# Audit DRAMBORA for Trustworthy Repositories: A Study Dealing with the Digital Repository of Grey Literature

Petra Pejšová, National Technical Library, Czech Republic
Marcus Vaska, University of Calgary, Canada

## Abstract

The credibility of a grey literature digital repository can be supported by a specialized audit. An audit of credibility declares that the digital repository is not only a safe place for storage, providing access and migrating to new versions of document formats, it also asserts the care components required of a digital repository environment, including the mandate, typology, policy, team, etc. This audit is very important in showcasing to participants and users the quality and safety of the data process.

This paper will present DRAMBORA (Digital Repository Audit Method Based on Risk Assessment), a methodology and tool for auditing a trustworthy digital repository of grey literature. DRAMBORA is an online instrument which helps organizations develop documentation and identify the risks of a digital repository. DRAMBORA is accessible from http://www.repositoryaudit.eu. The paper will also summarize prevailing advantages and disadvantages of DRAMBORA.

The second part of this paper will describe the audit of the National Repository of Grey Literature (NRGL) as a trustworthy digital repository using DRAMBORA as part of creating a digital repository of grey literature in the National Technical Library (NTK). The most important outcome of the audit was represented by the identified risks connected to the repository and potentially endangering its operation, quality, image, and other features. The main principle of the DRAMBORA audit and, at the same time, its main contribution, is its iteration (i.e. its repetition after a certain time period in new conditions when the original risks are reassessed; the measurements adopted for solution are assessed and new risks are identified).

**Topic Suited to Paper:** Repositories, Quality Control

**Keywords**: audit, credibility, gray/grey literature, methodology, repository, trustworthiness

**Introduction: Audit for Trustworthy Repositories**

"One of the central challenges to long-term preservation in a digital repository is the ability to guarantee the authenticity and interpretability (understandability) of digital objects for users across time"

(Susanne Dobratz and Astried Schoger, 2007)

In our technologically-enhanced environment, managing, preserving, and storing material for posterity is essential, regardless of whether the material in question is a paper file or a digital object (Ambacher, 2007). In fact, efforts at maintaining a stronghold over digital records has been attempted since the 1960s, however, awareness surrounding the true digital repository has only existed for the past decade. This has lead to a number of organizations, most notably the Research Libraries Group (RLG)/U.S. National Archives and Records Administration (NARA) to establish an audit for certifying and enhancing the credibility of grey literature digital repositories. As with any marketing campaign, creating awareness of an initiative and gaining the public's trust is fundamental to ensure success. The Audit Checklist developed by RGL and NARA in 2005 supports this notion with its goal to "develop criteria to identify digital repositories capable of reliably storing, migrating, and providing access to digital collections…a method by which...customers could gain confidence in the authenticity, quality, and usefulness of digitally archived materials" (Ambacher, 2007, p. 2).

Long-term preservation of the material contained within digital repositories functions similarly to the storing of paper documents in a traditional index file within an archive. Ever since institutional repositories arose and began gaining acceptance in the 1990s, efforts at sustaining the material within these storage banks for generations to come have been explored. The first such effort occurred in 1996 when the Task Force on Archiving of Digital Information drew attention to the need for a certification program for the long-term preservation of digital repositories, proclaiming that repositories "must be able to prove that they are who they say

they are by meeting or exceeding the standards and criteria of an independently-administered program for archival certification." (Dobratz and Schoger, 2007, p. 210).

While traditional publishing ventures often result in a considerable time lag between an author's manuscript submission, peer-review by a panel of experts, and subsequent publication in a leading journal within a particular discipline, digital libraries, and in particular digital repositories, allow an author to submit a presentation, thesis, report, etc., as soon as it is written. Further, the author is able to choose from a number of creative commons licenses, maintaining control over his/her data, and deciding how and by whom the data can be accessed (Ambacher, 2007).

**Credibility of Grey Literature Digital Repositories**

As with any research pursuit, guidelines must be followed and adhered to in order to gain credibility and reputation that a chosen research path is indeed the right one. The same holds true when evaluating the trustworthiness of institutional repositories. Although researchers caution that the approaches used in a national repository could well transcend boundaries and apply to international pursuits, it does not necessarily lead to only one universal tool for preserving digital material over the long-term (Dobratz and Scholze, 2006). Rather, the major task of any repository should be "evaluating and disseminating examples of good or best practice and by initiating and intensifying regional, national, and international collaboration" (Dobratz and Scholze, 2006, p. 583).

In order for a repository to be deemed trustworthy, it must operate "according to its objectives and specifications (it does exactly what it claims to do)" (Dobratz and Schoger, 2007, p. 212). Further, a repository must contain information that is complete and control for any unplanned changes, whether these changes are accidental technological glitches or deliberate sabotage. It therefore becomes essential that any edits to any part of a record, once it has already been placed in the depository, is meticulously noted.

