
Citation of the final article:

DOI: https://doi.org/10.1177%2F0955749017725938

This is the peer reviewed accepted manuscript.

©2017, The Author

Reprinted by permission of SAGE Publications.

Downloaded from DRO:
http://hdl.handle.net/10536/DRO/DU:30102541

Bernadette Houghton
orcid.org/0000-0001-5730-8805

Abstract
In 2013, Deakin University Library undertook a self-assessment of its research repository, DRO, against ISO 16363. ISO 16363 is heavily structured on the OAIS model, so an understanding of OAIS is critical for a full understanding of the ISO 16363 criteria. With over 100 ISO 16363 criteria, a self-assessment can be a large and complex undertaking. Not only are many of the criteria very complex to understand, their arrangement is such that many of them cover common ground, which may lead to duplication of effort. An effective and efficient self-assessment will include a preliminary review of all the criteria, including an assessment of their applicability and risk to the repository being assessed. If resources are limited, repository managers should also focus their self-assessment on high-risk or vulnerable areas. Regular self-assessments of research repositories are highly recommended. Repository managers also need to ensure that once a self-assessment is completed, that resources are allocated to addressing identified areas of improvement.

Keywords
trusted digital repositories, digital preservation, ISO 16363, ISO 14721, OAIS, self-assessment

Introduction

This paper describes Deakin’s self-assessment process, issues encountered, and recommendations for others considering a similar self-assessment. Three years have now passed since the original self-assessment was completed, and Deakin is looking to undertake a re-assessment in the next year. It is thus timely to reflect back on the original self-assessment and review how Deakin’s original methodology could be improved.

Apart from helping to improve the effectiveness and efficiency of Deakin’s next self-assessment, it is expected that this paper will assist other repository managers contemplating whether to undertake their own self-assessment. Hopefully, it will provide some guidance as to how to best structure the process. It may also, perhaps, inspire those who have not yet started thinking seriously about digital preservation.
ISO 16363
ISO 16363 provides a foundational basis for best practice in digital preservation. Based upon TRAC, which in turn is based upon the Open Archival Information System (OAIS) reference model, it provides a set of criteria against which to assess the trustworthiness of digital repositories.

There are around 105 ISO 16363 criteria, covering 3 areas:

- **Organizational infrastructure.** This covers governance, organisational structure and viability, staffing, accountability, policies, financial sustainability and legal issues
- **Digital object management,** covering acquisition and ingest of content, preservation planning and procedures, information management and access
- **Infrastructure and security risk management,** covering technical infrastructure and security issues.

For most of these criteria, ISO 16363 suggests documentation that can potentially be used as evidence to determine whether a repository meets a particular criteria. It also describes some of the ways a repository can demonstrate it is meeting the criteria. The suggested documentation is not prescriptive and, in practice, the relevant documentation will vary with each organisation.

DRO overview
DRO was established in 2007 in response to the Australian Government’s Research Quality Framework (RQF). RQF was intended to ensure public funding of research was directed to those projects that would most benefit the wider community. A happy side effect was that it encouraged most Australian universities, including Deakin, to set up a research repository for RQF reporting purposes. As it happened, RQF was cancelled before being fully implemented, and was replaced by the current Excellence in Research Australia (ERA) initiative in 2010. RQF’s influence lives on, however, in the entrenchment of research repositories within the Australian higher education sector.

DRO is based upon the Fez/Fedora repository software developed by the University of Queensland. One of the major reasons for choosing Fez/Fedora was its preservation-related functionality. This included support for versioning, JHOVE and PREMIS, as well as checksumming. Flexible digital object models, object relationships and robust object security were also major considerations.

At the close of 2016, DRO contained 67K records, mostly research works. 22% of records contained open access attachments.

