OCLC Systems & Services: International digital library perspectives

Scalable Decision Support for Digital Preservation: An Assessment:

Christoph Becker Luis Faria Kresimir Duretec

Scalable Decision Support for Digital Preservation: An Assessment

1. Introduction

This article continues the discussion in Becker et al (2014) and presents a systematic assessment and evaluation of the SCAPE decision support environment comprising PLATO, SCOUT and c3po. We discuss the improvements and identified limitations of the presented system. We furthermore discuss the quantitative and qualitative evaluation of advancing the state of art and report on a case study with a national library. Finally, we summarize the contributions and provide an outlook on future work.

2. Evaluation and assessment

While some of the questions that are raised by the design goals discussed in Becker et al (2014) can be readily evaluated using standard metrics, others require a detailed qualitative assessment. This section will discuss how to systematically assess improvements on the dimensions of trust and scalability. We will report on a typical case study conducted with the State and University Library Denmark, discuss key metrics that can be used for evaluation, apply them to assess recent advances, and discuss a set of limitations. We further discuss how these findings can be applied on a wider scale.

5.1 Evaluation dimensions and challenges

Five major design goals have been proposed in Becker et al (2014):

- **G1: Scalable content profiling** is required to create and maintain an awareness of the holdings of an organization, including the technical variety and the risk factors that cause difficulties in continued access and successful preservation.
- **G2: Monitoring compliance, risks and opportunities** is a key enabler to ensure that the continued preservation activities are effective and efficient.
- **G3: Efficient creation of trustworthy plans** is required so that preservation can function as a normal part of an organization's processes in a cost-efficient way.
- **G4: Context awareness** of the systems ensures that they can adapt to the specific situation rather than provide generic recommendations or require extensive manual configuration.

- **G5: Loosely-coupled preservation ecosystems**, finally, enable organizations to follow a stepwise adoption path and support continuous evolution of the preservation systems as new solutions and improved systems emerge.

Given this set of design goals, it is clear that a systematic evaluation has to be based on both qualitative and quantitative criteria and account for the various socio-technical dimensions of the design problem.

Scalable content profiling requires, first and foremost, efficiency in the data processing system. This can be measured in terms of the amount of data processed in a certain timeframe using a defined set of resources. This applies to the content profiling tool C3PO. The efficiency of decision making, on the other hand, can be measured in controlled experiments. This, however, has to be done in a real-world environment to be meaningful, which creates additional challenges for a large-scale assessment and has to be interpreted with caution.

The effectiveness of a preservation system composed of several heterogeneous and asynchronous processes, collaborating over time and controlled by decision makers in a real organization, is much harder to measure, since very often what needs to be measured in terms of the effects is time-delayed, and to a large extent defies objective measures in the present time. Similarly, trust is extremely hard to measure, and the preservation community has for a decade discussed different ways of assessing the trustworthiness of a repository (Ross & McHugh 2006; OCLC and CRL 2007). The resulting criteria catalogue ISO 16363 (ISO 2010) provides a useful checklist for assessing the assumed trust of an organization and hence can form a guideline for evaluation, but does not apply to the actual operations and the preservation lifecycle on the operational level. The Plato planning approach that forms the basis for the architecture presented here has been designed with these criteria in mind and evaluated for adherence with and support of these criteria (Becker et al. 2009). However, it can be argued that more holistic perspectives are required to assess and improve an organization’s trustworthiness, perspectives that emphasize enterprise governance of IT and the maturity of organizational processes (Becker et al. 2011).

The following discussion is designed loosely along the Goal-Question-Metric paradigm (Basili et al. 1994). Each goal is associated with a set of questions corresponding to the objectives outlined in Becker et al (2014). The answers to these should support an assessment as to how far the goal has been achieved. To this end, each question is further linked to a set of metrics that provide objective indicators to support an answer to the question. We discuss each of the design goals in turn and discuss the specific questions that need to be answered to provide an assessment of how the state of art is improved with the proposed system design and implementation. This forms the basis of a systematic discussion, taking into account the quantitative indicators and the qualitative discussion of the state of art.

5.2 Evaluation of design goals

G1 Provide a mechanisms for scalable in-depth content profiling

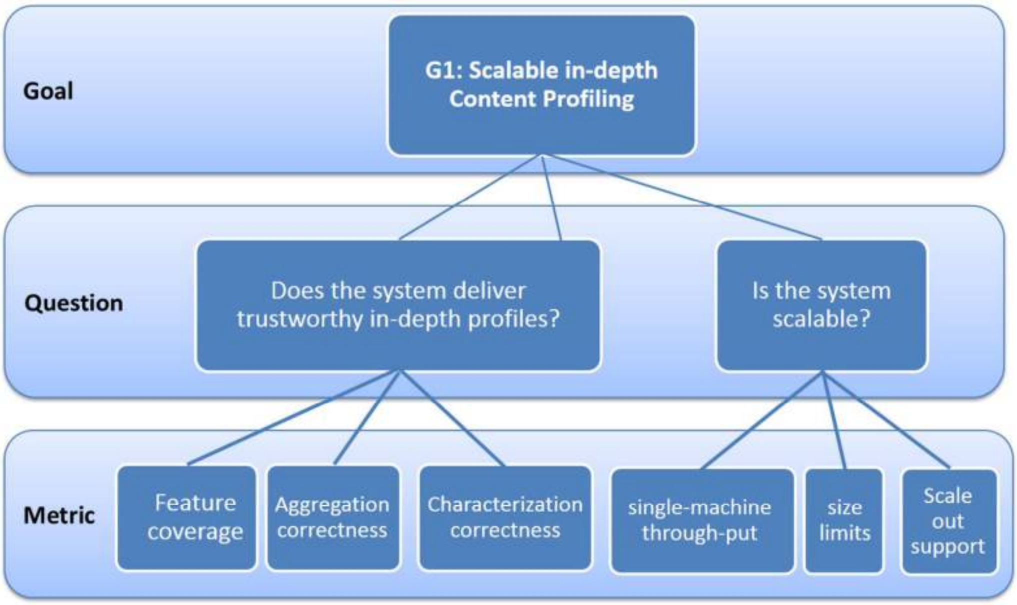

Goal
G1: Scalable in-depth Content Profiling
Question
Does the system deliver trustworthy in-depth profiles?
Is the system scalable?
Metric
Feature coverage
Aggregation correctness
Characterization correctness
single-machine through-put
size limits
Scale out support

Figure 14: Scalable profiling goals

Figure 14 poses the key questions we need to answer to evaluate the scalability and quality of in-depth profiling. Content profiles need to be meaningful, i.e. cover the interesting features that are known to be relevant for preservation, and trustworthy. Clearly, a profile covering only the size of files will be less meaningful than a profile including mime-types, formats, and validity. Additionally, a plethora of features influence the success in continued access, ranging from the presence of Digital Rights Management settings to the numbers of embedded tables in electronic documents and other dependencies.

The features needed for in-depth characterization process are broad coverage in terms of supported file formats and the features extracted, usage of a common vocabulary for identification of the formats, feature names and its values, and a reasonable low resource consumption and performance so they can be used in large-scale frameworks. By relying on the fits file information toolset, the C3PO profiler maximizes the coverage of features, arguably providing the highest possible feature coverage that can currently be achieved (Petrov & Becker 2012).

The correctness of aggregation itself can be verified in a straightforward way, since the operations are basic statistical calculations. The correctness of general map-reduce based operations in themselves can be assumed. On the other hand, the correctness of characterization components to provide accurate feature descriptors based on arbitrary input is far from proven. In fact, current data sets are entirely insufficient for proving correctness of the complex interpretation processes that take place. This, however, is a problem on the level of operations and cannot be attributed to the aggregation step of content profiling. Separate efforts are underway to verify the correctness of characterization tools using model-driven engineering to generate annotated test data (Becker & Duretec 2013). To enable future evolution, yet another aspect of scalability, a meaningful content profiler must be flexible enough to work on arbitrary property sets. C3PO supports this by relying on a generic data model, so that any additional property sets can be profiled.

This also supports the integration of further characterization tools. While FITS fulfils coverage and vocabulary requirements, it consumes considerable resources and takes a substantial amount of time to execute[i]. Hence, C3PO also supports Apache Tika, which in experiments showed far better resource consumption and performance with a throughput of up to 18 GB per minute[ii] when used with large-scale platforms such as Apache Hadoop[iii]. While in this case, Apache Tika was only used for file format identification, it supports feature extraction and has good coverage of file formats[iv], but does not yet use a well-defined vocabulary for the identification of extracted features.

