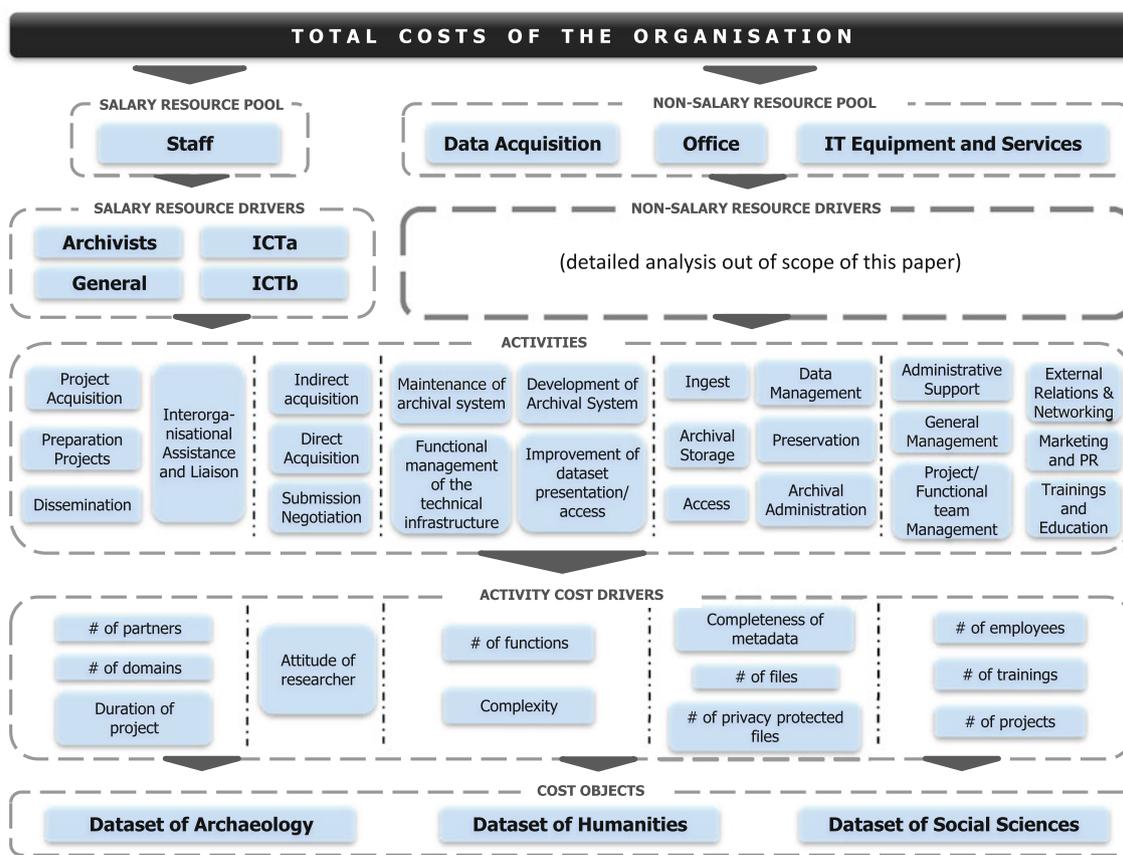


Fig. 3 Schema of the DANS-ABC model, an overview of the financial flows

4.2 Activities analysis: the DANS-ABRM

The *DANS-Activity-based Reference Model (DANS-ABRM)* describes activities taking place in a trusted repository. It is one of the most important components of the DANS-ABC model. The activities are conceptualised based on the OAIS reference model. In the activities description, no reference to DANS-specific terminology or particular technical details takes place. Organisation-specific parts of the DANS-ABRM are the categories of the activities and the tasks. Both categories and tasks are provided for the guidance of users and can be contextualised to the structure and language of the organisation.

Also, it is important to note that no activities taking place in an archive are excluded from the DANS-ABRM despite the fact that we always refer to costs of ‘preservation’. To support this decision of ours further, we would like to quote the definition of an archive from the OAIS study [13]: ‘Archive is an organisation that intends to preserve information for access and use by a designated community’. Accessibility and usability are as important as preservation, but require a different set of activities than the ones enabling preservation.

The activities of DANS are grouped in five categories:

- Networked services
- Data acquisition
- ICT (R&D and maintenance)
- Archiving
- Administration

Under these five categories, 23 activities in total are listed; each of which is provided with a short description and a set of tasks. The relationship of categories, activities and tasks is presented in Fig. 4. Table 2 contains the DANS-ABRM itself.

The DANS-ABRM, as any other reference model, is intended to improve communication between individuals involved in using the concepts it describes. A reference model breaks up a problem into entities, or ‘things that exist all by themselves’. This is often an explicit recognition of concepts that many people already share, but when created in an explicit manner, a reference model is useful by defining how these concepts differ from, and relate to, one another.

Another use of a reference model is to allow the comparison of different things. By breaking up a problem space into basic concepts, a reference model can be used to examine two different solutions to that problem, to compare how well each of the candidate solutions can be configured and to meet

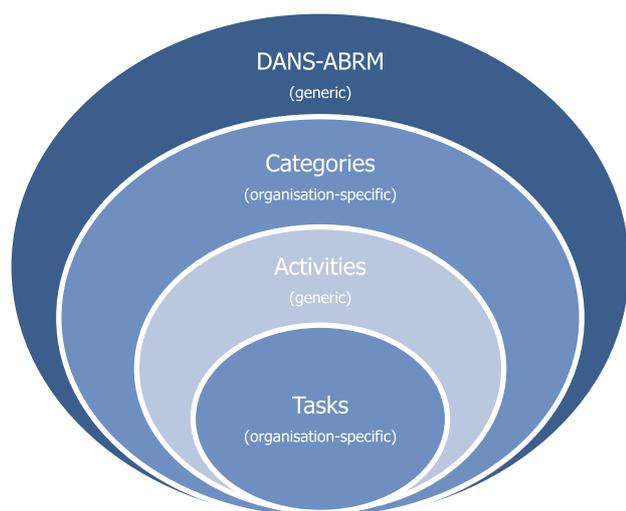


Fig. 4 The overlapping relationships of categories, activities and tasks in the DANS-ABRM

the needs of a particular business process. In doing so, the component parts of a solution can be discussed in relation to one another. That way, the DANS-ABRM can be mapped against other reference models and solutions (OAIS, LIFE, KRDS, NASA CET).

4.3 Cost drivers: connecting activities to cost objects

All the activities of the organisation are listed in the DANS-ABRM. One or more cost drivers correspond to each activity. The cost drivers connect the activity to the cost object (Fig. 4). On the left side of Fig. 5, one can see a concrete example of how cost drivers relate to the activities of *Direct Acquisition* and *Submission Negotiation*.

When establishing connections between activities and cost objects, it is important to consider that not all activities of the DANS-ABRM relate to each and every dataset. There are ‘easy’ datasets which only need to be published in the archival system (EASY) and there are datasets which require extent checking, conversions, new legal documents and development of new technologies before it is possible to access them. The extent and variety of activities performed for a dataset depends on its complexity and is a matter of

- (1) Technical characteristics of the dataset and
- (2) Quality standards (strategic decisions/policies) that the archive sets for its datasets.