Dobratz and Schoger (2007) also make mention of groups of users whose particular interests lie in ensuring that the trustworthiness of repositories is maintained. These include users who wish to access the information, data producers and content providers, and funding

agencies. In addition, repositories that wish to remain functional, trustworthy, and in business for many years down the road must "fulfill legal requirements…to survive in the market" (p. 212). A trustworthy digital repository puts the author's mind at ease, knowing that their information is secure, and will be preserved with the utmost integrity (Dobratz and Scholze, 2006). As previously mentioned, the RLG/NARA audit checklist and the Nestor certificate may be the most well-known means to prove the validity and trustworthiness of a repository, but they are by no means the only methods in existence.

**What an Audit Represents**

A question that should weigh heavily on the minds of any institution containing a digital repository is to assess what an audit represents to establishing criteria and trustworthiness, and what decisions must be made in order to either carry along the same work, or guide the repository in a different direction. Further, in order for a repository to be deemed trustworthy, it must meet its objectives, and contain information and material according to its mandate. There is certainly a strong tie between a trustworthy repository and its information technology infrastructure, dependent upon a number of competing factors. These include integrity, authenticity, confidentiality, and availability (Dobratz and Schoger, 2007). Authenticity precludes that the repository meets its objectives, containing information and material that, by its mandate, it is supposed to contain.

As Dobratz and Schoger (2007) explain, "availability is a guarantee of access to the repository…and that the objects within the repository are interpretable" (p. 212). This is essential for ensuring a repository's survival: repeated difficulties encountered with retrieving a specific item within a repository, or continuous maintenance resulting in repository downtime will result in clients choosing to deposit and/or access their material elsewhere. Allowing the owners of the repository to determine who should be granted permission to access the repository's contents instills a higher level of confidence for the depositing author, as he/she is able to upload and tag his/her own publications. Nevertheless, this level of access can be difficult to maintain. (Dobratz and Scholze, 2006).

Hou, Wojcik, and Marciano (2011) provide a voice that many institutions housing digital repositories can relate to: “integrity is an essential component of a trusted digital repository…all of the functional areas will have an audit trial” (p.182). Thus, establishing an audit for trustworthy repositories represents evidence gathered (usually by means of a checklist) measuring whether or not the repository adhered to pre-determined established evaluation criteria. Further, as digital repositories are primarily web-based programs relying on a server housed in the home institution, these repositories must have “a succession plan or escrow arrangements in place in case the repository ceases to operate.” (Ambacher, 2007, p. 6). Ambacher also posits that data loss, whether accidental or intentional, will inevitably occur, a potential weakness that can be exploited. Therefore, maintaining a sustainable repository with a firm foundation, along with establishing a back-up alternate route in the event of a digital disaster, is essential.

While gathering appropriate hardware, establishing a reliable and secure network connection, and ensuring that a digital repository is utilized to its full potential are all essential components of certification; having the appropriate software to run the repository cannot be overlooked. The *Audit Checklist for the Certification of a Trusted Digital Repository*, jointly created in 2005 by the RLG and NARA, comments on the framework used to evaluate such common repository software packages such as DSpace, Eprints, and Greenstone (Kaczmarek et al., 2006). Regardless of the software package that is chosen, it must be applicable and adaptable, in order to “facilitate data transfer…easily…to take advantage of future, unforeseen developments in computer software and technology” (p.2).

The goal of the RLG/NARA Audit Checklist is “to develop criteria to identify digital repositories capable of reliably storing, migrating, and providing access to digital collections” (Kaczmarek et al, 2006, p. 4). Adhering to the three key areas of digital preservation (namely, technology, resources, and management), the Audit Checklist consists of four key sections: organization; repository functions, processes, and procedures; designated community and the use of information; technologies and technical infrastructure (pp. 4-5).

**Reasons Why an Audit is Done**

If a digital repository is mapped out appropriately, it can have tremendous benefit to both the author depositing research material, and the institution responsible for its upkeep and maintenance. Therefore, an audit need not necessarily be seen as a negative or patronizing activity, but rather as a means of establishing credibility, and educating the repository owner as to any changes that may be required in order to help the repository gain trustworthiness among its users. Of the numerous reasons for why an audit is undertaken, the following are considered to be the core criteria that is often adhered to: an audit should maintain a sustainable, secure repository, with a user-friendly interface; it should establish and maintain a policy that will result in a long-term repository for data producers; it will benefit from a solid management foundation, ensuring that high-quality information is continuously deposited; finally, an audit must identify weaknesses and risks, and establish a process to overcome these challenges (Prieto, 2009).