In DRO’s first few years, the focus was on establishing workflows for the ingest and dissemination of material, and ensuring RQF/ERA requirements were met. Significant resources were also expended on making adjustments to the software to accommodate increasing expectations, such as open access, and integration with other university software. Several preservation-related initiatives were undertaken over these years. For example, file naming conventions were established, investigations undertaken into various digital preservation tools and standards, and an assessment made of optimal file formats for moving images. However, there was no formal assessment of DRO’s overall reliability in terms of digital preservation. In 2013, a window of opportunity opened up that enabled Deakin to begin investigating the compliance of DRO and its associated workflows with digital preservation best practices.
Why do a self-assessment

A self-assessment of a digital repository helps to identify weaknesses and problem areas that can then be addressed, thus improving the performance and reliability of the repository. A self-assessment does not have to be a formal process; it can be as simple as evaluating a repository’s performance against a list of jotted down criteria. However, the more formalised and independent the criteria and the more rigid the assessment framework, the greater the likelihood of a robust and trustworthy assessment. Assessments done by independent third parties also strengthen the likelihood of trustworthy results.

It should be kept in mind that a self-assessment is, by its nature, not the same as an independent audit. In an independent audit, the auditor is at arms’ length from the repository, and undertakes a series of tests to ensure that workflows are operating effectively and efficiently. Self-assessments may not necessarily involve such depth of investigation. Additionally, local knowledge can be both an advantage and a hindrance. An advantage is that loopholes in workflows and areas of weakness may already be known. At the same time, local knowledge can make it easier to inadvertently overlook existing, as yet unidentified, problems. The main benefit of a self-assessment is that areas of improvement or risk can be identified and addressed without the expense of an external audit.

First steps towards self-assessment

The first step in assessing DRO’s compliance with digital preservation best practices was to undertake a literature review. The review identified several tools that could be used for a self-assessment. These included:

- DRAMBORA (Digital Curation Centre, 2015).
- Data Seal of Approval (Data Seal of Approval Board, 2016).

ISO 16363 and DRAMBORA appeared to be the most widely accepted and well-documented tools for assessing the reliability of digital repositories, so Deakin chose to focus on these.

ISO 16363 and DRAMBORA are actually complementary tools. DRAMBORA focuses on whether a repository meets its own self-defined goals, while ISO 16363 measures a repository against pre-defined criteria. However, the take-up of DRAMBORA appears to have lost momentum in recent years. Consequently, this article focuses on the results of the ISO 16363 assessment, referring only to DRAMBORA in an administrative context.

Together, the ISO 16363 and DRAMBORA assessments took around 6 months to complete. They were resourced via an estimated 25% of the assessor’s time. Fortuitously, the assessor had 4 years internal audit experience, so was familiar with assessment methodologies and compliance checking processes. She had also worked with DRO since its inception, so was also familiar with its workflows and technical issues.

---

1 At the time of researching this paper, the main DRAMBORA website had been offline from at least mid-November 2016 to mid January 2017; it had reappeared by 17 January 2017.
2 The author, a fulltime Deakin University Library staff member.
Assessment methodology

Following the literature review, a project plan was developed, with the objective of assessing DRO’s compliance with both ISO 16363 and DRAMBORA. Work on the ISO 16363 self-assessment began in August 2013, once the DRAMBORA component of the project was fully completed.

Initially, the self-assessment was undertaken one criteria at a time. For each criteria, attempts were made to obtain Deakin equivalents of all ISO 16363’s suggested documentation, as well as any other documentation considered relevant to that criteria. As much of the documentation was spread over different parts of the University, this was a time-consuming task. However, it was made easier by some documentation already having been retrieved as part of the DRAMBORA assessment. As the assessment progressed, it was found to be more effective to review criteria in like groups.

In the early stages of the assessment, it became evident that much of the documentation suggested by ISO 16363 was not materially relevant to Deakin’s circumstances. So from then on, the assessor relied primarily on her own investigations into Library and University policies and procedures to identify relevant documentation.