The quality standards define the type of work/activities needed to be carried out for each dataset and the quality of the equipment/software which is used. The technical characteristics define the duration of these activities and the actual material costs. The quality standards and technical character-

istics of a dataset are broadly used as cost drivers in our study. The values given to the cost drivers indicate how complex or large a dataset is.

4.4 How to estimate costs by means of the DANS-ABC model?

The exact procedure to follow to estimate costs with the DANS-ABC is described with the help of a simplified example in Fig. 5.¹ The selections of cost object, resource pools, resource drivers and activities made by the user appear on the left side. The values that the user inserts and the calculations done by the model appear on the right side.

Let us assume, we want to know how much it did cost us (or it will cost us) to acquire a certain *dataset of archaeology*. The resources used for that include the personnel and the material costs involved. An *archivist* and an employee of *general responsibilities travelled* to a meeting with a client/potential depositor. The activities they performed appear on the left side of the figure. In order to calculate the exact time spent on these activities, the user needs to give values to the activity cost drivers. The time spent on both activities is dependent on whether the client is an *existing or a new customer* (is he familiar with the services provided by the archive and its regulations?) and on the *attitude of the researcher* (is he willing to cooperate?). The choices *existing or new* and *cooperative or non-cooperative* correspond to fixed values estimated by the data from previous experience of the organisation. As a result, after the calculations taking place on the right side of the figure, we have a certain amount of money indicating the acquisition costs for this specific dataset. Acquisition costs are only a part of all costs related to the preservation of the dataset. In a similar way, all other costs can be attached to the dataset. A description of the concrete formulas connecting all the DANS-ABC elements exceeds the purposes of this paper.

5 Grounding the DANS-ABC in empirical data: examples of findings and discussion

In this section, we will look into how the model worked for DANS. The testing of the model generated quantitative data which, if accumulated and interpreted over time, can track the financial impact of any changes implemented in the organisation. Senior managers view such knowledge as an important organisational resource [1, 3, 33], and its acquisition, articulation, and enhancement over time contributes to the uniqueness of the organisation [16]. In the following sections, we elaborate on the potential use of DANS-ABC

¹ The numbers in the figure are fictional.

Table 2 Activities of the DANS–ABRM with examples of tasks into which these activities translate in DANS

Category	Activities	Description	Tasks
Networked services	Project acquisition	Market research for projects that could potentially contribute to the aims of the organisation (e.g. generate research data) ^a	Reporting on ideas and tasks Review/assess/feedback Communication with partners
	Dissemination	Increasing awareness amongst the potential data consumers and data producers ^a	Presentations about existing/new services
	Preparation projects	Preparing funding proposals and applications for projects that are internal initiatives ^a	Research and proposals Building Up a network of partners
	Interorganisational assistance and liaison	Providing technical assistance and information in liaison with other organisations, private individuals, educational institutions, etc. ^a	Develop community data standards and best practices Share development of preservation systems and tools
Data acquisition	Indirect acquisition	Promotion of data-preservation-and-access-friendly ideology that will potentially result into more research data deposits (initialization of policy making, depositor support, etc.) ^a	Outreach and depositor support Promotion of data reuse Initialize policy development Symposia Expansion into new disciplines
	Direct acquisition	The processes involved in acquiring research data for an archive ^c	Facilitate data creation Contracting with producers, clients, funders Retro archiving (ADA) Negotiate costs
	Submission negotiation	The specification of submission requirements for producers/depositors together with communication and negotiation with producers/depositors ^b	Agree on metadata Discuss data formats Agree on access rights
ICT (R&D and maintenance)	Maintenance of archival system	Maintaining the archival system supporting the core services needed to operate and administer the datasets ^a	Fixing bugs Monitoring system performance and operations Technical management Disaster recovery Testing
	Development of archival system	Further developing the archival system; providing an interface between new application software and the platform. ^{1,3}	Develop infrastructural systems Meetings with users/customers Develop new functions Testing
	Improvement of dataset presentation/access	Software and system design or standards development facilitating data discovery, presentation and access ^a	Requirements Analysis Facilitate discovery of data System Design Web design Generate customised software Trainings on usage of new technologies
	Functional management of the technical infrastructure	Selection of and communication with ICT and storage services providers. Hardware maintenance and setting up of service processes ^a	Setting the standards, choosing the providers Negotiating with the providers, setting up service processes, adapting processes Day to day taking part in the service processes (dealing with incidents, implementing changes)

Table 2 continued

Category	Activities	Description	Tasks
Archiving	Ingest	The “ingest” functional area includes receiving, reading, quality checking, cataloguing, of incoming data (including metadata, documentation, etc.) to the point of insertion into the archive. Ingest can be manual or electronic with manual steps involved in quality checking, etc. ^{d,e}	Receive submission (SIP) Quality check Verification SIP creation Error checking and correction
	Archival storage	Error checking and correction and update of Archival Information Package (AIP) entries ^a	Update database (AIP entry) Replace data (AIP) & updating
	Data management	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 ^e	Create metadata Structure data
	Access	Services and functions for populating and accessing a wide variety of data by the project. Services and functions which make the archival information holdings and related services visible to Consumers ^e	Publish data Check AIP Create jump off page Inform depositor (check) Storage hierarchy management Transfer to another archive
	Preservation	The services and functions for monitoring, providing recommendations, and taking action, to ensure that the information stored in the archive remains accessible over the long term, even if the original computing environment becomes obsolete. Preservation covers the process of performing actions on digital objects in order to ensure their continued accessibility. It includes evaluation and quality assurance of actions, and the acquisition or implementation of software to facilitate the preservation actions ^c	Backup Preservation Planning Checking data integrity Knowledge Monitoring Development of preservation system (software) Migration to preferred formats
	Archival administration	User support to users of the archival system and establishing and maintaining the archive’s standards and policies (e.g. initial format standards, documentation standards, model deposit agreements, the archive’s selection policy and procedures to be followed during the Ingest process). The latter normally involve a large initial effort to develop and then regular review and small updates over time and rarer major re-drafting ^c	Legal/licenses/rights protection

Table 2 continued

Category	Activities	Description	Tasks
Administration	Administrative support	Administrative support and control provided by office managers, personal assistants and secretaries. Management of supply chain, movement of goods, and recording of purchases and deliveries ^c	Users support Office management Correspondence Travel arrangements Catering arrangements Personnel paperwork, etc.
	General management	Management and administration at the data service provider level (“front office”) and direct management of functional areas. Overall responsibility for internal and external science activities, information technology planning, and data stewardship ^c	Legal Policy Finance Human resources Quality Assurance Visiting Conferences
	External relations & networking	Maintaining liaison with and assisting in the coordination of activities related to regional and international organisations and other authorities. Promoting the visibility and image of the archiving entity and its services ^a	Attending and organizing symposia Communication with partners Planning Editing publications/website/etc Advertising Branding Writing articles Press (release/news items/contacts) Newsletters and website update
	Marketing and PR	Creation and implementation of a communication plan, promotion of the archiving entity or a specific project, media relations, advertising, sponsorship acquisition and market research ^a	Visiting international project meetings Progress monitoring/reporting Communication with team/partners Project Planning Project marketing
	Project management	Planning, organising, and managing resources to bring about the successful completion of specific project goals and objectives ^a	Attending trainings
	Trainings and education	Trainings of new staff or education about new tasks, bonding trips, etc. ^a	Participating in company trips

^a Definition created in DANS

^b Definition from the LIFE project

³ Definition from the Keeping Research Data Safe (KRDS) Activity Model

^d Definition from the Cost Estimation Toolkit (CET) of NASA

^e Definition from the OAIS Reference Model

outputs in management control, decision-making and organisational learning.