As the recent copyright issues in Canada indicate, particularly the current Access Copyright befuddlement that exists at some academic institutions, there are a number of legal ramifications that must be taken account when depositing material into any repository. The repository ownership must allow material to be uploaded, stored in an archive, and modified, as required, for posterity (Dobratz and Scholze, 2006, p. 587). Additional challenges faced by these institutions result from the speed in which some repositories have been established. As Downs and Chen (2010) explain, methods for storing and preserving digital content have not yet reached the level of organization used to house non-print material. This can raise doubts about the content of a digital repository, as "trust encompasses not only the integrity of the digital data, but also the authenticity of the links between the data and the data sources and documentation" (Downs and Chen, 2010).

Security of the contents within a repository will always play a prominent role. Repositories should be accessible around-the-clock, and include digital signatures as well as digital object identifiers (DOI) to be able to easily retrieve a requested file. In addition, the establishment of a consistent archiving format will ensure that documents are preserved for

many years into the future. In fact, “the minimum availability of a document [should] be no less than five years” (Dobratz and Scholze, 2006, p. 590).

While supporters of the Open Access Movement would declare that the full contents of a repository should be freely and publically available to all (and indeed, this is the case with a number of institutional repositories, including DSpace at the University of Calgary), there are nevertheless a number of interest groups for whom trustworthiness holds particular merit. These include users who wish to access reliable information immediately and well into the future, content providers who rely on the audit of a repository to support their effort at ensuring high-quality information in a repository is maintained (i.e. a warranty for data producers), and corporations that determine whether or not a repository will receive adequate funding and for how long. Finally, as previously mentioned, entering the digital repository environment is indeed a competitive venture, and all repositories are therefore required to “fulfill legal requirements” (Dobratz and Schoger, 2007, p.212) in order to survive.

One methodology posited by Kaczmarek and colleagues (2006) is the creation of a matrix to function as a tool which will aid in the decision-making process of certifying a repository as a trustworthy source of information. Kaczmarek et al (2006) explain that settling on which software package best suits a particular repository will lead to a rubric “to determine how critical each particular point of functionality is and if that point is absolutely required” (p.2). Such steps were taken by the Exploring Collaborations to Harness Objects in a Digital Environment for Preservation (EXCHO DEPository) project, a joint effort between the National Digital Information Infrastructure and Preservation Program (NDIIPP) at the Library of Congress, and the University of Illinois at Urbana-Champaign.

While the above examples of digital repositories comment on the importance of establishing policies that are firmly adhered to in order to establish trustworthiness and acuity, repositories must also be established in such a way that they can be easily customized if necessary. Such is the case with DCAPE, the Distributed Custodial Archival Preservation Environments project, originating out of the University of North Carolina Chapel Hill (Hou, Wojcik, and Marciano, 2011). Adhering to the three key preservation policies, namely

"management of archival storage, validation, and trustworthiness" (p. 181), DCAPE supports one of the fundamental reasons why an audit of a repository is undertaken. Ensuring that high quality material is continuously deposited is certainly one way of ensuring a repository's livelihood, however without a user-friendly interface, authors and researcher's alike may become frustrated and choose to deposit their publications elsewhere, which, in turn, reflects negatively on the purpose of sustaining the repository for generations to come.

**Existing Audit Methodologies and Tools**

DINI, the Deutsche Initiative für Netzwekinformation, is aimed at supporting the Open Access movement in Germany. The aim of this guideline is to enhance the cooperative partnership between German educational institutions with a goal to "provide a tool for repository operators that could be used to raise the visibility, recognition, and importance of the digital repository within the university." (Dobratz and Scholze, 2006, p. 584).

As exemplified in many repositories, DINI criteria are based on two categories, the first of which explains the minimum requirements that must be captured in order for the repository to be deemed credible. These requirements include visibility and server policy, support for authors, legal issues, authenticity and integrity, indexing, impact and access to statistics, as well as long-term availability (Dobratz and Scholze, 2006, p. 585). Nevertheless, despite these rather strict requirements, Dobratz and Scholze comment on the challenges involved in deeming a repository to be both trustworthy and credible, hence the need for an audit. These include the establishment of a server policy, creating a visible service for authors, and implementing persistent identifiers (p. 586).

In addition to the aforementioned repository requirements, DINI also supports the need for creating open access to archived materials, and posits that a policy needs to be established to allow for each repository to be registered and recognized by large-scale collectives, namely the Directory of Open Access Repositories, OpenDOAR. (Dobratz and Scholze, 2006). As DINI proclaims, creating an open access policy showcases "a clear commitment to support the 'green way' to open access" (p. 587).