Once the relevant documentation was obtained for a criteria or group of criteria, an assessment was made of DRO’s compliance. This involved reviewing the documentation and, where relevant, checking that underlying processes and workflows were working properly and compliant with policies and procedures. This strategy of verifying that processes and workflows were working helped to overcome some of the subjectiveness resulting from the assessor’s familiarity with DRO.

All findings were documented in a wiki, with 3 levels of compliance indicated: Full, Part or Not Compliant. Strategies to address areas of weakness were also identified as each criteria was completed.

ISO 16363 and OAIS in practice

Deakin’s self-assessment against ISO 16363 highlighted several issues with the standard itself and the underlying OAIS model.

OAIS conceptuality

When the self-assessment process began, the assessor had only superficial knowledge of the OAIS model. At that stage, ISO 16363 appeared to be fairly straightforward and self-contained, and in-depth knowledge of OAIS was not thought necessary. However, as the assessment proceeded, it became apparent that a greater understanding of OAIS was required to fully understand the ISO 16363 criteria. Chief among this was the conceptual nature of OAIS.

There is a common and pervasive misunderstanding about the OAIS model (Schumann 2012, p.6), which can potentially lead to misinterpretations of ISO 16363, as happened in Deakin’s case. The misunderstanding relates to the fact that OAIS is a conceptual model only. It does not aim to describe workflows, processes and technologies used in an actual archival system. It describes at a theoretical level the framework and requirements of an ideal archival system. The intent of OAIS is to be implementation-independent; that is, independent of any particular software, hardware or set of workflows.
To provide a specific example from Deakin’s self-assessment, ISO 16363 treats Submission, Archival and Dissemination Information Packages (SIPs, AIPs and DIPs) as separate entities. This lead Deakin to initially believe that compliance with ISO 16363 meant that separate submission, archiving and dissemination versions of a file had to be retained. However, this is not so. In practice, a repository can still be compliant with ISO 16363 if it retains only one version of a file\(^3\). The critical factor is that there must be sufficient information provided for the repository to be able to identify or re-create the submission, archival and dissemination files on request. This is where the “package” part of SIPs, AIPs and DIPs comes from.

The creators of OAIS made a concerted effort to address the potential misunderstanding of its conceptual nature in its second incarnation, ISO 14721:2012. Several references are made throughout the text to stress the point that OAIS is a reference model only and not a design or implementation of a specific archival system (Consultative Committee for Space Data Systems 2012, pp. 1-2, 4-3, 4-20 et al). However, ISO 16363 mentions the conceptual nature of OAIS only in passing (Consultative Committee for Space Data Systems, 2011, p.1-1).

A review of the literature, post-assessment, indicated that misunderstanding of the intent and purpose of OAIS is still very common (Giaretta, 2016; Lavoie, 2014; Schumann, 2012).

The ISO 16363 criteria are more closely aligned with actual repository infrastructure, operations and workflows than is the OAIS model. However, there are certainly some areas where the conceptual nature of the ISO 16363 criteria may cause confusion for those unfamiliar with the model; the SIP, AIP and DIP example being just one area, also identified by Lavoie (2014).

**Terminology**

ISO 16363’s use of OAIS terminology caused some confusion in the early stages of the assessment. The assumption behind ISO 16363 is that those undertaking a self-assessment are already familiar with OAIS and its terminology. ISO 16363 explicitly states that “Because the OAIS has become a foundational document for digital preservation, the common terms are well understood and are therefore used within this document” (Consultative Committee for Space Data Systems, 2011, p.1-1). However not everyone who undertakes a self-assessment against ISO 16363 is necessarily familiar with all the details of OAIS and its terminology.

Another terminology issue encountered by Deakin relates to non-OAIS terminology used to describe operational-level workflows and concepts. Some of these terms are outdated or ambiguous. For example, ISO 16363 uses “linking/resolution services” to refer to persistent identifiers and links. However, this term is today more commonly associated with OpenURL link resolver services such as Innovative Interfaces’ WebBridge and Ex Libris’ SFX.