5.1 Size of resource pools

Estimation of resource costs according to the resource grouping of the DANS–ABC model based on traditional accounting figures depicted the following results (Table 3).

It is obvious that staff is the major resource pool in DANS. This fact supports our decision to follow an ABC approach to establish a costing system for DANS. That is also the reason for giving priority to the activities analysis in this paper.

5.2 Cost- and labour-intensive activities

In this section, we present the results of mapping the hours of a full calendar year extracted from the time-sheets of 51

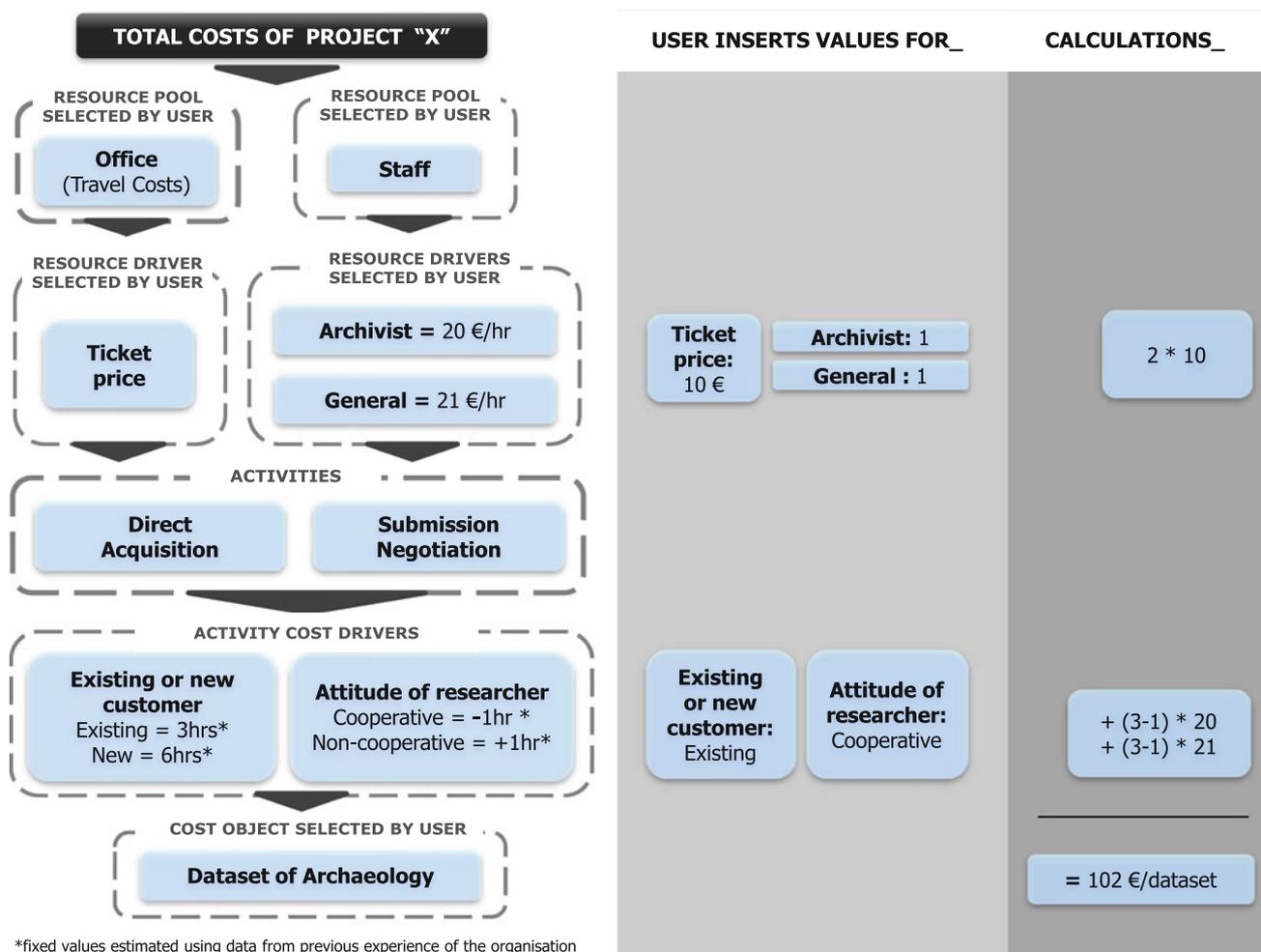


Fig. 5 Procedure and example of costs estimation using the DANS–ABC model

Table 3 Total costs per each resource pool of the DANS–ABC model

Data Acquisition (%)	Office (%)	IT services and equipment (%)	Staff (%)	Total ^a (%)
14.19	14.33	6.94	64.54	100

^a Average total of multiple years (2006–2008)

employees of DANS against activities of the DANS–ABRM. One of the main objectives of this study was to find out which are the most cost and labour intensive activities in DANS.

In Fig. 6, we notice that the most labour intensive activity is Preservation with a percentage of 17.02%. Other activities comprising very high percentages are: Development of Archival System, Project Management, Administrative Support, Improvement of Dataset Presentation and Access and General Management. An interesting observation is that all other activities included in the ‘Archiving’ category except Preservation have much lower percentages.

In Table 4, we can see and compare the cost and labour intensity of the top-rated activities. Cost-wise, once again, the activity of Preservation appears on the top. This high percentage clearly reflects the decision of the organisation to outsource a lot of work for a preservation project. Development of Archival System comes second and Project Management third. We would like to remind that in these estimations, costs of hardware or storage services are not included. These cost percentages represent staff expenses only.

The average time spent by each group of employees for each activity they carry out is presented in Table 5 (time spent on each activity divided by the total working time used by a group).

Based on Table 5, one can say the following about DANS:

- Archivists spend most of their time *managing data, ingesting and administrating archives*, but they spend almost the same amount of time on *marketing and project management*.