Originally created with cultural heritage organizations in mind, the Nestor Catalogue of Criteria for Trusted Digital Repositories serves as a guide for planning and maintaining digital repositories well into the future (Dobratz and Schoger, 2007). The criteria raised by Nestor include the following key concepts which can be applied to virtual any repository framework: compliance with terminology created by the Open Archival Information System (OAIS), abstraction, adequate documentation, transparency (essential to gain trust), adequacy, and measurability. As Dobratz and Schoger (2007) proclaim, these criteria will function as "indicators showing the degree of trustworthiness" (p. 214). The organizational structure for Nestor is divided into three top-level categories, each with a number of subdivisions. These are depicted as follows: organizational framework (defined goals, adequate usage, legal and contractual rules, organizational form, quality management), object management (integrity, authenticity, strategic plan for technical preservation, acceptance from producers adhering to established criteria, archival storage, usage, data management system), and infrastructure and security (adequate IT infrastructure, protecting the repository and the objects contained within it) (pp. 215-216). [See Appendix 1].

**(DRAMBORA): A Methodology and Tool for Auditing a Trustworthy Digital Repository**

**DRAMBORA description: tool and methodology**

Launched in 2008, as the result of a joint effort between the Digital Curation Centre and Digital Preservation Europe, the Digital Repository Audit Method Based on Risk Assessment (DRAMBORA) functions as a toolkit "to make the self-auditing process easier and more efficient for repository managers" (Donnelly et al., 2009). Although digital repositories had already been in place for some time prior to the establishment of DRAMBORA, there was no standard guideline for determining the key components required for successfully implementing, initiating, and sustaining a digital archive. This issue led to the Centre for Research Libraries (CRL), widely credited as the developers of DRAMBORA to produce a list of 10 core requirements that all digital repository owners must be made aware of and should follow to preserve their archival storehouses for generations to come (see Appendix 3). As can be seen from this list, technological infrastructure plays only one part in ensuring that the data within a

repository is adequately stored and maintained over time. Creating a manageable process and action plan, along with accounting for any legal ramifications that may manifest themselves along the way, is equally important.

While DRAMBORA is a relatively recent phenomenon, it nevertheless underwent a series of pilot tests in the two years prior to its official unveiling. More than merely serving as another toolkit, it is primarily responsible for presenting "a methodology for self-assessment, encouraging organizations to establish a comprehensive self-awareness of their objectives, activities, and assets before identifying, assessing, and managing the risks implicit within their organizations" (Donnelly et al., 2009). Undoubtedly, attempting to maintain any form of electronic storage method implies a certain amount of risk, perhaps even more so than a traditional print collection. DRAMBORA has attempted to ease this risk process by positing a series of stages: authors are required to develop an organizational profile, describe and document a mandate, and list objectives/goals, activities, and assets (Donnelly et al., 2009). These stages are, however, only meant to serve as guidelines; the DRAMBORA team cautions that the entire purpose of this audit method is to serve as a living document, with revisions being made along the way as the need arises.

**How to Use DRAMBORA**

There are presently 18 institutions that rely on the DRAMBORA toolkit to conduct self-assessment audits of their digital repositories. A number of these organizations also hold strong ties to the grey literature community. Before describing how to use this methodology it is necessary not only to discuss the purpose of DRAMBORA (see Appendix 2), but also its three primary applications: as a web-based tool, DRAMBORA can assess the effectiveness of a repository infrastructure, and offer suggestions for its improvement; it acts as a preparatory resource for external auditors who may wish to serve as aggregators of the DRAMBORA movement; finally, it anticipates any potential weaknesses or challenges, and subsequently adjusts its plans to overcome these boundaries (Donnelly et al., 2009).

Existing in both an online and offline format, DRAMBORA is a user-friendly program that guides the user through four phases. First, the user is encouraged to register for a personal

account, as well as provide details regarding the repository at his/her institution. This allows DRAMBORA to present a customized self-assessment profile for each user. Further, additional staff members from the institution in question will be identified, where they will be able to contribute to the self-assessment process. The subsequent phase of using DRAMBORA refers to the actual self-assessment audit. The goal of this stage is to ensure that the repository undergoing an audit establishes clear objectives with documented sources. The organization's mandate and/or vision/mission statement, along with any potential legal or technical issues should be listed here.

Following the actual self-assessment, any potential risks must be identified and assessed. For evaluative purposes, all of the risks identified are categorized and assessed according to their potential impact, accounting for the frequency or probability that any potential negative effects and subsequent risks could appear. (Donnelly et al., 2009). Finally, once all risks have been assessed and identified, a plan should be established to develop counter-measures, anticipated outcomes, and a timeline to reassessment the repository, to ensure that any issues have been resolved. The careful mapping of a repository using the DRAMBORA tool may already give an overview of what is finished and what is not, which documents, procedures, tools, and measurements are missing and where most critical risks reside.