The objectively measurable throughput and resource usage in profiling, then, is the crucial final question. To measure the time and resource behavior of C3PO, a set of controlled experiments has been conducted. The first measured the throughput of C3PO on a single standard machine, while the second employed a server with strong hardware and explored the boundaries of scalability by attempting to profile up to 400 million resources (12 Terabyte) in a single profile, enabling a further extrapolation of these results to the entire set of 300 TB in this collection. The third examined the limits of the web visualization platform to cope with these amounts of data.

The first experiment (Petrov & Becker 2012) tested the performance on limited resources and showed that on a standard computer with 4 GB of RAM and 2.3 GHz CPU, the ingesting and generation of a profile of 42 thousand FITS takes about 1.5 minutes.

Large scale tests were performed by Niels Bjarke Reimer from the Danish State and University Library[v]. A 12 TB sample was taken from a dataset with 300 TB of the Danish web archive. FITS was run on the sample content, resulting in 441 million FITS files. This characterization process took about a year to complete[vi].

For content profiling, two processing parts need to be considered: 1) the gathering of files into the internal data structure and 2) the analysis of that data set using mapreduce queries. The experiments were executed on a single machine with the specifications described in Table 1.

Processor	2 X Intel Xeon x5660 2.8 GHZ (12 core)
RAM	72 GB
Storage	Isilion storage box with 20 TB storage and 400Gb SSD, connected by a 1Gbit/s Ethernet network
Operating System	MongoDB Linux x86 64-bit v2.4
MongoDB	Version 2.4
Application service	Apache Tomcat version 7

Table 1: C3PO scalability test machine specifications

The first step, which ingests the FITS files into a MongoDB server, was tested with the 441 million FITS items. The graph depicted in Figure 15 shows the import time for samples of around 3.600 files. The Y-axis unit shows time in milliseconds and the X-axis unit is a sample number, which can also be considered as a timeline.

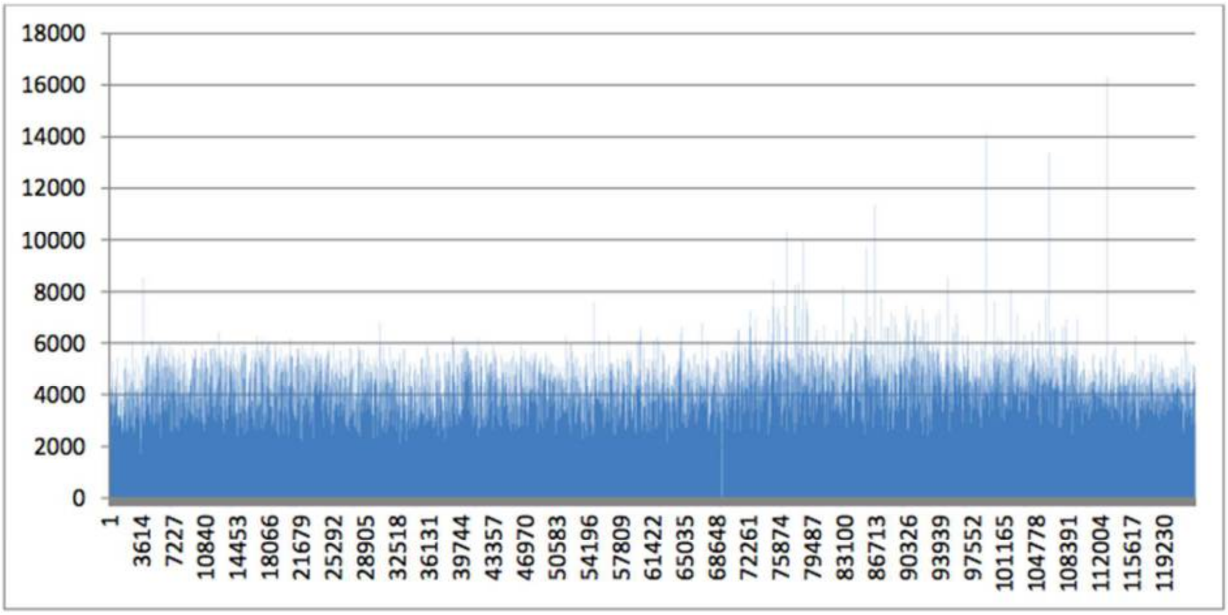


Figure 15: Performance of C3PO import process using FITS metadata (Reimer, et al. 2013)

The complete import process took less than 80 hours, with an average execution time of 0.65 milliseconds per FITS file. This import time is quite constant with only a few outliers, which implies that the platform and the software are acceptable for importing large amounts of data.

The second step, the analysis of the data using map reduce queries, was tested with a data set of about 12 million FITS files and took 15 hours and 18 minutes, which is about 4.63 milliseconds per FITS file. Using sharding and mapreduce technologies, the processing time of the second step should also be linear.

In conclusions, both steps are linear and together take on average 5.28 milliseconds per FITS file. This means that processing the current 300TB dataset would take about 16170 hours, or about 677 days, on a single machine. As both processes are massively parallelizable and the MongoDB platform already supports sharding and map reduce, the processing time can be highly reduced by distributing the load on several servers. Substantial resources may be needed to bring the processing time down to a practical time, but this profile does not have be re-generated frequently.

The C3PO tool also provides a web interface that supports real-time analytics on the gathered data. While this is not one of the requirements strictly required for automated monitoring, it provides interesting insights into content profiles that is considered highly valuable by the decision makers. However, in this scenario the limits of the web application are revealed. Several test runs were made with different data set sizes to ascertain the limits of the application. For each test run, two manual procedures on the web interface were made: 1) opening the overview page that calculates, in real time, distributions of several extracted features, and 2) drilling down into into the characteristics of a subset of the collection, as all the files of a defined format.

Test run	# FITS files	Elements size GB	number of properties	Overview processing time	Drill down processing time
1	13,962	0.03	80	Fast	Fast
2	108,348	0.26	96	18 sec	11 sec
3	363,991	1.00	106	30 sec	34 sec
4	1,020,514	2.46	113	2 min 25 sec	1 min 42 sec
5	1,639,842	3.95	119	3 min 52 sec	2 min 50 sec
6	2,683,596	6.44	119	6 min 28 sec	4 min 25 sec
7	11,905,935	28.63	211	not finished within 3 hours	N/A
8	441,923,560	1183.50	5122	N/A	N/A

Table 2 - Testing the limits of real-time analytics in the C3PO web application (Reimer, et al. 2013)

Table 2 shows the results of the tests, which show acceptable results up to around 2.5 million files, with a waiting time of around 6.5 minutes. Above this limit, the web system does not respond within 3 hours, which is considered unacceptable. Hence, it is not feasible to perform real-time analytics with the current solution on the set of 440 million FITS files of the 12TB data set. It should be noted that the analysis of the entire dataset provides a view of over 400 million rows in a table with over 5000 columns, with a resulting database size of over a Terabyte.

G2 Enable monitoring of operational compliance, risks and opportunities

This section analyzes how the mechanisms presented in Becker et al (2014) can be used to accomplish proactive monitoring of operational compliance, risks and opportunities in a preservation environment.

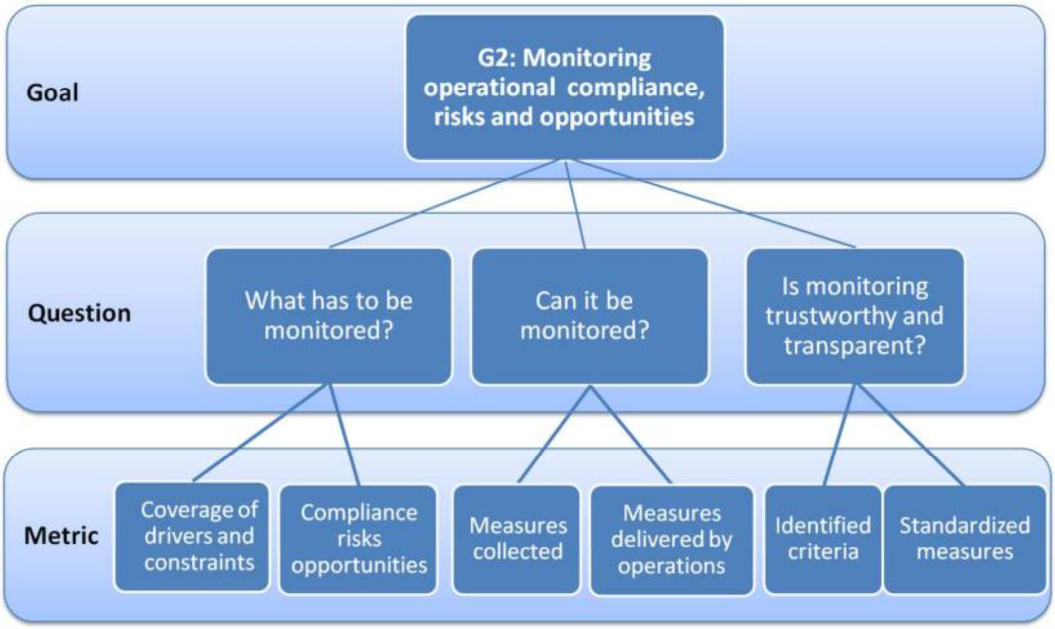

Goal
G2: Monitoring operational compliance, risks and opportunities
Question
What has to be monitored?
Can it be monitored?
Is monitoring trustworthy and transparent?
Metric
Coverage of drivers and constraints
Compliance risks opportunities
Measures collected
Measures delivered by operations
Identified criteria
Standardized measures