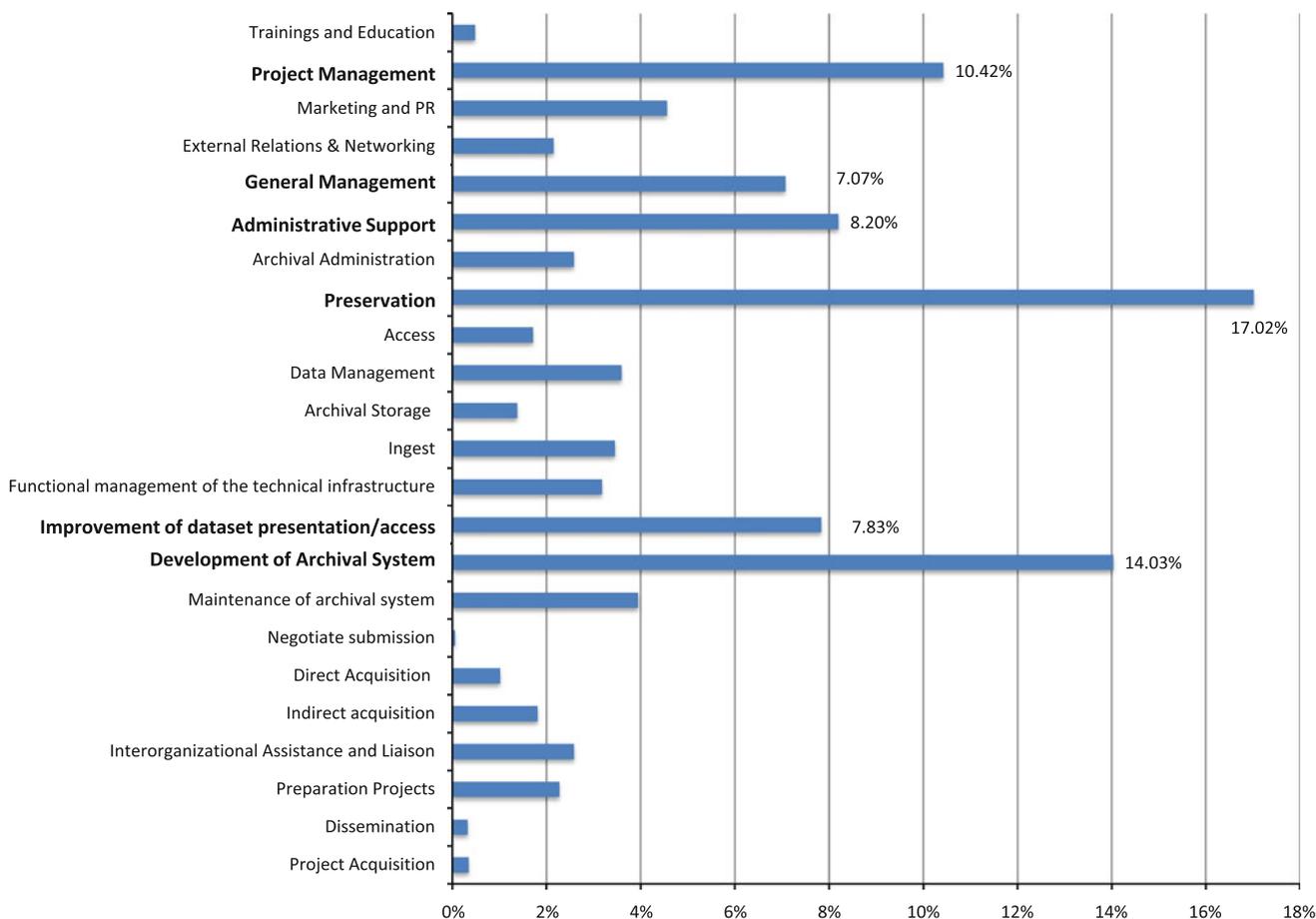


Fig. 6 Labour intensity of activities in DANS

Table 4 The most labour- and cost-intensive activities in DANS

Activity	Labour intensity (%)	Cost intensity (%)
Preservation	17.02	29.83
Development of archival system	14.03	14.43
Project management	10.42	7.85
Administrative support	8.20	7.44
Improvement of dataset presentation and access	7.83	7.08
General management	7.07	6.24

- ICT people spend more time on *developing the archival system* than on *maintaining* it, which usually indicates that the ICT department works well.
- ICT people spend twice more time on *Project Management* than people in the group General do. *Project Management* actually is the second most time consuming activity for ICTa people.
- The involvement of ICTb people in activities of DANS is little. That is explained by the fact that almost all of them are outsourced employees.

■ Archivists are the most ‘multitasking’ people, while ICT people are more specialised.

Whether the observations described above are ‘good’ or ‘bad’ can be examined with the help of the BSC of DANS (Fig. 8) analysed in Sect. 6 of this paper.

5.3 Archiving activities per discipline

The information provided in Fig. 6 and Table 5 can be analysed in even more detail. For example, we can look into the three main disciplines of the organisation.² For each discipline, there is a different mix of proportion of archiving activities (Fig. 7). The reason for the differences can be explained by the nature of the datasets (technical characteristics) and the policies applying to these datasets (preservation and dissemination quality standards) both described in Sect. 4.3.

² For Fig. 7, only the percentage of the preservation activities which is directly related to archiving of datasets for these disciplines was taken into account. The rest of the 17.02 % of preservation relates to a certain project (MIXED) running during our sample period.

Table 5 Labour intensive activities for each group of employees

Cat.	Activities (%)	Archive (%)	General (%)	ICTa (%)	ICTb (%)	Total time (%)
Networked services	Project acquisition	0.52	0.85	–	–	0.35
	Dissemination	–	0.71	0.43	–	0.32
	Preparation projects	5.13	3.33	0.72	–	2.27
	Interorganisational assistance and liaison	4.21	4.79	1.11	–	2.59
Data acquisition	Indirect acquisition	3.09	4.05	–	–	1.81
	Direct Acquisition	3.06	1.20	–	–	1.01
	Submission negotiation	–	0.21	–	–	0.06
ICT (R&D and maintenance)	Maintenance of archival system	1.32	–	12.60	–	3.94
	Development of archival system	1.08	0.12	32.53	20.60	14.03
	Improvement of dataset presentation/access	2.35	1.95	18.28	7.00	7.83
	Functional management of the technical infrastructure	–	2.55	8.52	–	3.18
Archiving	Ingest	15.56	–	–	–	3.45
	Archival storage	6.20	–	–	–	1.38
	Data management	16.17	–	–	–	3.59
	Access	6.82	0.71	–	–	1.71
	Preservation	1.28	4.42	1.10	72.18 ^a	17.02
	Archival administration	11.35	0.21	–	–	2.58
Administration	Administrative support	0.55	29.05	–	–	8.20
	General management	2.95	21.89	1.15	–	7.07
	External relations & Networking	1.16	4.16	2.54	–	2.15
	Marketing and PR	10.09	7.63	0.67	–	4.55
	Project management	6.49	11.96	19.36	0.23	10.42
	Trainings and education	0.60	0.20	1.00	–	0.48
	Total	100	100	100	100	100
	Group time/total time	22.20	27.79	28.98	21.04	100

^a The large percentage of costs for external ICT contractors sourced from a single preservation project (MIXED) running during our sample period

We have already referred to the nature of the datasets and the quality standards that apply to the preservation process in previous paragraphs. Here, we are giving concrete examples of how the above affect the proportions illustrated in Fig. 7.