**Complexity**

In addition to the issue of terminology, many of the ISO 16363 criteria are not well explained, are open to different interpretations or are overly complex. Sometimes this makes them very difficult to understand for those not well-versed in OAIS (Elstrom, 2014). While most of the criteria do include some explanatory text, this text doesn’t always provide sufficient clarity. Below are just 2 examples

\(^3\) Backups not factored in.
of complex criteria which have little or no explanatory text and may be difficult for many users to grasp:

- **4.2.1** The repository shall have for each AIP or class of AIPs preserved by the repository an associated definition that is adequate for parsing the AIP and fit for long-term preservation needs.
- **4.2.7.3** The repository shall bring the Content Information of the AIP up to the required level of understandability if it fails the understandability testing.

Familiarity with OAIS and its terminology will impact on the degree to which ISO 16363 criteria are considered complex; as will their compatibility with a specific repository’s workflows and processes.

While most people who work in the preservation arena are likely to be familiar with the OAIS acronym, it is probable that many will not be fully conversant with the details of the model. This is especially so for smaller organisations. This fact has been well acknowledged in the digital preservation field, and simpler tools, such as the NDSA Levels of Preservation, have been established to help smaller organisations manage digital preservation risks (Owens, 2012b). Interest also has been expressed for an ‘OAIS lite’ version to make OAIS more accessible to smaller and less well-funded institutions (Paradigm, 2008).

**Length and repetitivity**

Another factor that contributes to the complexity of an ISO 16363 assessment is the large number of criteria, not all of which will be relevant to every organisation, and many of which overlap. The number of criteria and the duplication of many of them contribute to the complexity, time-consuming and potentially resource-intensive nature of an ISO 16363 self-assessment. An example of duplication is in relation to checksums, which is covered to varying degrees by at least 8 separate criteria4.

**Self-assessment findings**

The assessment found that DRO fully met 67 of the 105 criteria, partially met 32 and did not meet 5 criteria; with one criteria being not applicable to Deakin. In practice, different criteria will have different degrees of importance and risk levels for different repositories. ISO 16363 does not provide any risk assessment or ranking for any of its criteria.

Deakin’s compliance with ISO 16363 was most robust in the infrastructure and security risk management area, with organizational infrastructure being the area of greatest weakness. Although Deakin had significant amounts of documentation that indicated digital preservation was considered of high importance, at that time there was no formal preservation policy or preservation strategic plan. Nor were there guidelines regarding issues such as preferred file formats for most media, nor procedures for dealing with fragile digital media.

Some of the issues identified by the assessment related to known preservation-related functionality built into the Fez/Fedora software, where the functionality had either not been activated in the first place or was not working as expected. For example, versioning was available at both the metadata and object level, but had been activated only partly at the metadata level, and not at all at the object

---

4 Checksums or fixity checks were relevant, to varying degrees, to the following ISO 16363 criteria: 3.3.5, 4.1.1, 4.1.5, 4.1.6, 4.2.1, 4.2.8, 4.2.9, 4.4.1.2, 5.1.1.3.
level. Another example related to checksumming, which also had not been fully activated; and where it had been activated, there was no verification mechanism in place.

**Post-assessment progress**

Deakin has made significant progress in addressing the areas of improvement identified by the self-assessment. A major achievement was the establishment of a Research Data Store linked to DRO. This provides a space to preserve research data and accommodate AIPs for items that are unable to be fully accommodated in DRO due to their large file size or number of associated files. Fixity checks have also been implemented in both DRO and the Research Data Store.

A formal preservation policy for DRO has been created, although not yet officially adopted, and a preservation strategic plan established. DRO file naming conventions have been reviewed and guidelines established regarding acceptable file formats. Additional preservation-related tools have been incorporated into workflows, such as Bagger/BagIt. A suite of regular integrity checks has also been established, based on some of the issues identified by the assessment.