**DRAMBORA: Advantages and Disadvantages**

Undoubtedly, as a web-based self-audit tool, DRAMBORA far surpasses a number of competitors in this field, and thus it would not do the tool justice without mentioning some of its key benefits. First and foremost, the online version of this program allows the user to view the internal activity of a repository, identifying any potential problems, and rectifying them as quickly as possible. In addition, the user is able to interact with the content on the screen, navigating to sections of interest without having to flip through an entire text. Second, the methodology and tools are well implemented, clearly identifying the organizational role and structure of each institution taking part in the audit. Third, the descriptions used and examples presented are pertinent, intuitive, and applicable to the task at hand. This includes a clearly-

defined mission statement, complete with key aims and objectives. Finally, the scale of evaluation compared to the risks is adequately assessed: "an internal understanding of the successes and shortcomings of the organization [enable it]…to effectively allocate or redirect resources to meet the most pressing issues of concern" (Donnelly et al., 2009). "For inspiration and possible direct help, the DRAMBORA tool contains a number of links to supporting documents and a range of practical examples of completed entries, whether in the preparation phase or the audit phase. In the area of risk identification, predefined risks can be directly used and modified or unique risks may be formulated." (Karlach, 2010, p. 127)

Despite the obvious benefits of DRAMBORA, there are nonetheless a few disadvantages that must also be considered. Namely, at present, the implementation and methodology is only available in English. Although English is seen as the universal language of communication, it is important to note that the majority of DRAMBORA users hail from European countries. Thus, offering the DRAMBORA interface is a variety of languages is a task that the developers of this toolkit would be wise to consider. An additional disadvantage relates to the technical, albeit programming aspect of this product. DRAMBORA functions perfectly fine on a standard Windows or Mac interface, but is not compatible with the Czech Windows operating system (i.e. it does not support the Czech character set – iso-8859-2/windows-1250). While it is understandable that this program cannot comply with every possible computer operating system, perhaps a text-based (DOS) version, in addition to the current HTML version, should be brought forward to the development team. The final two arguments surrounding the negative aspects of DRAMBORA center on access issues. Presently, read-only access is not permitted (a user must be fully registered and log in to a created account, in order to make use of al of the program's features). In addition, exporting records (either via e-mail or to a bibliographic management program) is currently not possible.

**Audit of the National Repository of Grey Literature (NRGL)**

**Introduction to NRGL**

The NRGL project, *The Digital Library for Grey Literature – Functional Model and Pilot Implementation*, started thanks to the support from the *Ministry of Culture of the Czech*

*Republic* as part of both research and development programs. The project is divided into three phases, lasting from 2008 to 2011. Its main goals are the systematic collection, long-term archiving and provision of access to specialized grey literature, pertaining specifically to research and development, civil service and education, as well as with the business sphere and "open access" at the national level. To support this goal, the NTK created a functional network of partner organizations, a working model, and a pilot application. In addition, on the basis of verified technology and methods defined under the project, recommendations and standards are created for other institutions electing to build their own digital grey literature repositories. Recommendations and standards consist mainly of a preferred metadata format, exchangeable formats and templates, examples of licensing models and of legal issues resolved, preservation methodology, archives, and the provision of access to digital data.

**NRGL Audit**

The first audit of the NRGL as a trustworthy digital repository using the toolkit and methodology of DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) was performed at the end of 2009 as a part of creating a digital repository of grey literature in the National Technical Library (NTK). The audit results and experience from its course were summarized in a final report, and published in the book *Grey Literature Repositories* in 2010. The most important outcome of the audit was presented by identifying risks connected to the NRGL and potentially endangering its operation, quality, images, and other features; these risks were eliminated or moderated by the NRGL team during 2010. The main principle of the DRAMBORA audit and, at the same time, its main contribution and iteration, resulted in its repetition after a certain time period in new conditions when original risks were reassessed, the measurements adopted for solution assessed, and new risks identified.

The second audit of the NRGL digital repository was performed after one year. In this audit, the actual state of the repository was assessed, with progress achieved during 2010. New potential risks were identified, as well as possible ways to eliminate them or to reduce their impact. The NRGL documentation, the description of the whole project, its processes,

procedures, and related documents developed significantly during 2010, which, in turn, made a solid basis for the audit.

Work with the DRAMBORA Tool went on without any of the issues and problems experienced with the previous audit. On the basis of lessons learned from 2009, the NTK communicated with the authors of the tool and methodology, and proposed some improvements and modifications. As a result, the web tool DRAMBORA appeared more stable after a one-year pause; however, the most unexpected modifications have not been introduced yet, especially regarding the elimination of the rather unpleasant fact that the online version of DRAMBORA does not support languages other than English at this time. Nevertheless, as the results of the audit are intended to be presented in the international field, namely in the area of grey literature projects, we will continue to use English.