Example 1 The time spent on *Data Management* varies a lot across the three disciplines depicted in Fig. 7. Archaeology accounts for the largest share at 44%, while social sciences for 25% and history for 0%. This is partially explained by the fact that the archival of archaeological datasets in DANS is a self-funded project (eDNA), which gives the project leader

the opportunity to define individual quality standards rather than adhere to the organisation's general standards. These project-specific policies are the following:

- File-specific metadata is added to all files of an archaeological dataset. Metadata file lists are requested from the depositors, but they always need extra processing to upload them in the system and synchronize them with the files.
- Data producers of archaeological datasets are allowed to deposit their data in non-preferred formats. Archivists

Fig. 7 Workload breakdown of archiving activities for datasets of history, social sciences and archaeology

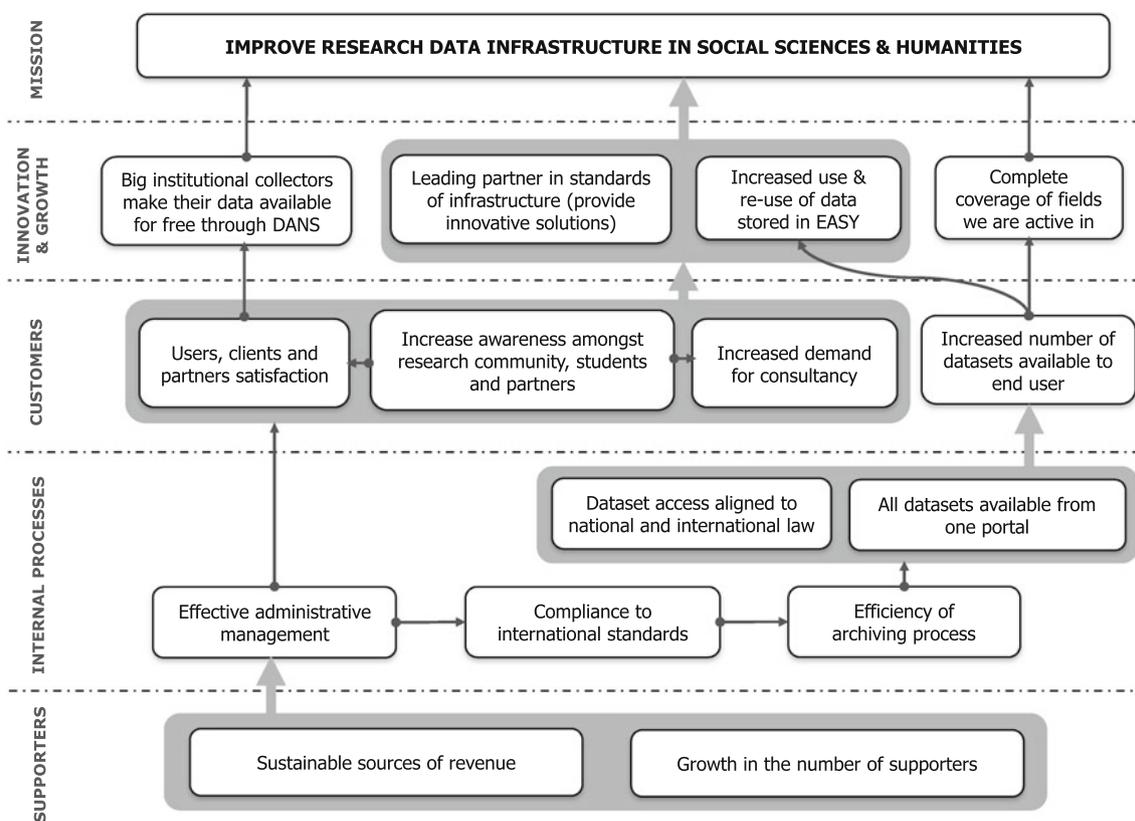
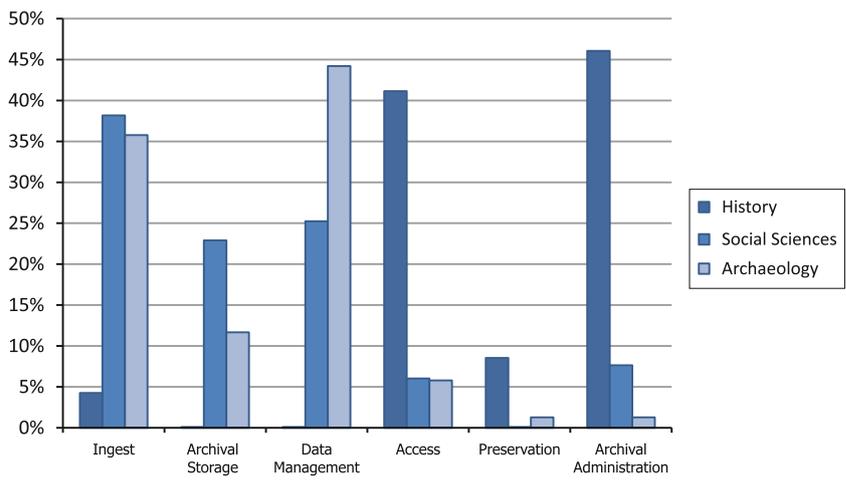


Fig. 8 The BSC of DANS

dealing with archaeological datasets convert these files in preferred formats.

Both the file format conversions and the addition of file-specific metadata take a lot of time and have never been generic tasks applying to all disciplines in DANS.

Furthermore, the latter policy goes hand in hand with the nature of the datasets. Archaeological datasets consist

of a much wider variety of file formats than other disciplines’ datasets. There are documents, spreadsheets, databases, images and even GIS and CAD cartography files. Other disciplines generally handle a smaller variety of data formats (SPSS, PDF).

The above explain the big difference in the percentages of time spent on *Data Management* from the perspective of the nature of datasets and quality standards applying to different disciplines.

Example 2 A percentage of 41% devoted to Access (see Table 5) and a huge 46% representing Archival Administration (legal issues and user support) in the archiving process for history is another example of how the nature of the datasets affect the proportions illustrated in Fig. 7.

The datasets of history are more ‘delicate’ compared to the ones in the other two sciences; they contain much more privacy-sensitive information. Special permissions are required which is translated into a lot of *legal work*, *users support* and special *access* procedures. All these tasks were carried out by the archivists during our target period increasing disproportionately the percentage of time spent on Access and Archival Administration.

5.4 Variance analysis for staff costs

In order to check the accuracy of the estimations, we compared the total of staff costs generated by the ABC model with the total stated in the income statements.³ This comparison revealed a variance of 17%, which can be further eliminated by reviewing interview notes and discussing with key-employees to understand its source.

6 Further use of ABC: linking costs to strategy

As we already said the BSC is a tool which helps an organisation to define its strategy and plan. In fact, no action or result can be clearly termed as ‘good’ or ‘bad’ in an organisation. There are only ‘planned’ and ‘not planned’ actions and results. The BSC translates an organisation’s mission and existing business strategy into a limited number of specific strategic objectives that can be linked and measured operationally [29–31]. Many public, private and not-for-profit organisations have adopted the scorecard as part of their strategic management approach [21]. The *espida* project is an attempt of using the BSC for digital preservation [14].