There is also a greater awareness within Deakin of the potential preservation-related implications of new projects or tools. This has resulted in new integrity checks being set up to identify problems with the tools. For example the introduction of support for NISO RP-22 triggered the establishment of a new set of integrity checks to identify contradictory content in NISO RP-22 fields.

Some of the issues identified by Deakin’s assessment require ongoing attention; for example, new file formats need to be reviewed as they come into DRO. However, most of the high priority areas of improvement identified in the self-assessment have been addressed. Progress continues to be made. Indeed, digital preservation is an ongoing task that will never actually be finished; areas of improvement can always be identified.

The self-assessment has also inspired preservation-related improvements to other Deakin systems. For example, risk assessments have been undertaken of the different media stored in the Library’s Special Collection, and preservation strategies established for openjournals@Deakin and Deakin’s exhibition platform, Fusion.

Some of the issues identified during the self assessment have turned out to be more complicated than originally envisioned, so quick decisions to address these have not been feasible. For example, while metadata versioning has been activated, object versioning is a much more complicated issue due to the large amount of extra space required. A manual versioning procedure for objects is already in place (e.g. post prints, pre-prints etc), but DRO has not as yet been configured to turn on object versioning.

**Lessons learnt/Recommendations**

Looking back at Deakin’s self-assessment with the benefit of passing time and experience gained, there are several things Deakin would choose to do differently next time. Some of these things relate to the overall management and organisation of the self-assessment; while some relate to the assessment itself. The lessons learnt from the self-assessment, and recommendations for others considering a similar journey include:
• **Regular re-assessments.** It is highly recommended that repository managers undertake regular self-assessments of their repository to ensure the repository continues to remain robust. While the frequency is something best left to the judgement of repository management, given the fact that both OAIS and ISO 16363 have an update schedule of 5-year intervals, a maximum of 5 year intervals for re-assessment is reasonable. Obviously, any external audits or accreditations, as well as resource availability, will influence decisions in this area. Additionally, re-assessments don’t necessarily have to be of the same intensity each time around.

• **Choose the best tool for your circumstances.** Not all repository managers will have the time, resources or staff expertise to undertake a self-assessment, particularly against ISO 16363. The fact is that ISO 16363 is a long, detailed and complex standard. It requires specific skills and knowledge, not only of repository workflows and processes, but also of the standard itself. It also requires, at the very least, an understanding of OAIS’s conceptual nature and a familiarity with OAIS terminology. You can, however, choose to adapt it to suit your circumstances and available resources.

Alternative, simpler, tools to ISO 16363 do exist, such as NDSA Levels of Digital Preservation (Owens, 2012b), DRAMBORA (Digital Curation Centre, 2015) and Data Seal of Approval (Data Seal of Approval Board, 2016). So if ISO 16363 appears too complex and overwhelming, consideration should be given to utilising a different tool. The Digital Preservation Coalition (2015) lists criteria to consider when choosing a self-assessment tool.

There are also ISO 16363-specific tools available to assist with self-assessments (Archivmatica, 2016; Primary Trustworthy Digital Repository Authorisation Body Ltd, 2017).

• **Do a self-assessment before an external audit.** Repository managers considering an external audit or certification of their repository should ideally undertake a self-assessment beforehand, even if only a quick and superficial one. This is because external audits are expensive, and receiving a low accreditation is not good for a repository’s reputation. At the very least, a preliminary self-assessment should be undertaken, and the most critical issues addressed before an external audit or certification process is initiated. This also reduces the risk and pressure of the repository having to undergo another external audit or certification process within a short time, and thus is more cost-efficient in the long run.

• **Determine the depth and focus of the assessment.** The depth of any ISO 16363 self-assessment should be determined at the outset, as this will impact on the allocation of resources. A standard self-assessment will involve the collection and review of relevant documentation for each criteria, and a determination as to whether a repository meets those criteria. However, repository managers may choose to focus the self-assessment on areas of highest risk or known vulnerabilities, and may also choose to structure the assessment to fit available resources and staff expertise. They can also choose to replicate an external audit, and collect evidence to assess whether workflows and internal controls.