Figure 16: Monitoring

As outlined in Figure 16, the key questions relate to the identification of aspects that need to be monitored and to the coverage of measures available to provide indicators related to these aspects. Based on an analysis of a reference model for drivers and constraints (Antunes et al. 2011), which classifies each of the influencers a preservation organization should be aware of, the discussion in (Becker, Duretec, et al. 2012) showed that relevant questions and measures can be derived for each of the influencers of interest. This enables the development of appropriate adaptors for measuring specific indicators pertaining to this driver. Table 3 shows key examples, while a full discussion and detailed table is provided in (Becker, Duretec, et al. 2012).

Driver	Question	Indicator	Sources
Content	Is the content volume growing unexpectedly?	Rate of growth changes dramatically in Ingest	content profile, Repository Report API
Operations	Are our content profiles policy-compliant?	Mismatch between content profiles and policy statements	content profiles, control policy statements
Format	How many organizations have content in this format?	number of shared content profiles containing a format	content profiles shared by organizations
Format	What is the predicted lifespan of format X?	lifespan estimates based on historic profiles	model-based simulation

Table 3: Selected preservation drivers and related information sources (Becker et al, 2012)

In practice, the achieved coverage of measures is by no means complete, but increasing. Currently supported sources include format registries, semantic policies, content profiles, and an automated rendering and comparison tool (Law et al. 2012). A prioritization approach is taken to target first and foremost those aspects that are perceived most critical. The open nature of the adaptor design, the data model, and the licensing model has the effect that additional sources can be integrated by anybody in the preservation community, and the coverage is rising steadily.

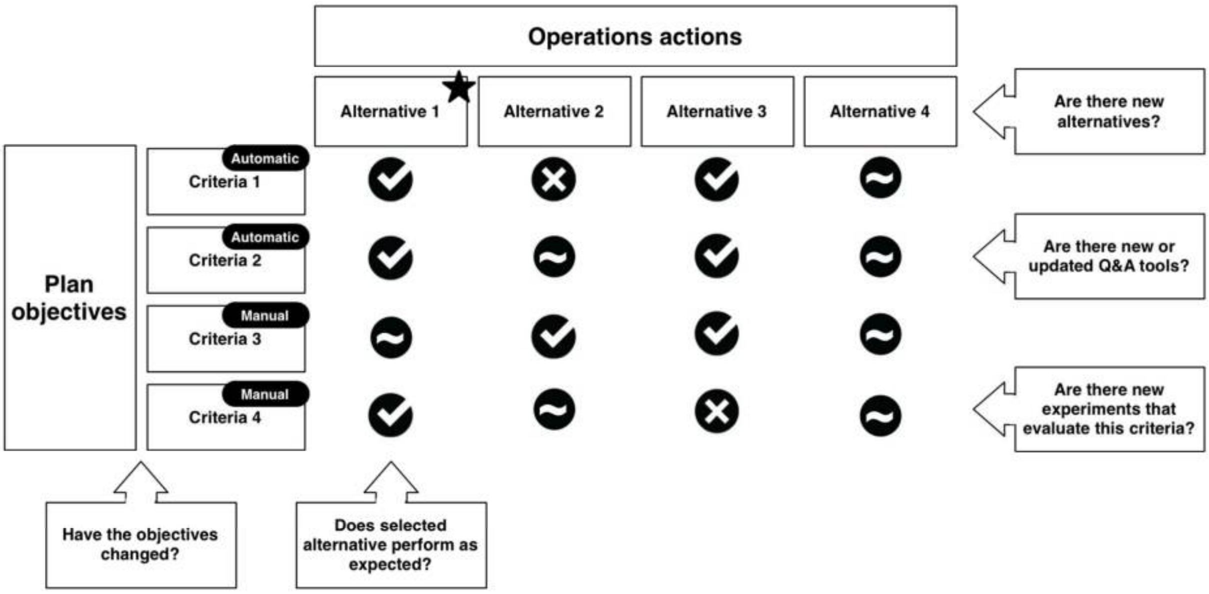

Operations actions
Alternative 1
Alternative 2
Alternative 3
Alternative 4
Are there new alternatives?
Plan objectives
Automatic
Criteria 1
Automatic
Criteria 2
Manual
Criteria 3
Manual
Criteria 4
Are there new or updated Q&A tools?
Are there new experiments that evaluate this criteria?
Have the objectives changed?
Does selected alternative perform as expected?

Figure 17: Evaluation of monitoring compliance

Monitoring of operational plans is illustrated schematically in Figure 17 on a simplified example. Consider a preservation plan that evaluates four potential actions ("alternatives") against a set of four decision criteria. These criteria evaluate the important aspects of the data to be preserved, the environment, and the actions to be applied. Based on these criteria, preservation actions in question are evaluated and a ranking is calculated. The planner then chooses the best suited action and adopts it. In this example, a check mark denotes a best-in-class performance, a tilde denotes acceptable performance, and a cross reflects an unacceptable performance, for example a process that did not terminate or an image conversion that shows a distorted image. We can see that two alternatives have been rejected, and alternative 1 has the highest score and will be selected.[vii]

Since the decision criteria identified during planning lead to the adoption of a certain action, they must be monitored during operational executions as well to enable the organization to track whether the action keeps performing according to the expectations. This is shown on the bottom left of the figure. However, this is not the only aspect prone to evolution:

(1) New alternatives will emerge over time that may perform better than the chosen alternative. In cases where no alternative was acceptable, this will sometimes be the only thing monitored,

since the organization would wait for a better solution to become available before embarking on premature preservation actions. For example, this was the case in (Kulovits et al. 2009).

(2) Updated or new Quality Assurance tools can emerge that provide more reliable or more efficient measures for Quality Assurance or even the first automated way to measure relevant quality. For example, these could be of the kind described in (Jurik & Nielsen 2012) or (Bauer & Becker 2011).

(3) Related to this, experiments including certain criteria may be conducted by other individuals or organizations that can reveal risks and opportunities related to this plan. For example, the chosen Quality Assurance tool might be shown to malfunction on similar objects, which poses a major risk (Bauer & Becker 2011).

(4) Finally, the organization's objectives themselves may shift over time as goals change. This would be reflected by a change in the control policies.

The tool suite described in this article is designed to provide full support for this monitoring scenario. The upcoming release of Plato generates specifications describing the expected quality of service (QoS), similar to a service-level agreement (SLA), for the set of decision criteria considered, linked to the corresponding organizational policies, and deposits corresponding monitoring conditions in Scout upon deployment of a preservation plan. Such QoS specifications are created for those criteria in a tree which are influenced by the dynamic behavior of the service - i.e. the components. That means that they are not created for aspects relating to the format, such as the ISO standardization of PDF versions, but they include criteria such as whether the created files are well-formed. QoS is then measured within executable workflows and monitored for fulfillment. Aspects pertaining to the format and other non-dynamic aspects are monitored as risks and opportunities using Scout.

While Scout is able to collect a wide variety of measures, these are naturally limited by the availability of operations that support such measures. The controlled vocabulary encourages developers to declare which measures their tools deliver to support discovery, but the coverage of measures will naturally vary across different scenarios. It is important to note, however, that any required measures can be integrated by any organization due to the open nature of the ecosystem.

Finally, transparency of the monitoring process is achieved through the usage of the permanent shared vocabulary and the explicit declaration of tolerance levels in the QoS, corresponding to the specified acceptance thresholds that are derived from the organization's control policies.

G3 Improve planning efficiency

Previous work has shown that the key challenge in planning is to make the decision making process more efficient (Kulovits et al. 2009; Becker & Rauber 2011c). In **Becker et al (2014)**, we reflected on the dimension of trust that should not be sacrificed along this quest. Correspondingly, the key questions shown in Figure 18 relate to the aspect of effort: How long does it take to create one preservation plan now, and how much further improvement is possible?