The strategy of DANS is aligned to the responsibilities mentioned in the OAI RM as ‘mandatory responsibilities that an organisation must discharge to operate an OAI archive’. The BSC of DANS, appearing in Fig. 8, is based on the Strategy Plan document of DANS created by the management and the steering group of DANS. A concise description is given below.

The BSC of DANS has four levels/perspectives:

- The financial perspective is called ‘Supporters’ and lies at the bottom of the BSC.

³ The costs of staff refer to the period August 2007–July 2008. A special customised report was conducted by the financial department to realise the comparison.

- The ‘Internal Processes’ of DANS are gathered at the second level. It represents in a way the goals of different groups of people working in the institute.
- The third level named ‘Customers’ contains the effects of the processes listed at the second level on users, partners and clients.
- Having succeeded in all these goals, we can precede with measuring our impact on the business and social environment of the organisation expressed in terms of ‘Innovation and Growth’.

The ultimate goal of DANS—its ‘Mission’—is to improve the research data infrastructure in the social sciences and humanities and it is the ‘head’ of the BSC. By viewing the organisation from all four perspectives, the balanced scorecard provides a more comprehensive understanding of current performance [50] which in turn helps organisations act in their best long-term interests.

DANS is a not-for-profit organisation and its budget is fixed. The income usually comes from long-term, repeat and performance-based funding sources—foundation, individual, government, corporate and fee-based—which provide predictable funding, despite conditions of market failure. That is why the financial dimension is placed at the basis of the BSC structure.

The long-term interest of DANS is to ensure that its *sources of revenue are sustainable* and that there is a *growth in the number of its supporters*. These two phrases in the language of BSC are called Success Factors (SFs). Success factors describe the strategic objectives of the organisation. Each of the 15 illustrated SFs is described further by a set of Performance Indicators (PIs). They are an indication of how it shall be known that the outcome has come to pass. Figure 8 contains all the SFs of DANS.

6.1 How does strategy influence costs?

The SFs reflect decisions which are not independent of the final costs an organisation bears. Let us analyse for example a major strategy of DANS—innovation. The SF representing it is listed in the dimension of *Innovation & Growth* at the top of the BSC structure under the title *Leading partner in standards of infrastructure*. Beagrie et al. [4] say the following about the ‘First-Mover Innovation’ costs: ‘where preservation functions and file formats are evolving a high-degree of R&D, expenditure might be required in implementation phases and in developing the first tools, standards and best practices. Many of the disciplines and archives covered in this study have made considerable investments as communities in evolving shared standards, practices, and tools and we believe this could be making a significant impact on their long-term digital preservation costs’. According to that description DANS

is certainly a ‘first-mover innovator’ and many activities described in the ABRM contribute to that. The underlying idea of this choice could also be described with the example from the Dutch Digitale Bewaring project cited by Beagrie et al. [4] ‘creating a batch of 1,000 records costs 30 times less than ‘repairing’ a batch of 1,000 records with badly created metadata’. Being a ‘First-mover Innovator’ might cost more in the short-run but might be profitable in the long-run (benefitting both the organisation and the society).

Another evidence of strategy and organizational performance affecting costs is the effort that is being put lately by various organizations in a number of countries to specify criteria for the certification of digital repositories. Amongst other goals, these criteria are intended to bring about greater efficiency in data management which in its turn would lead to costs minimization by, for example, minimizing the amount of manual work involved in the documentation, management and making available of data, for both researchers/data producers and data managers [43]. DANS in collaboration with eight international partners have also developed such guidelines, the so-called Data Seal of Approval (DSA) [45]. The DSA ensures that in the future, research data can still be processed in a high-quality and reliable manner, without this entailing new thresholds, regulations or high costs.

6.2 How to use the DANS–ABC and BSC to improve organisational performance?

Table 6 illustrates a BSC working template for goal-setting, evaluation and action steps indication. It is populated with one SF, its PI and the DANS–ABRM activities corresponding to them as an example. In the frames of our project, all SFs of DANS have been linked to activities from the DANS–ABRM.

The first column of Table 6 contains the SFs of the BSC. The second contains the corresponding PIs. There can be more than one PIs linked to each SFs. Targets of the third column are set periodically (e.g. yearly) for each PI during the planning process. The progress is examined periodically (e.g. quarterly) by obtaining data on the PIs. Further, the current state of PIs is compared to the goal set. If the planned results differ from the ones we got, a problem-solving action step has to be implemented. The ABC-elements (activities) work as a hint allowing management spot faster the source of the problem.

Let us assume that the annual target for ‘increasing the number of datasets available to end users’ is 10 contracts with researchers who will deposit the data coming from their projects to DANS. At the end of the year, the management notices a result of only three contracts signed. Something is not going as planned. To investigate the problem, we check which are the activities related to this specific PI. In our exam-

ple, the activities are: *Direct Acquisition* and *Indirect Acquisition*. The management can do two things:

- (1) Look into the ABC outputs and
- (2) Discuss the problem with the personnel responsible for the activities mentioned above.

The discussion with the staff does not necessarily mean that staff underperforms. A discussion with staff in combination with the ABC outcomes can reveal the source of the real problem. For example, it might be the case that during the year very few working hours have been dedicated to acquisition, because the staff involved in these activities is also responsible for too many other activities. Possible action steps solving this problem could be: to hire more staff or to re-allocate responsibilities.

7 Conclusions & future research

This study intends to help funders, managers and decision-makers in archiving entities understand the connections between strategy, resources, everyday work and ‘products’ of their organisations. In other words, this paper advances the state of the art on the cost analysis of digital preservation and access in the following ways:

- (1) It implements fully the ABC method identifying not only activities but also all its components—activities, resource pools, resource cost drivers, activity cost drivers and cost objects;
- (2) It proposes a new unit of costs measurement—euros per dataset—and an idea for cost object customisation by classification of datasets based on their complexity;
- (3) It links the cost model to the strategy of the organisation, bringing the decision making process to another level.

Because of the novelty of the approach in this context, the paper provides a detailed description of how the proposed ABC model was created. Then, it grounds it in empirical cost data to demonstrate how the model works in a real world environment.

The most challenging task of creating the model was to conceptualise the workflow and activities of an archive in a generic, but at the same time acceptable by the organization form. The difficulty sourced from the facts that many activities in this field are related to research and innovation—and not routine activities—and that archiving and preservation is a new industry and its processes have not been standardized enough yet. Many activities of the DANS–ABRM have not been previously described in the existing literature. The level of detail chosen for the description of the ABRM activities

Table 6 Working template of the BSC elements linked with elements of the DANS–ABC model

	BSC elements				Action step	ABC elements
	SF	PI	Target for PI (set during planning)	Current state of PI (obtained during evaluation)		
Customers	Increase number of datasets available to end user	# of contracts signed with research funders	[e.g. 10]	[e.g. 3]	[to be decided by management]	Direct acquisition Indirect acquisition

aimed at helping the management make decisions, not at putting a price-tag to each dataset. The ABC approach supplements the information available from traditional accounting methods with detailed costing figures that highlight the impact of different policies on costs. Another issue future researchers should take notice of is the fact that the most time and effort consuming stage of the study was the selection of cost data. Access to information about historical projects or types of datasets was limited. The latter was often a problem due to lack of good archival records and turnover of the personnel involved.