**NRGL Audit: Preparation and Definition**

DRAMBORA Interactive was used during the preparation phase in addition to the audit phase (Karlach, 2010, p. 126-127). The preparation phase consisted of acquiring all relevant information and documents on the status of the repository, its description, standards, procedures, staff, material, budget, etc. (Donnelly et al., 2009). This information served as input data for the preparation phase of the audit and was entered into the DRAMBORA Interactive in the section *Before the Assessment*. Here, the repository was described, the scope of relevant areas (Functional Classes) of the audit were defined, and the repository staff, including a detailed description of individual team members and their roles, was listed. The definitions of staff roles were especially important since, at the subsequent risk identification stage, it was necessary to relate risks to respective roles. Even during the preparation phase, a substantial contribution might be made to an audited repository. This helps the staff see the repository from a global vantage point, to map and accumulate the most important descriptive data about the repository, and to point to possible deficiencies and defects, offering the opportunity for problems to be remedied and missing materials to be completed. The audit was run using the portion of the DRAMBORA tool called the Assessment Centre (Donnelly et al., 2009). Here, the repository mandate, including its mission, purpose, founders, etc., was defined. Other

repositories were also identified that influenced its activity, both external (e.g., legislative) and internal (e.g., organization, content type restrictions, etc.). The audit continued by defining repository goals, activities, and the means used to achieve these goals.

**NRGL Audit: Identified Risks**

In addition to the mapped repository and the relevant environment, the producers of the methodology and the tool consider the most important output to be the analysis of identifiable risks endangering the repository, its quality, readiness, reputation and position in the eyes of both specialists and ordinary users. In the 2010 audit, 16 risks from the previous audit were assessed, primarily regarding the progress in their elimination, and an additional 8 new risks were identified. The NRGL repository is still in the pilot project stage; however, it is run on final software versions and real data are being stored. Identified and reassessed risks mainly refer to the description of activities and procedures of the repository, the state and development of the staff, project funding, hardware and software sources, including their backup and relationship to the NRGL environment (see Appendix 5).

After all necessary information is entered, the Reporting Centre function helps to create output reports on the identified risks for the repository, with respect to their relationship and plausible solutions. Two types of output report formats are available, either PDF or HTML. Other saved descriptive information cannot be exported easily, however, it is possible to copy saved snapshots of the audit page. Besides the mapped repository and its relevant environment, the producers of the methodology and tool consider the most important output to be the analysis of identifiable risks endangering the repository, its quality, readiness, reputation, and position in the eyes of both specialists and ordinary users. Since the DRAMBORA tool does not provide read-only access, it is regrettably not feasible to allow free access to the audit at this time.

Generally, risk elimination is much easier in a case where the respective area is fully under control and in charge of the NRGL management and team. If the risk relates to the cooperation within or even outside the NTK, the situation is considerably more complicated. The creation of a knowledge database NRGL Wiki indicates great progress; this database should

be further developed and strictly adhered to, as the continuous documentation of procedures, activities and results of the NRGL team is of crucial importance for the elimination or minimization of the impact of most risks. Such progress may be seen in the development of the NRGL repository since the last audit along with the new activities and goals that have been added. Therefore, the documentation of activities and analysis of risks are most important. A large portion of risks reflect the topic of building the NRGL partner network, i.e. the partner network of providers of the repository content. Consequently, this area should be of priority especially for promotion and education. In relation to NRGL partners, sufficient attention should be paid to legal issues connected to the Author Act.

**Conclusion**

A yearly repetition of the audit under new conditions, identification of new or modified risks, and creation of another action plan make the audit an iterative process that contributes to the trustworthiness of the NRGL. Despite the valiant efforts of libraries, information technology specialists, and researchers, who devote considerable amounts of time and effort to maintain credible digital repositories, it can seem like a tall barrier to overcome. While Downs and Chen (2010) caution that "no organization can absolutely guarantee long-term preservation and access", efforts to establish methods of audit and recognize trustworthy digital repositories must continue. As DRAMBORA and the subsequent audit of the National Repository of Grey Literature have shown, the task at hand may not yet been complete, but it is certainly moving in the right direction. It is thus perhaps fitting to conclude with the mission statement of Columbia University, which reflects not only on the goals of this particular institution, but which speaks to the efforts of raising awareness of grey literature in all topic fields and venues. Namely, one must "advance knowledge and learning at the highest level and…convey products of its efforts to the world" (Columbia Mission Statement, 2011). We therefore recommend that an audit be undertaken on an annual basis, identifying any associated risks, and creating an action plan to make the audit an iterative process that contributes to the trustworthiness of the digital repository.