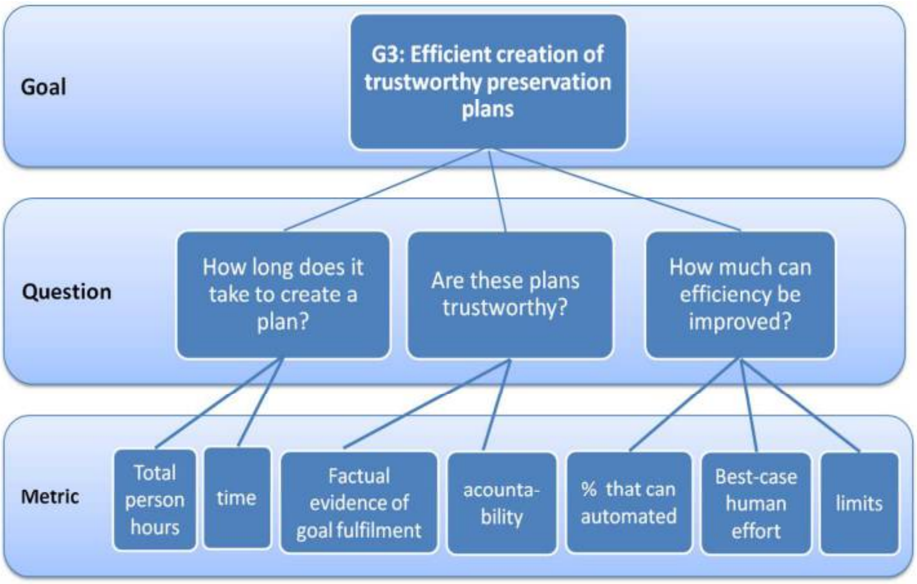

Goal
G3: Efficient creation of trustworthy preservation plans
Question
How long does it take to create a plan?
Are these plans trustworthy?
How much can efficiency be improved?
Metric
Total person hours
time
Factual evidence of goal fulfilment
acounta-bility
% that can automated
Best-case human effort
limits

Figure 18: Efficient creation of trustworthy plans

Previous discussions have shown the trustworthiness of plans produced by Plato (Becker et al. 2009), which is based on the evidence-based measures of decision criteria directly linked to organizational goals and based on factual evidence, documented with full change tracking assigned to acting users. These strengths continue to form the backbone of trustworthy planning. While it is clear that fully automated, i.e. *autonomous*, preservation planning is contradicting the goal of trustworthiness in this domain, the goal nevertheless must be to achieve a substantial increase in efficiency (Becker & Rauber 2011c).

We will focus our discussion on measurements of effort on a controlled case study conducted with the Danish State and University Library, described in detail in (Kulovits et al. 2013). In this study, a set of responsible decision makers and experts from the library set out to create a preservation plan, with the assistance of a planning expert and a moderator who kept time of all activities throughout the planning process. The goal of planning was to create a preservation plan for a large set of audio recordings; the drivers and events motivating the plan included the goal to homogenize the formats of the library's holdings and provide well-supported and efficient access to authentic content. The team at the library has comprehensive expertise in all relevant areas, which range from technical knowledge on audio formats and quality assurance mechanisms for comparing audio files to a documented understanding of the designated communities and preservation purpose of the content set at hand.

The preservation plan was created using the then-current version 3 of the planning tool Plato, which is the precursor of the solution presented here. The goal was to identify the major areas of decision making effort and measure the potential improvement that can be realistically achieved.

The total time required to create a preservation plan amounted to 35.5 person hours, completed over a period of two days. This shows on the one hand that efficient teams in well-established settings can already plan quite efficiently. Nevertheless, the effort must be further reduced to make planning truly a part of "business-as-usual" preservation in practice. To contextualize the effort required in this case, it is important to understand that this effort strongly depends on a well-defined understanding of the decision making context, including the understanding of the goals and constraints; the expertise of decision makers; and the technical proficiency of the staff carrying out the experimental steps of preservation planning. Finally, a strong driver for cost is the homogeneity of content: For large object sets that are very diverse, several preservation plans will have to be created, each respecting to a certain degree the specific aspects of a subset of the content and the means available to ensure access to this subset.

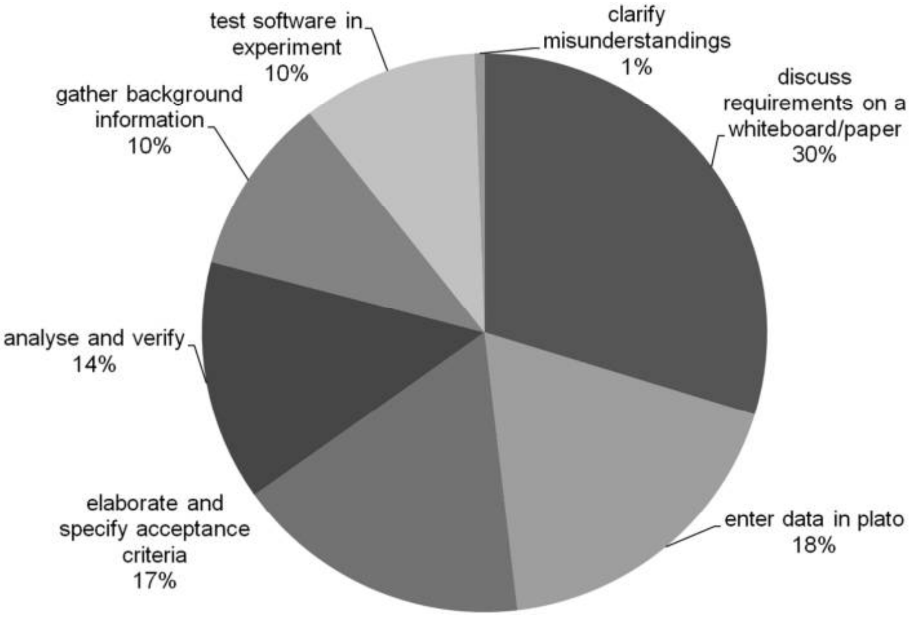


Figure 19: Distribution of effort across activities in preservation planning (Kulovits et al, 2013)

Figure 19 shows the distribution of effort across each of the types of activities that were part of this planning process. It should be noted that several of these activities were in fact on the upper end of the efficiency range for several reasons:

- Experiment execution often takes more time. The experiments conducted were highly efficient due to the minimum number of alternatives evaluated, the high technical proficiency of staff, the homogeneity of content, and the quality assurance mechanisms employed. In many cases, the experimentation processes consumes a multiple of the time. The integration of Taverna workflows and myExperiment can reduce this massively, since potential components can be discovered and automatically invoked within planning. This automation is similar to an existing integration of automated measures in Plato (Becker & Rauber 2011a), but makes these mechanisms available on an open, standardized and easily extensible basis.

- Background information often is unavailable. This applies in particular to the user communities and the statements of *preservation intent* that many organizations are only now beginning to document systematically (Webb et al. 2013). The organization in question, however, has a stable and well-supported definition of collections and user communities, from which the preservation goals could be derived rather efficiently. Formal policy specification makes this background explicit and known by the systems, so that the effort can be further reduced.

- Analysis and verification is complex. Even with the support of a planning expert, 14% of the time was spent in sense-making, analyzing the completed set of evidence and assessment in the decision making tool to arrive at a conclusion that was well understood by the stakeholders. This points to the need for improving the decision support tool in visualizing results in a more easily understandable and user-friendly way. Improved summaries in Plato are planned to this end.
- Entering data into the system is tedious, in particular for users not familiar with the tools. This was alleviated by the involvement of a planning expert familiar with the tool. Similar to other aspects, this benefits greatly from the integration of the tool with workflows and from the explicit endowment of Plato with an understanding for the policy models of organizations. A subsequent controlled experiment showed that Plato 4 reduced this effort by over 50% (Kulovits et al. 2013).

In an ideal case, the effort required to cover the above aspects (software testing, background information, analysis and verification, and data input) can be removed almost entirely. Still, 50% of the time in this case would be spent discussing requirements. However, the majority of these concerns objectives about formats, significant properties, and technical encoding or representation (Becker & Rauber 2011a). For all of these aspects, standard definitions are now available as part of the controlled vocabulary that enable decision makers to reuse definitions and formalize these aspects on a policy level, removing this activity from the operational planning process. This applies to the designated community and preservation intent statement as well as to format and risk factors and technical constraints. The control policy statements thus can improve his effort by enabling reuse of these goals and constraints across plans. As the discussion on Goal 4 will show, the context awareness of Plato can eliminate the need for in-depth discussions of requirements *as part of planning* almost entirely.