7.1 Lessons learned and recommendations

For the perspective of practitioners trying to replicate our work, to increase the benefits received, we recommend the following:

- (1) Activities and cost drivers must be clearly defined with a list of specific examples to simplify counts and minimize inter-ratter variances (heterogeneity occurring from the different ways users might perceive the elements of the model); providing users with contextualised tasks and drivers is likely to increase understanding of the model and accuracy in counts, resulting in more accurate time and cost estimates.
- (2) Strong employees' participation in identifying activities and cost drivers is necessary to minimize resistance to adoption of the model and to ensure that the model is relevant and useful [39, 44]. Involvement of the staff, in the contextualization of the elements of the model and in the results analysis, is very useful for organisational learning. It helps, for example, overcoming barriers like 'the lack of clarity in responsibilities for digital preservation' mentioned in the survey conducted recently by the Digital Preservation Coalition (DPC) [51].
- (3) Regular updates of the model which will reflect changes in organisational resource mix and archiv-

ing practices; otherwise, the estimates will be compromised [19].

- (4) Automated collection of the data required for the model (e.g. integration of ABC in time-sheets) to minimize the effort required to maintain the model.

All recommendations are essential for eliminating the implementation costs of the model derived from the need of the continuous presence of a 'model-expert' in the organisation.

Our model also has its limitations related to its transferability and extension. While the approach presented in this paper for the building of the model can be easily applied to other organisations, the model itself may require significant customisation before it can be used elsewhere. The extent of the customisation needed will depend on how close is the organisation's philosophy and structure to the one of DANS. The activities and cost drivers identified in this study reflect the philosophy of a trusted digital repository [42] compliant with the 16 guidelines listed in the Data Seal of Approval (DSA) [45]. Organisations following these guidelines will find it fairly easy to adapt the model to their needs. Activities and drivers of non-compliant organisations might differ more from the ones listed in this paper. Comparison of activities is possible using the description columns of the Activity-based Reference Model. The resources required to adjust the model to another organisation may introduce significant setup costs if the organisation is not familiar with the OAI model and terms. Comparison and transferability of some cost drivers might be more challenging since that depends on the preservation software used in the organisation. Also, in the future, more activities and drivers can be added to the model. However, the inclusion of additional drivers, while possibly increasing accuracy, will also increase the complexity of the model. The problems of increased model complexity include *user resistance*, *increased inter-ratter variances* among employees in activity and driver counts, and *delay in estimation* [9–11] and should be seriously taken into account.

7.2 Further research

Further research goals in regards to our model development include:

- Improving the *reliability* and *generalisation* of the results obtained from the data analysis by replicating the case study in similar environments,
- Investigating questions regarding the acceptance of the model in the organisation,
- Testing the process of estimating and predicting costs of datasets further to allow more general conclusions and
- Further analysing the DANS–ABC data and findings comparing them with findings of other studies.

All the above can eventually provide management with even more useful insights in their business and potentially eliminate costs further.

Acknowledgments We would like to thank all the employees of DANS for their active support and participation in this work. Special thanks goes to Valentijn Gilissen, Maarten Hoogerwerf, Rutger Kramer and Lucas Pasteuning.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Adler, P.S.: When knowledge is the critical resource, knowledge management is the critical task. *IEEE Trans. Eng. Manag.* **36**(2), 87–94 (1989)
2. Ayris, P., Davies, R., McLeod, R., Miao, R., Shenton, H., Wheatley, P.: The LIFE2 final project report. Research report. LIFE Project, London, UK (2008). <http://eprints.ucl.ac.uk/11758/1/11758.pdf>. Retrieved 5 Oct 2008
3. Baskerville, R., Pries-Heje, J.: Knowledge capability and maturity in software management. *DATA BASE Adv. Inf. Syst.* **30**(2), 26–42 (1999)
4. Beagrie, N.A., Chruszcz, J., Lavoie, B.: keeping research data safe: a cost model and guidance for UK Universities, JISC, London (2008). <http://www.jisc.ac.uk/publications/reports/2008/keepingresearchdatasafe.aspx>. Retrieved 5 Oct 2008
5. Bellinger, M.: Cost and business models for digital preservation: Developing digital life cycle management services at OCLC. In: 4th Forum of Digital Preservation Coalition “Preservation of e-Learning Materials and Cost Models for Digital Preservation”. London, 15 October 2002. http://www.dpconline.org/graphics/events/presentations/pdf/BellingerDPCForum_CostsBusinessModels.pdf. Retrieved 1 Nov 2009
6. Blue Ribbon Task Force (BRTF) Interim Report: Sustaining the digital investment: issues and challenges of economically sustainable digital preservation. Technical Report on Sustainable Digital Preservation and Access, National Science Foundation, December 2008. http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf. Retrieved 5 May 09
7. Booth, B., Banks, M., Hunolt, G.: Cost estimation tool enhanced operational comparables database (2006). <http://opensource.gsfc.nasa.gov/projects/CET/Doc.zip>. Retrieved 10 Dec 2008, 10 Oct 2008
8. Bradway, B., Ross, S.: Measuring corporate customer profitability: the role of activity-based cost analysis. *Corp. Cust. Manag.* **4**(Research Brief 6), 1–10 (2000)
9. Cokins, G.: *Activity-Based Costing: Making It Work*. Mc Graw Hill, Boston (1998)
10. Cokins, G.: Learning to love ABC. *J. Account.* **188**(2), 37–39 (1999)
11. Cokins, G.: Overcoming the obstacles to implementing activity-based costing. *Bank Account. Finance* **14**(1), 47–52 (2000)
12. Cooper, R., Kaplan, R.S.: Measure costs right: make the right decisions. *Harv. Bus. Rev.* **66**, 96–103 (1988)
13. Consultative Committee for Space Data Systems (CCSDS): Reference model for an open archival information system (OAIS), CCSDS 650.0-B-1 Blue Book (2002)
14. Currall, J., McKinney, P.: *espida Handbook* (2007). http://www.gla.ac.uk/espida/model_download.shtml. Retrieved 15 May 11
15. DeMarco, T.: *Controlling Software Projects: Management, Measurement and Estimates*. Prentice Hall, PTR, Upper Saddle River (1986)
16. Dodgson, M.: Organisational learning: a review of some literature. *Org. Stud.* **14**(3), 375–394 (1993)
17. Eisenhardt, K.M.: Building theories from case study research. *Acad. Manag. Rev.* **14**(4), 532–550 (1989)
18. ERPANET: Cost orientation tool (2003). <http://www.erpanet.org/guidance/docs/ERPANETCostingTool.pdf>. Retrieved 10 Dec 2008
19. Fontaine, K., Hunolt, G., Booth, A., Banks, M.: Observations on cost modelling and performance measurement of long-term archives. NASA research paper in PV2007 Conference Proceedings (2007). http://www.pv2007.dlr.de/Papers/Fontaine_CostModelObservations.pdf. Retrieved 10 Dec 2008
20. Grady, R.B., Caswell, D.L.: *Software Metrics: Establishing a Company-Wide Program*. Prentice-Hall Inc., Upper Saddle River (1987)
21. Gumbus, A., Lussier, R.N.: Entrepreneurs use a balanced scorecard to translate strategy into performance measures. *J. Small Bus. Manag.* **44**(3), 407–425 (2006)
22. Hall, T., Fenton, N.: Implementing effective software metrics programs. *IEEE Softw.* **14**(2), 55–65 (1997)
23. Helmi, M., Hindi, N.: Activity-based costing in banking: a big challenge. *J. Bank Cost Manag. Account.* **9**(2), 5–19 (1996)
24. Hockx-Yu, H.: Digital preservation in the context of institutional repositories. *Program Electron. Libr. Inf. Syst.* **40**(3), 232–243 (2006)
25. Horngren, C.T., Datar, S.M., Foster, G.: *Cost Accounting—A Managerial Emphasis*, 11th edn. Pearson Education, Inc., New Jersey (2003)
26. Hunolt, G.: Users’ guide cost estimation toolkit (CET), version 2.1 (2006a). <http://opensource.gsfc.nasa.gov/projects/CET/Doc.zip>. Retrieved 10 Dec 2008
27. Hunolt, G.: Technical description document cost estimation toolkit (CET), version 2.1 (2006b). <http://opensource.gsfc.nasa.gov/projects/CET/Doc.zip>. Retrieved 10 Dec 2008
28. Hunolt, G., Booth, B., Banks, M.: Cost estimation toolkit (CET), version 2.1 (2006). <http://opensource.gsfc.nasa.gov/projects/CET/CET%20V2p1.xls>. Retrieved 10 Dec 2008
29. Kaplan, R.S., Norton, D.P.: The balanced scorecard—measures that drive performance. *Harv. Bus. Rev.* **70**(1), 71–79 (1992)
30. Kaplan, R.S., Norton, D.P.: Putting the balanced scorecard to work. *Harv. Bus. Rev.* **71**(5), 134–142 (1993)
31. Kaplan, R.S., Norton, D.P.: Why does business need a balanced scorecard?. *J. Cost Manag.* **11**(ER), 5–11 (1997)
32. Kroll, K.M.: The ABCs Revisited. *Industry Week* (254:22), December, pp. 19–21 (1996)