---

5 Both OAIS and ISO 16363 are due to be updated late 2017.
are actually being complied with, and are working as they should.

- **Identify relevant document at the start.** When reviewing the criteria, potentially relevant documentation should be identified up front. While it is probable that some relevant documentation will be identified and need to be retrieved mid-assessment, it will save time and effort if as much of the relevant documentation as possible is identified and retrieved at the start. This is especially so given it is inevitable that some documents will be relevant to multiple criteria. Keep in mind, also, that ISO 16363’s suggested documentation are intended as guidelines only and won’t necessarily correspond to your own organisation’s documentation.

- **Become familiar with relevant policies and procedures.** Ideally, the person in charge of the self-assessment will be familiar with the organisation’s and the repository’s policies and procedures that relate to the assessment. This will also, hopefully, include knowledge of relevant documentation. This will make it quicker and easier to locate documentation as evidence for the repository meeting ISO 16363’s criteria.

- **Review the criteria before you start.** A preliminary review should be undertaken of the ISO 16363 criteria before the assessment itself begins. The review should ensure that the criteria are fully understood; this will save time and help to avoid misunderstandings during the actual assessment. It will also help with the process of identifying relevant supporting documentation for the criteria.

  Ideally, the review should also assess the risk probability and impact of each of the ISO 16363 criteria. Time and possible duplication of effort can also be saved by grouping similar criteria together, and weighing the relevance of each criteria or criteria group to the repository. Not all criteria are equally as important or relevant to a particular repository, so identifying the criteria or criteria group at greatest risk or of higher relevance makes it easier for the assessor to focus resources on those criteria.

  It is also a good idea to flag potentially problematic criteria at the start of a self-assessment, and keep them in mind as the assessment proceeds. This increases the likelihood that the criteria may be explained in the context of other criteria or a broader view of the repository workflows.

- **Allocate sufficient resources.** Keep in mind that ISO 16363 self-assessments are resource-intensive; although the extent to which this applies will depend very much upon the depth and focus of the assessment. Also, resources will be needed not just for the actual assessment itself, but also to implement any recommendations. The amount of resources that may be needed post-assessment can be difficult to assess, but it should be recognised that a self-assessment is not over once the final report of findings has been written.

**Conclusion**
Deakin’s self-assessment provided it with a greater understanding and knowledge of the inner workings of its repository software and the importance of various repository workflows. It enabled
areas of weakness to be identified, and provided the opportunity to strengthen those weaknesses, thus improving the overall security and trustworthiness of the repository. Some weaknesses were already known, and for these, the self-assessment provided Deakin with the incentive to address them sooner rather than later. The 2013 self-assessment provides a reference point for Deakin’s next assessment, which is likely to be more streamlined and cost-effective as a result of the lessons learnt.

The self-assessment also highlighted some of the interdependencies between OAIS and ISO 16363. In particular, Deakin found that at least a basic understanding of OAIS and its terminology is needed to undertake a self-assessment against ISO 16363; with a more thorough understanding required for intensive self-assessments. Most critically, the conceptual nature of OAIS needs to be fully understood. The familiarity gained with OAIS and ISO 16363 is a valuable asset, and has lead to further improvements in digital preservation within the Library as a whole.

The assessment improved staff skills and has resulted in a more robust repository, with a stronger underlying infrastructure. Deakin still has more work to do in this area; but good digital preservation practice involves constantly striving to improve the robustness of a repository and checking that everything is working as expected. Deakin now has a clearer vision of what to do in the future.

**Bibliography**


Biography

Bernadette Houghton is the Digital Preservation Librarian at Deakin University. Her work involves preservation of digital materials held in Deakin’s Special Collections and research repository, DRO. She was involved in the establishment and ongoing support of DRO and has a strong background in systems librarianship. She also has 4 years of experience as an internal auditor with the former State Electricity Commission of Victoria.