For an organization that establishes planning as a proper function in its roles and responsibilities and possesses a solid skills and expertise base, we estimate that preservation planning should on average take about one to two person days per plan, provided that policies and content profiles are known and documented. However, a large variance across organizations is to be expected. This estimate will strongly depend on a variety of specific factors and certainly needs to be further validated in longer-term empirical studies. These should in particular also cover the question of homogeneity of content sets covered in a plan: How many plans are required to safeguard a particular heterogeneous set of objects? A detailed discussion on the activities in this process and the relevant skills and expertise is presented in (Kulovits et al. 2012).

G4 Make systems aware of their context

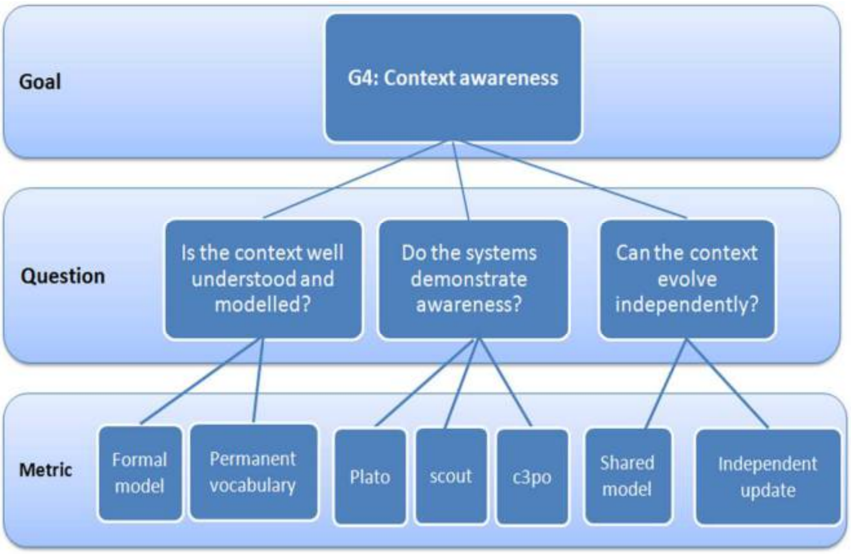

Goal
G4: Context awareness
Question
Is the context well understood and modelled?
Do the systems demonstrate awareness?
Can the context evolve independently?
Metric
Formal model
Permanent vocabulary
Plato
scout
c3po
Shared model
Independent update

Figure 20: Context awareness

By providing a model that enables decision makers to formulate policies so they can be understood by automated processes, the systems can understand their context and stay informed about its state. To assess the context awareness of the systems in question, we investigate three distinct aspects: On the one hand, the context needs to be well understood and modeled in order to ensure a solid approach has been taken. On the other hand, each of the systems needs to demonstrate that it can use the part of the context that is relevant for its function appropriately. Finally, it is crucial to ensure that this does not come at the cost of coupling the context too closely with the systems. To this end, we discuss how this context can evolve independently from each of the systems. This is illustrated in Figure 20.

The modular approach of the semantic models has been discussed in Becker et al (2014). A detailed documentation of the model is provided in (Kulovits et al. 2013). The models are based on W3C-approved standards and follow established Linked Data principles. At the heart of the model is the Resource Description Framework[viii] (RDF), a standard model for enabling the representation of data and metadata in subject-object-predicate triples. The Web Ontology Language[ix] (OWL) provides the mechanisms for the description of vocabularies, defining classes and properties. These are used to annotate, describe and define resources. Having well-defined semantics, OWL facilitates reasoning, ontology management and querying of data. The model is represented as an RDF graph and queried using SPARQL. The vocabulary domains have permanent identifiers according to the following ontologies:

- http://purl.org/DP/preservation-case contains the basic elements that link a preservation case together.
- http://purl.org/DP/quality describes the quality ontology, linking attributes and measures in a domain-specific quality model.

- http://purl.org/DP/quality/measures contains the vocabulary individuals that are used for annotating, describing and discovering measures and the mechanisms for measuring.
- http://purl.org/DP/control-policy, finally, defines the classes of objectives relevant for making a preservation case operational.

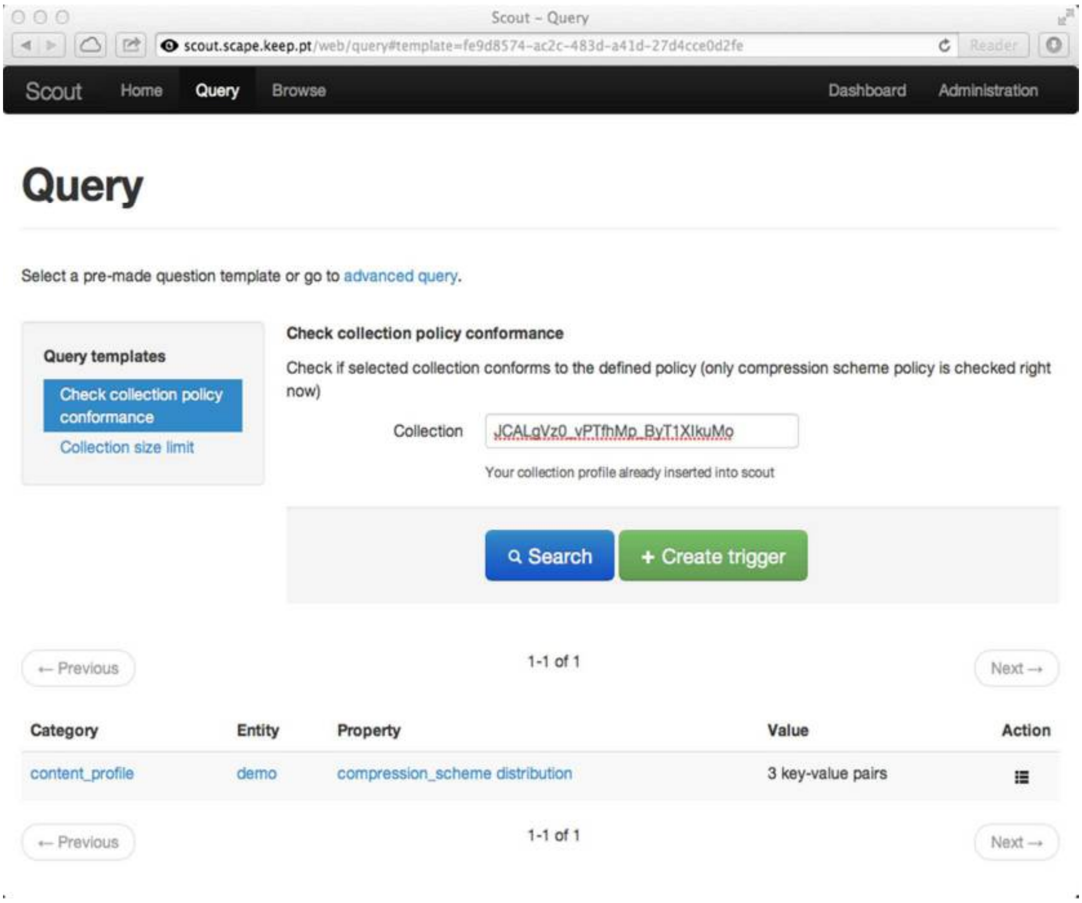
Each of the systems presented is aware of those parts of the model that are relevant for its domain. Correspondingly, each system shows its awareness of this model in different manners.

Plato uses the control policy model in several ways. On the one hand, the preservation case provides the basic cornerstones of planning. Instead of providing the documentation of the planning context in textual form, as it used to be standard (Becker & Rauber 2011c), a planner who has specified the policy model selects a preservation case to start planning, and the contextual information from this case is extracted from the policy model. Additionally, the objectives and measures specified in the control policy enable the decision support tool to derive the complete goal hierarchy automatically from the model, leaving it to the decision maker only to revise, verify and confirm the decision criteria to be used for the experimental evaluation. In the case study discussed above, this requirements specification alone accounted for 30% of the effort. While the policies of course require a similar discussion, much of the objective specification has to be discussed only once and can then be carried forward across preservation cases, which represents a substantial efficiency gain as soon as more than one plan is created. Similarly, the acceptable values, and hence the utility functions associated with each measure, can be computed in a straightforward way based on the objectives specified in the control policies, which presented a potential gain of another 17% in our case study discussed above.

C3PO uses the vocabulary relevant to characterization in the content profile, referencing elements from the quality measures catalogue (such as http://purl.org/DP/quality/measures#55). Given that it only provides objective analytics of factual statements about the domain elements, it has no understanding of the policy model and does not require any.