33. Larsen, T.J., Levine, L.: Information systems: current issues and future changes. *DATA BASE Adv. Inf. Syst.* **30**(2), 7–11 (1999)
34. Mcleod, R., Wheatley, P., Ayris, P.: Lifecycle information for e-literature: full report from the LIFE project (LIFE Project, London, UK) (2006). <http://eprints.ucl.ac.uk/archive/00001854/01/LifeProjMaster.pdf>
35. Nationaal Archief: Costs of digital preservation, version 1.0, May 2005 (Digital Preservation Testbed, The Hague, Netherlands) (2005a). <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/CoDPv1.pdf>. Retrieved 10 Dec 2008
36. Nationaal Archief: Digital preservation cost model, version 1.0, 20 April 2005 (Digital Preservation Testbed, The Hague, Netherlands) (2005b). [http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Kostenmodel_in_Excel_versie_1.0_\(final\).xls](http://www.digitaleduurzaamheid.nl/bibliotheek/docs/Kostenmodel_in_Excel_versie_1.0_(final).xls). Retrieved 10 Dec 2008
37. Netherlands Coalition for Digital Preservation (NCDD): A future for our digital memory permanent access to information in the Netherlands: Interim report—summary in English Netherlands (2009). <http://www.ncdd.nl/en/documents/Englishsummary.pdf>. Retrieved 5 Jan 2010
38. Olsen, R.: Activity-Based Costing: A Decision-Making Tool, The Manufacturing Report. Lionheart Publishing, Inc., Atlanta (1998). http://www.offtech.com.au/abc/ABC_Resources.asp. Retrieved 10 Oct 2008
39. Ooi, G., Soh, C.: Developing an activity-based costing approach for system development and implementation. *DATA BASE Adv. Inf. Syst. Summer* (2003) **34**(3), 54–70 (2003)
40. Pflieger, S.L.: Lessons learned in building a corporate metrics program. *IEEE Softw.* **10**(3), 67–74 (1993)
41. Reimain, B.C., Kaplan, R.S.: The ABCs of accounting for value creation. *Plan. Rev.* **18**(4), 33–34 (1990)
42. RLG/OCLC Working Group on Digital Archive Attributes: Trusted Digital Repositories: Attributes and Responsibilities. RLG, Mountain View, CA (2002). <http://www.rlg.org/longterm/repositories.pdf>. Retrieved 15 May 2011
43. Ross, S., McHugh, A.: The Role of Evidence in Establishing Trust in Repositories. *D-Lib Magazine*, vol. 12, No. 7/8 (2006). <http://www.dlib.org/dlib/july06/ross/07ross.html>. Retrieved 15 May 2011
44. Roztock, N., Valenzuela, J.F., Porter, J.D., Monk, R.M., Needy, K.L.: A procedure for smooth implementation of activity based costing in small companies. In: Proceedings ASEM national conference, Virginia Beach, October 21–23, pp. 279–288, (1999)
45. Sesink, L., Horik, R., Harmsen, H. (eds): Data Seal of Approval—quality guidelines for digital research data. The Hague, Data Archiving and Networked Services, 2nd ed. DANS (2010). ISBN 9789490531027. http://www.datasealofapproval.org/sites/default/files/DSA%20booklet_2-0_engels_mei2010.pdf. Retrieved 15 May 2011
46. Shenton, H.: Life cycle collection management. *LIBER Q.* **13**, 254–272 (2003). <http://liber.library.uu.nl/publish/articles/000033/article.pdf>. Retrieved 10 Dec 2008
47. Soin, K., Seal, W., Cullen, J.: ABC and organizational change: an institutional perspective. *Manag. Account. Res.* **13**(2), 249–271 (2002)
48. Stephens, A.: The application of life cycle costing in libraries. *Br. J. Acad. Librariansh.* **3**, 82–88 (1988)
49. Stephens, A.: The application of life cycle costing in libraries: a case study based on acquisition and retention of library materials in the British library. *IFLA J.* **20**, 130–140 (1994)
50. Stewart, A.C., Carpenter-Hubin, J.: The balanced scorecard: beyond reports and rankings. *Plan. High. Educ.* **29**(2), 37–42 (2000)
51. Waller, M., Sharpe, R.: Mind the Gap: Assessing Digital Preservation Needs in the UK. Digital Preservation Coalition, York (2006). <http://www.dpconline.org/docs/reports/uknamindthegap.pdf>. Retrieved 10 Dec 2008