## References

Ambacher, B. (2007). Government archives and the digital repository checklist. *Journal of Digital Information, 8*(2), 1-10.

Columbia University. (2011). *Mission Statement*. Retrieved November 20, 2011 from http://www.columbia.edu/content/mission-statement.html

Dobratz, S., & Schoger, A. (2007), Trustworthy digital long-term repositories: The Nestor approach in the context of international developments. *Research and Advanced Technology for Digital Libraries, Proceedings, 4675,* 210-222.

Dobratz, S., & Scholze, F. (2006). DINI institutional repository certification and beyond. *Library Hi Tech, 24*(4), 583-594.

Donnelly, M., Innocenti, P., McHugh, A., & Ruusalepp, R. (2009). *DRAMBORA Interactive User Guide*. Glasgow. Retrieved November 22, 2011 from http://www.repositoryaudit.eu/help/

Downs, R. R., & Chen, R.S. (2010). *Self-assessment of a long-term archive for interdisciplinary scientific data as a trustworthy digital repository*. Retrieved October 22, 2011 from journals.tdl.org/jodi/article/download/753/642

Hou, C.Y., Wojcik, C., and Marciano, R. (2011). Trusted digital repository design: A policy-driven approach. *Archiving, 7*, 181-186.

Kaczmarek, J., Hswe, P., Eke, J., & Habing, T.G. (2006). Using the Audit Checklist for the Certification of a Trusted Digital Repository as a framework for evaluating repository software applications. *D-Lib Magazine, 12*(2), 1-10. Retrieved August 3, 2011 from http://www.dlib.org/dlib/december06/kaczmarek/12kaczmarek.html.

Karlach, P. (2010). An audit of the National Repository of Grey Literature using the DRAMBORA tool. In Pejšová, P [ed.]. *Grey Literature Repositories*. Zlin: VeRBuM, p. 126-127. Available as an E-book at: http://nrgl.techlib.cz/images/Book.pdf.

National Technical Library. (2008). *Audit of the National Repository of Grey Literature (NRGL) in the NTK using the DRAMBORA tool: Second audit, 2010.* Retrieved August 25, 2011 from http://nrgl.techlib.cz/images/DRAMBORA_2010_EN.pdf

National Technical Library (2008). *National Repository of Grey Literature: An audit of the NRGL as a trustworthy digital repository.* Retrieved August 25, 2011 from http://nrgl.techlib.cz/index.php/Audit

NUŠL. (2011). *National Repository of Grey Literature [NRGL].* Retrieved November 20, 2011 from http://nrgl.techlib.cz/index.php/Main_Page

Pejšová, P. (2010). The development of grey literature in the Czech Republic. In Pejšová, P [ed.]. *Grey Literature Repositories*. Zlin: VeRBuM, p. 34. Available as an E-book at: http://nrgl.techlib.cz/images/Book.pdf.

Prieto, A.G. (2009). From conceptual to perceptual reality: Trust in digital repositories. *Library Review, 58*(8), 593-606.

Ross, S. (2006). The role of evidence in establishing trust in repositories. *D-Lib Magazine, 12*(7-8). Retrieved November 18, 2011 from http://www.dlib.org/dlib/july06/ross/07ross.html

## Appendix 1: TRAC-Checklist

(adapted from: Dobratz, S., & Schoger, A. [2007]. Trustworthy digital long-term repositories: The Nestor approach in the context of international developments. *Research and Advanced Technology for Digital Libraries, Proceedings, 4675,* 210-222.)

Common principles:

1. Continued maintenance of digital objects
2. Organizational fitness
3. Acquires and maintains contractual legal rights; fulfills responsibilities
4. Effective and efficient policy framework
5. Acquires digital objects based on criteria, corresponding to commitments and capabilities
6. Maintains and ensures integrity, authenticity, and usability of digital objects over time
7. Creates and maintains metadata about actions to take on digital objects during preservation, as well as relevant production, access support, and usage process context before preservation.
8. Dissemination requirements
9. Strategic program (preservation planning and action)
10. Technical infrastructure adequate (maintenance and security)

## Appendix 2: Purpose of DRAMBORA Toolkit

(from Donnelly et al., 2009)

1. Defining the mandate and scope of functions of the repository
2. Identifying the activities and assets of the repository
3. Identifying the risks and vulnerabilities associated with the mandate, activities, and assets
4. Assessing and calculating the risks
5. Defining risk management measures
6. Reporting on the self-audit

## Appendix 3: The Ten Core Requirements for Digital Archives

(from Donnelly et al., 2009)

1. Mandate & Commitment to Digital Object Maintenance
2. Organizational Fitness
3. Legal & Regulatory Legitimacy
4. Efficient & Effective Policies
5. Adequate Technical Infrastructure
6. Acquisition & Ingest
7. Preservation of Digital Object Integrity, Authenticity & Usability
8. Metadata Management & Audit Trails
9. Dissemination
10. Preservation Planning & Action

## Appendix 4: NRGL Audit - Examples

DRAMBORA Interactive creates connections among individual parts of audit. See connections marked bold.