Scout, finally, leverages the policy model for monitoring the alignment of operations and plans to the policies, and also monitors the policy itself: If it is updated, that means that affected plans should be re-evaluated. Specific standard queries are provided as templates that monitor policy compliance. These can be activated by the user. For example, Figure 21 shows Scout starting a monitoring activity on the policy conformance of a specific content set (identified by a collection key). In this case, it shows in a preview that the property *compression scheme* is violated by 3 entries, and provides the option to create a continuous monitoring process by specifying a trigger with a condition and an event.

It can be seen that the model of the context is shared between the tools, with the decision maker updating the ontology independently of the tools. A crucial requirement is that the context model can evolve independently of the systems. This is especially important considering that the current model is very much focused on operational support and can benefit greatly from being expanded to cover aspects of decision making that are further removed from operations. Similarly, it can be expected that meaningful linkages will surface that connect the existing ontologies to emerging ontologies from neighboring areas ranging from software quality and ontologies for describing software dependencies and platforms to preservation metadata and related policies. The potential for such evolution is guaranteed by the choice of representation and languages, since the Linked Data principles that the model adheres to are designed with these very goals in mind.


Scout – Query
scout.scape.keep.pt/web/query#template=fe9d8574-ac2c-483d-a41d-27d4cce0d2fe
Reader
Scout
Home
Query
Browse
Dashboard
Administration

Query

Select a pre-made question template or go to advanced query.

Query templates

- Check collection policy conformance
- Collection size limit

Check collection policy conformance

Check if selected collection conforms to the defined policy (only compression scheme policy is checked right now)

Collection: JCALgVz0_vPTfhMp_ByT1XIkuMo

Your collection profile already inserted into scout

1-1 of 1

Category	Entity	Property	Value	Action
content_profile	demo	compression_scheme distribution	3 key-value pairs	☰

1-1 of 1

Figure 21: Checking collection policy conformance in Scout

G5 Design for loosely-coupled preservation ecosystems

The design goal of loosely-coupled systems is relevant for several reasons. On the one hand, it is crucial to enable the stepwise adoption approach preferred by many organizations (Sinclair et al. 2009). On the other hand, it ensures that evolution can take place independently, enabling each organization to replace parts of its system without negatively affecting continued operations, and enables each component of the ecosystem to be sustained independently (to a degree) of the others.

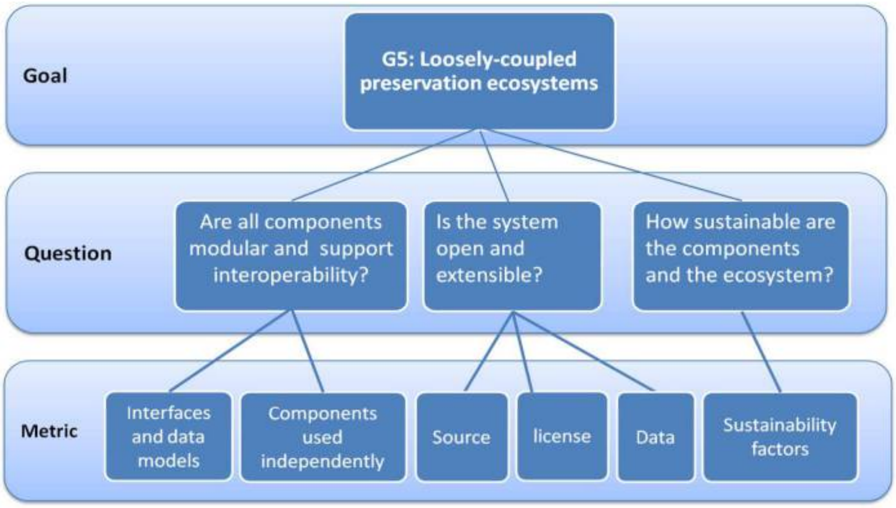

Goal
G5: Loosely-coupled preservation ecosystems
Question
Are all components modular and support interoperability?
Is the system open and extensible?
How sustainable are the components and the ecosystem?
Metric
Interfaces and data models
Components used independently
Source
license
Data
Sustainability factors

Figure 22: Loosely-coupled preservation ecosystems

Figure 22 relates these goals to more specific questions. While it is clear that the components are open source, licensed under OSI-approved conditions[x], and highly modular, it is useful to consider closely both the functional specifications and the data structures. The API specifications for the SCAPE Planning and Watch suite are in the process of being published openly on github. Data exchanged between components is standardized and supported by schemas, as shown in Table 4.

	Plato	Scout	C3PO
All functional interfaces openly published	In progress	In Progress	In progress
All data structures documented using standards and schemas	XML schemas published for each version	Linked Data model Policy model	XML schema published
Component is used independently	Yes	Yes	Yes
Component follows the controlled vocabulary	objectives, measures, control policies, preservation cases	objectives, measures, control policies	measures

Table 4: Interoperability of components

The controlled vocabulary as the glue that connects much of the ecosystem is maintained on github[xi]. Curating this vocabulary over the long term will be sustained by a community effort. Recent discussions in the communities of metadata and preservation have brought forward long-term requirements for such evolution that will be considered carefully. (Gallagher 2013a; Gallagher 2013b)

The components are functionally independent in that every component can and is actually used independently. Nevertheless, it is clear that the compound value proposition is larger than the sum of its parts, serving to encourage take-up of the suite as a whole. Similarly, the usage of this tool suite benefits greatly from integrating also with the workflow development, execution and sharing platforms Taverna and myExperiment, whose latest releases provide specific support for semantic annotation, driven by the requirements outlined in this article.

Since such an ecosystem should be built with sustainable evolution in mind, we consider a recent discussion that identified eleven factors affecting the sustainability of a modular preservation system (Gallagher 2013a; Gallagher 2013b). Table 5 shows how our system performs on each of these criteria.

Sustainability factors[xii]	How does the SCAPE Planning and Watch suite perform?
Ability to view and modify source code	All components are openly licensed, and all source code elements are freely available on a github repository.
Widely used	C3PO and Scout are relatively new, but enjoying quick take-up in the community, while Plato has been growing to over 1000 user accounts since first publication 2008. However, usage so far has been limited to prototypical evaluation rather than production-level deployment, mostly due to the level of effort involved.
Well tested, few bugs or security flaws	All tools support automated tests and have an active ticketing system, and the major releases are considered very stable. No security incident has been reported so far.
Actively developed, supported	All tools are part of an active development community, continuously supported, and the development platform is hosted by the Open Planets Foundation[xiii].
Standards aware	All components follow standards on multiple levels wherever possible. This ranges from standard technologies such as Java Server Faces to XML Schema declarations and Linked Data principles.
Well documented	All components have extensive code documentation, manuals, built-in help and tutorials, as well as scientific publications explaining the theoretical foundations and practical implications of the software.
Unrestricted licensing	All software components are licensed under OSI-approved open licenses such as LGPL and Apache Software License 2.0. All documentation is licensed under the Creative Commons license.
Ability to import and export data and code	Preservation plans, executable plans and content profiles can be freely imported and exported, and shared between users. The Scout knowledge base is a Linked Data triple store and hence equally portable.
Compatible with multiple platforms	Being based on standard server technologies, all components are compatible with multiple platforms. Plato even integrates with multiple platforms at once in the case of preservation action discovery (Kraxner et al. 2013).
Backward compatible	This is very relevant in the context of Plato, which is an online service since 2008. Here, there is full backward compatibility with a fully traceable forward conversion upon import of legacy preservation plans. All plans created on the online service have been automatically migrated for all releases. Similarly, the knowledge base of Scout is designed to keep growing incrementally,

	without disposing of accumulated historical data.
Minimal customization	There is almost no customization required, since all contextual adaptation of the systems' behavior can be achieved through the configuration of API endpoints and the corresponding definition of control policies.

Table 5: Sustainability evaluation of the SCAPE Planning and Watch suite

While the ecosystem is well positioned for future sustainability, there is still room for improvements. This includes the development and publication of Technical Compatibility Kits that can automatically test the functional compliance of a component to an API specification, as it has been done for the Data Connector API[xiv], but also the long-term evolution of vocabularies and any future extensions of the tool suite.

5.3 Practical adoption

Considering the preservation lifecycle outlined in Becker et al (2014), what does the availability of the described system mean for an organization that has content and a preservation mandate, has set up a reasonable organizational structure and defined corresponding responsibilities, but has not yet ventured to create and maintain specific, actionable preservation plans? The exact measures to be taken will certainly depend on the specific institutional context, but essentially, such an organization can follow a series of steps.

1. *Getting started entails several aspects.*

    a. *Start content profiling*. Run format identification and characterization components such as fits on the set of content to extract metadata, deploy the content profiling tool C3PO, gather the metadata, and conduct an analysis of the content profile.

    b. *Sign up with SCAPE Planning and Watch*, either on the online service[xv] or on an organization-specific deployment based on a public code release[xvi].

    c. *Connect the organization's repository* with SCAPE Planning and Watch, either through configuring a standard adaptor or implementing a specific adaptor.

2. *Specify control policies* based on a thorough analysis of the organization's collection, the user communities, and the preservation cases that are considered relevant.

3. *Activate the monitoring* of policies and content profiles in Scout to detect policy violations.

4. *Create preservation plans* to increase the alignment of the organization's content and operations to the goals as declared in the policies. This planning will be done by evaluating action components using characterization and QA components in Taverna workflows, all integrated in planning. The finished Plans contain workflow specification including QoS criteria that can be automatically monitored.

5. *Deploy the operational plans* to the repository through the plan management API, connected to a workflow engine such as Taverna.

6. Establish responsibility for continuous monitoring. This is supported by Scout, which will monitor the compliance of operations to plans and detect risks and opportunities connected to these plans and policies.

5.4 Limitations

From the discussion above, a number of limitations can be observed. These can be divided into limitations of the current capabilities of available tools, which can be expected to grow, and more fundamental limitations of current approaches which require new perspectives to be overcome; limitations of the problem space that set natural limits to further improvement; and limitations on the quantitative evaluation that can feasibly and meaningfully be conducted. This section discusses those limitations that are seen as central to the further advancement of the state of art.

Coverage and correctness of available measurement techniques

The availability of tools and mechanisms to deliver objective and well-defined measures that are shown to be correct and reliable is a key challenge holding back operational preservation today (Becker & Rauber 2011c; Becker & Duretec 2013). Scout supports a growing set of adaptors to feed in measures into the knowledge base, and by nature of the design alleviates some of the shortcomings and gaps in existing tools through the free combination of multiple information sources, but still is limited by the availability of these information sources. Similarly, experiment automation in Plato and, equally important, the feasibility of large-scale preservation operations in general, is entirely dependent on the existence of well-tested, efficient and effective mechanisms for Quality Assurance. Recent work is showing promising advances (Jurik & Nielsen 2012; Bauer & Becker 2011; Law et al. 2012), but there is still a wide gap to be addressed for preservation operations to be broadly supported. It seems crucial that this gap is made explicit and shared with a wide community so that efforts to close it can be based on a solid assessment of the shortcomings of existing tools rather than isolated ad-hoc identification of application scenarios within single institutions, as is often practiced today. Scalable preservation operations are only possible with fully automated, reliable and trustworthy Quality Assurance; and such quality assurance is expensive to develop and difficult to verify. Only through coordinated community efforts based on solid experimentation can the evidence be constructed to make a convincing case on authenticity (Bauer & Becker 2011).

The utter lack of solid, reliable and open benchmark data set with full ground truth is a fundamental inhibitor to validating the correctness of such measures. To address this gap, we are investigating innovative approaches to turn around the publication of test data sets from ex-post annotation, inherently plagued by unreliable ground truth and copyright problems, to an open, model-driven generative approach (Becker & Duretec 2013).

Scalable distributed and cost-efficient processing: How to profile a Petabyte?

As shown above, the content profiling tool C3PO provides support for scaling out on distributed platforms. However, it requires considerable resources if the content to be profiled is approaching the Petabyte range, and visual analytics are not currently supported on such amounts of data. Yet, it is important to point out that the core goals of content profiling are achieved regardless of the collection size: Visual analytics is an additional capability on top of the processing activity.

To enable cost-efficient creation of large content profiles without visual analytics requirements, we are exploring purely sequential profilers with a small footprint as a low-cost alternative, and we are investigating a set of techniques for feature-space pruning and dimensionality reduction prior to the more expensive processing steps.

Similar considerations apply to preservation operations such as actions, characterization, and quality assurance. As noted, the execution of *fits* on the 440M resources on the Danish web archive took a year to complete, which clearly indicates the need for improvements. Similarly, automated QA mechanisms are computationally demanding (Bauer & Becker 2011). These processes need to be supported by parallel execution environments and more efficient algorithms to be truly applicable on large-scale volumes.

The element of human decision making

As observed above, trustworthy preservation should always be driven by careful decision making and factual evidence. While this element of human decision making can be reasonably minimized, replacing it entirely will only be possible once a solid, substantial knowledge base of real-world cases populates the ecosystem described above. Eventually, the human element can in the ideal case be reduced to a policy specification activity and a monitoring oversight function. This is clearly out of scope for this article, but will provide the logical next step in research on preservation planning and monitoring.

Trust and maturity

The assessment of complex socio-technical systems such as the one presented is challenging. Arguably, it will not be complete without an enterprise governance view incorporating a set of dimensions on the level of organizational process performance and maturity. A first view on this perspective has been presented in (Becker et al. 2011), where a process and maturity model for preservation planning was outlined that was aligned with the IT Governance framework COBIT (IT Governance Institute 2007). Current efforts are building on this work to develop a full-fledged process and capability maturity model that shall support organizations in systematic improvement of their preservation capabilities.[xvii]

5.5 Summary

This section discussed each of the key design goals of the architecture and system presented in Becker et al (2014) and conducted a quantitative and qualitative evaluation of the key objectives for each of the goals. We showed that the system significantly improves on the existing state of art in digital preservation by combining a context-aware business intelligence support tool with a scalable mechanism for content profiling, both integrated with a successor of the standard preservation planning tool Plato that is showing substantial efficiency gains over previous solutions. While there are limitations on the scale of content that can be profiled, analyzed and preserved in limited amounts of time, the improvements show that preservation planning and monitoring can be realistically advanced to a continuous preservation management function integrated with operational systems. This will provide a substantial step forward for the many organizations that are looking for ways to enable their repositories for truly supporting the long-term access promise that digital preservation has set out to deliver (Hedstrom 1998). We pointed out a number of limitations that currently hold back further progress, and outlined current efforts to tackle them.

3. Conclusion and Outlook

Ensuring the longevity of digital assets across time and changing social and technical environments requires continuous actions. The volumes of today’s digital assets make effective business intelligence and decision support mechanisms crucial in this enterprise. While purely technical scalability of data processing can be handled using state of the art technologies, curators require specific decision support to enable large-scale management of digital assets over time. This demands a set of systems and services that facilitate scalable in-depth content analysis, intelligent information gathering, and efficient decision support, designed as loosely-coupled systems that are able to interact and connect to the wider preservation context.

This article presented a systematic assessment and evaluation of the SCAPE Planning and Watch suite presented in **Becker et al (2014)**. The results of the assessment demonstrate the possibility to deploy full preservation lifecycle support into preservation systems of real-world scale by adopting a loosely-coupled, open and extensible suite of preservation tools that each support particular aspects of the core preservation planning and monitoring capabilities:

1. **Scalable content profiling** is supported by the highly flexible and efficient content profiler C3PO, which has been tested on a data set of 441 million files.

2. **Monitoring of compliance, risks and opportunities** is supported by the monitoring system Scout, which provides an extensible open platform for drawing together information from a variety of sources to support the much-needed business intelligence insights that are key to continued preservation success.

3. **Preservation planning efficiency** is being continuously improved as the ecosystem grows, and recent advances show that planning can become a well-understood and managed activity of repositories.

4. **Context awareness** of each of the systems is supported by a shared permanent vocabulary set to grow over time through extensions with related ontologies, connecting the domains of

solution components and the preservation community with the organizational policies and the decision support and control systems presented here.

5. **Loose coupling of the components** in this ecosystem guarantees that organizations can follow an incremental approach to improving their preservation systems and capabilities.

We discussed the evaluation of key aspects of each tool as well as the ecosystem as a whole and outlined the key benefits and advances over the existing state of art. Based on a number of limitations, we define a number of key goals for future research. These include real-time profiling of very large data sets in the Petabyte range; benchmarking of automated tools against solid, reliable ground truth in open, fully transparent experiments with shared data sets; and a systematic framework for assessing the performance of organizations in terms of process metrics and organizational maturity.

References

Antunes, G. and Borbinha, J. and Barateiro, J. and Becker, C. and Proenca, D. and Vieira, R. (2011), "Shaman reference architecture", version 3.0. SHAMAN project report.

Basili, V.R. and Caldiera, G. and Rombach, H.D. (1994), "The Goal Question Metric Approach", Encyclopedia of Software Engineering, Volume 2, John Wiley, pp.528–532.

Bauer, S. and Becker, C. (2011), "Automated Preservation: The Case of Digital Raw Photographs" , in Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation Proceedings of 13th International Conference on Asia-Pacific Digital Libraries (ICADL 2011) in Beijing, China, 2011, Springer-Verlag.

Becker, C. and Antunes, G. and Barateiro, J. and Vieira, R. and Borbinha, J. (2011), "Control Objectives for DP: Digital Preservation as an Integrated Part of IT Governance", In Proceedings of the ASIST Annual Meeting, 2011, New Orleans, USA: American Society for Information Science and Technology.