**1. Mandate and equipment**

NTK status: To build national repository of grey literature and to make the information and findings contained in the repository accessible for NTK users using modern information technology.
http://www.techlib.cz/default/files/download/id/1747/dodatek-c-1-k-zl-ntkpdf.pdf
Repository Hardware
Hardware used to run the repository software and database - SUN SUNXFIRE 4500 server, OS SOLARIS 10

**2. Functional classes**
Supporting Functions:
Legal & Regulatory Legitimacy
Functions and characteristics corresponding to legislative, regulatory or common law rights and responsibilities of the repository.
Operational Functions:
Acquisition & Ingest
Functions and characteristics corresponding to the repository's negotiation, submission, receipt and ingestion of data from creators and suppliers.

**3. Staff**
Position: Manager
Unique Staff ID: 1
Telephone: +420232002485
Staff Email: petra.pejsova@techlib.cz
Address: NTK, Technicka 6/2710, 160 80 Praha 6
Status: Coordinator
Username: petrapej
Name: Miss Petra Pejsova
Alt. Email: petra.techlib@gmail.com
Roles: Management

**4. Roles**
Role Name: Management
Description: Establishes strategy and objectives of the repository, Establishes strategy of the repository content provider network...
**Corresponding Staff Members: Manager**
Activity Responsibilities: Budget Management, Cooperation Network, Team Management, NUSL Publicity
Risk Responsibilities: Loss of Staff Members, Pilot Project End, Disaster Recovery, Partner Network Voluntary, Backup Tapes...

**5. Constriction**
Name: Documents Publication Status
Description: The Repository is devoted to grey literature, so it accepts only unpublished or semi-published documents
Type: Policy
**Functional Class(es): Supporting Functional Classes - Acquisition & Ingest**
Web Links:
http://nusl.techlib.cz/index.php/Typologie_dokument%C5%AF_NU%C5%A0L

**6. Goals**
Name: Best Practices
Description: Best practices for building similar cooperating institutional repositories are one of the planned outputs of the project
2010: Best practices for partners created in the technical and methodical areas, see section
**Constraints - Methodology of the cooperation with NUSL etc.**
**Functional Class(es)*: Supporting Functional Classes - Efficient & Effective Policies**

**7. Activities**
Activity Name: Repository Backup
Activity Desc: To create a backup copy of the system to preserve the current setup and of the repository database to preserve it's content
**Activity Role(s): Administrator**
**Related Assets: Repository Hardware, Repository Software**
**Related Objective(s): Main Function**
**Functional Class(es)*:Preservation of Digital Object Integrity, Authenticity & Usability**
**Related Risks: Backup Tapes Storage**

## Appendix 5: NRGL Audit – Identified Risks

NRGL was analyzed in all functional classes, however, only the most severe and most obvious risks were recorded according to Pareto's rule of 80/20 – 20% of risks are responsible for 80% of the danger.

16 originally identified risks in 2009:

**Risk Number 1**: Loss of Staff Members

**Risk Number 2**: Pilot Project End

**Risk Number 3**: Disaster Recovery

**Risk Number 4**: Partner Network of Volunteers

**Risk Number 5:** Backup Tapes Storage

**Risk Number 6**: Financial Shortfall

**Risk Number 7**: Budget for Services

**Risk Number 8**: FAST Trial Version

**Risk Number 9**: Weak Mandate

**Risk Number 10**: No Ingest Policy

**Risk Number 11**: Document Formats

**Risk Number 12**: Software Administration

**Risk Number 13**: Undocumented Policies

**Risk Number 14**: Long Term Preservation Strategy not described

**Risk Number 15**: Staff Skills insufficient

**Risk Number 16**: Deliberate System Sabotage

Risks newly identified in 2010:

**Risk Number 17:** Duplicate project

**Risk Number 18**: Partners do not supply full text

**Risk Number 19**: Slow growth of partner network

**Risk Number 20**: Sample partner contract has limited usability (applicability)

**Risk Number 21**: Migration to new HW platform

**Risk Number 22**: New CDS Invenio version

**Risk Number 23**: Legal Risk - Authors Act

**Risk Number 24**: Termination of legal support