Becker, C. and Kraxner, M. and Plangg, M. and Rauber, A. (2013), "Improving decision support for software component selection through systematic cross-referencing and analysis of multiple decision criteria", in Proceedings of 46th Hawaii International Conference on System Sciences (HICSS)*,* 2013*,* Maui, USA, pp 1193-1202.

Becker, C. and Duretec, K. and Petrov, P. and Faria, L. and Ferreira, M. and Ramalho, J.C. (2012), "Preservation Watch: What to monitor and how", in Proceedings of the 9th International Conference on Preservation of Digital Objects (iPRES)2012, Toronto, Canada.

Becker, C. and Duretec, K. and Faria, L. (2014). "Scalable Decision Support for Digital Preservation". OCLC Systems & Services, volume 30, no. 4.

Becker, C. and Kulovits, H. and Guttenbrunner, M. and Strodl, S. and Rauber, A. and Hofman, H. (2009), "Systematic planning for digital preservation: evaluating potential strategies and building preservation plans", International Journal on Digital Libraries, Volume 10, Issue 4, pp 133–157.

Becker, C. and Duretec, K. (2013), "Free Benckmark Corpora for Preservation Experiments: Using Model-Driven Engineering to Generate Data Sets", in Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital libraries (JCDL), 2013, Indianapolis, USA, pp 349-358.

Becker, C. and Rauber, A. (2011a), "Decision criteria in digital preservation: What to measure and how", Journal of the American Society for Information Science and Technology, Volume 62, Issue 6, pp 1009-1028.

Becker, C. and Rauber, A. (2011c), "Preservation Decisions: Terms and Conditions Apply. Challenges, Misperceptions and Lessons Learned in Preservation Planning", in Proceedings of the 11th annual international ACM/IEEE Joint Conference on Digital libraries (JCDL), 2011, Ottawa, Canada, pp 67-76.

Gallagher, M. (2013a), "Improving Software Sustainability: Lessons Learned from Profiles in Science" in Proceeding of Archiving 2013, Washington D.C., USA, pp 74-79.

Gallagher, M. (2013b), “Why can’t you just build it and leave it alone?”, retrieved from http://blogs.loc.gov/digitalpreservation/2013/06/why-cant-you-just-build-it-and-leave-it-alone/ .

Hedstrom, M. (1998), “Digital Preservation: A time bomb for digital libraries”, in Journal of Computers and the Humanities, 1997, Volume 31, Issue 3, pp 189–202.

ISO (2010), “Space data and information transfer systems - Audit and certification of trustworthy digital repositories (ISO/DIS 16363)*”*, International Standards Organisation.

IT Governance Institute, 2007. COBIT 4.1 Framework.

Jurik, B. and Nielsen, J. (2012), “Audio Quality Assurance: An Application of Cross Correlation” in Proceedings of the 9th International Conference on Preservation of Digital Objects (iPRES)2012, Toronto, Canada.

Kulovits, H. and Rauber, A. and Kugler, A. and Brantl, M. and Beiner, T. and Schoger, A. (2009), “From TIFF to JPEG2000? Preservation Planning at the Bavarian State Library Using a Collection of Digitized 16th Century Printings”, in *D-Lib Magazine* ,2009, Volume 15, Number 11/12.

Kulovits, H. and Becker, C. and Rauber, A. (2012), “Roles and responsibilities in digital preservation decision making: Towards effective governance”, in *The Memory of the World in the Digital Age: Digitization and Preservation* 2012, Vancouver, Canada.

Kulovits, H. and Becker, C. and Andersen, B. (2013a), “Scalable preservation decisions: A controlled case study”, in proceeding of *Archiving 2013*. Washington D.C., USA , pp 167-172.

Kulovits, H. and Kraxner, M. and Plangg, M. and Becker, C. and Bechofer, S. (2013b), “Open Preservation Data: Controlled vocabularies and ontologies for preservation ecosystems”, in Proceedings of the 10th International Conference on Preservation of Digital Objects (iPRES)2013, Lisbon, Portugal.

Law, M.T. and Thome, N. and Gançarski, S. and Cord, M. (2012), "Structural and visual comparisons for web page archiving", in Proceedings of the 2012 ACM symposium on Document Engineering (DocEng'12), 2012, New York, NY, USA, pp 117-120.

OCLC and CRL (2007), "Trustworthy Repositories Audit & Certification: Criteria and Checklist".

Petrov, P. and Becker, C. (2012), "Large-scale content profiling for preservation analysis", in Proceedings of the 9th International Conference on Preservation of Digital Objects (iPRES)2012, Toronto, Canada.

Ross, S. and McHugh, A. (2006), "The Role of Evidence in Establishing Trust in Repositories", in *D-Lib Magazine*, 2006, Volume 12, Number 7/8.

Sinclair, P. and Billenness, C. and Duckworth, J. and Farquhar, A. and Humphreys, J. and JArdine, L. (2009), "Are you Ready? Assessing Whether Organisation are Prepared for Digital Preservation", in Proceedings of the 6th International Conference on Preservation of Digital Objects (iPRES)2009, San Francisco, USA, pp 174-181.

Webb, C. and Pearson, D. and Koerbin, P. (2013), "Oh, you wanted us to preserve that?! Statements of Preservation Intent for the National Library of Australia's Digital Collections", in *D-Lib Magazine*, 2013, Volume 19, Number 1/2.

Acknowledgements

Part of this work was supported by the European Union in the 7th Framework Program, IST, through the SCAPE project, Contract 270137, and by the Vienna Science and Technology Fund (WWTF) through the project BenchmarkDP (ICT12-046).

Author Biographies:

Christoph Becker is Assistant Professor at the Faculty of Information and Director of the Digital Curation Institute at the University of Toronto, and Senior Scientist at the Information and Software Engineering Group of the Vienna University of Technology in Austria. His research is focused on understanding sustainability in a digital world and enabling better decision making, especially in the area of digital libraries, digital curation, digital stewardship, digital preservation, digital archiving.

It pursues an Information Systems approach to Digital Curation and Preservation, emphasizing sustainable software design, scalable analytics, systematic experimentation, and organizational improvement. He completed his doctoral degree in computer science at the Information and Software Engineering Group of the Vienna University of Technology in Austria in 2010. In 2010-2011, he was a visiting researcher with INESC-ID in Lisbon, Portugal. He was involved in DELOS, DPE, PLANETS and SHAMAN, is Principal Investigator of the project BenchmarkDP and until 2013 was leading the sub-project Scalable Planning and Watch of the European Commission funded project SCAPE.

Kresimir Duretec is a Researcher at the Department of Software Technology and Interactive Systems at Vienna University of Technology. He is currently pursuing his PhD at the same department. Previously he graduated with MSc and BSc in Computer Science from the University of Zagreb in 2011 and 2009 respectively. In 2013-2014, he was leading the sub project Scalable Planning and Watch at the EU funded project SCAPE and is currently a researcher in the BenchmarkDP project. His current major interest is centered around applying model driven engineering principles to the systematic assessment and benchmarking of digital preservation systems.

Luís Faria is a Systems and Informatics Engineer at the University of Minho, Portugal. He was part of the original development team of RODA (Repository of Authentic Digital Objects) and has engaged in R&D tasks dedicated to systems design, service oriented architectures, format migration services and database preservation. Currently, he is the Innovation Director at KEEP SOLUTIONS. He is involved in European research projects focused on digital preservation such as SCAPE, 4C and EARK, and is pursuing his Ph.D. in digital preservation at the University of Minho.

[i] http://www.openplanetsfoundation.org/blogs/2013-01-09-year-fits

[ii] http://www.openplanetsfoundation.org/blogs/2012-11-06-running-apache-tika-over-arc-files-using-apache-hadoop

[iii] http://hadoop.apache.org

[iv] https://tika.apache.org/1.4/formats.html

[v] http://en.statsbiblioteket.dk

[vi] http://www.openplanetsfoundation.org/blogs/2013-01-09-year-fits

[vii] In Plato, the scoring functions range between 0 and 5, with 0 being unacceptable, and are aggregated across the goal hierarchy. This is discussed in detail in (Becker et al, 2013).

[viii] http://www.w3.org/RDF/

[ix] http://www.w3.org/TR/owl2-overview/

[x] http://opensource.org/licenses

[xi] https://github.com/openplanets/policies

xii (Gallagher 2013, Gallagher 2013a)

[xiii] http://openplanetsfoundation.org/

[xiv] https://github.com/fasseg/scape-tck

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

[xvi] https://github.com/openplanets/plato

[xvii] www.benchmark-dp.org