



JISC MDC Project Plan

Overview of MDC Project

1. Background

A recent inventory study of OAI-compliant facilities indicates that less than 10% of content is primary data ([www.driver-support.eu/documents/DRIVER Inventory%20study%202007.pdf](http://www.driver-support.eu/documents/DRIVER%20Inventory%20study%202007.pdf)). This represents a poor utilization of the research investment (a) because the personnel and equipment costs allocated to performing experiments represents a large proportion of the total budget of many projects and (b) the data almost certainly has value beyond the scope of the project in which it is generated. Efforts to promote data capture and preservation will make experimental data available that could for example be used (a) to validate and develop new models, (b) as input into knowledge discovery systems, (c) for the purposes of auditing, etc.

The engineering materials sector is characterised by a wide diversity of facilities and data. This perhaps explains why solutions to data curation and systems interoperability issues are in their infancy compared to other domains, such as crystallography, astronomy, and the biological sciences. Where other domains can rely on tailored solutions to federate distributed facilities that find acceptance through a standardization process that can be approved by a relatively small group of stakeholder (such as OAI-PMH in the case of crystallography), the engineering materials community has to accommodate and conform to procedures that require longer and more involved consultative processes. Such processes are already underway, as evidenced by the NIST Toward Ontology-based Standards initiative (<http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2009>), the W3C Product Modelling Incubator (www.w3.org/2005/Incubator/w3pm), the CODATA Materials Data Task Group (<http://www.codata.org/taskgroups/TGmatlsdata/index.html>), and the EC/CEN Economics and Logistics of Standards-compliant Schemas and Ontologies for Interoperability - Engineering Materials Data initiative (<http://www.cen.eu/cenorm/businessdomains/businessdomains/iss/activity/ws-elssi.asp>).

2. Aims and Objectives

The project aims to mobilise and develop the data capture and curation resources in the engineering materials domain to establish a Materials Data Centre to serve the materials community. An OAI-compliant Materials Data Centre will promote data capture and systems interoperability in this domain. The work will extend existing DDLs in the engineering sector to incorporate standards-compliant schemas and ontologies. The work will also enhance the fractography repository at <http://www.fract.ses.soton.ac.uk>. The work will be undertaken as a

PhD project, with the University of Southampton providing the matching funds during the second half of the PhD (as approved by the JISC policy department). The workplan addresses three key issues:

- The establishment of a repository to manage materials test data
- The formulation of schemas and ontologies that can be utilized to address systems interoperability
- The development of quality assurance and access management processes

Detailed Objectives:

- To deliver an extensible Web-enabled materials data centre that is structured according to the existing MatML schema for materials property data.
- To integrate new standard-compliant schemas
- To formulate an empirical quality measure based on the standards-compliant XSDs, including the extent, quality, and importance of the parameters featured in the XSDs.
- To formulate a data access management that promotes Open Access but that accommodates data ownership and confidentiality issues. Where possible, due attention will be given to integrating with data access management mechanisms promoted by the JISC e-infrastructure programme, such as promoted by the JISC Core Middleware initiative
- To develop standards-compliant test schemas for the most common materials tests.
- To extend the standards-compliant schemas to equivalent ontologies and rulesets.
- Review existing formats for micrographs and explore their application to fractographic data
- Examine the options for integrating the Fractography Online Resource and Archive repository at <http://www.fract.ses.soton.ac.uk/> with the Materials Data Centre
- To integrate ontologies and rulesets and the quality measure and the data access models

3. Overall Approach

The solution that the MDC project proposes to the issue of interoperability and data capture in the engineering materials sector is to leverage existing procedural standards in a discipline where there is the requirement for the schemas, ontologies, and rulesets. In combination with an existing schema for materials properties (MatML) the test specific schemas will be employed to deliver a data store, which combined with a Web interface, will form the Materials Data Centre.

XML Schema is a mature technology that provides an immediate solution to the capture and curation of materials data. It also addresses the 'integrity constraint' issue (as described in 'Opening, Closing Worlds — On Integrity Constraints', E Serin et al, OWLED 2008), which is presently a limitation of the OWL 2.0 ontology language specification. Schemas

development will also establish viable hierarchies and structures that are expected to contribute the development of ontologies and rulesets.

Unlike many other science and engineering domains, the problems present by the curation of materials data are not related to volume of data, but to the extent of the metadata required to give meaning to the results beyond the context in which they were generated. It is extremely important that a materials data set is comprehensive with respect to the various parameters that are recorded. It is also clearly important that the data are of an inherently high quality. To address these issues, the MDC project will incorporate an effort to develop an empirical quality measure that accommodates these and other characteristics of materials data. Similarly, the MDC project will develop a data access workflow that balances the need for Open Access with data ownership. It is only reasonable that while a research group has an obligation to share their data, that they are given sufficient opportunity to publish and exploit their results. It should also be expected that where there is an interest to access data generated by other research groups, that there should be a corresponding willingness to share data. These and other data access management issues will be explored and developed by the MDC project.

The work will be undertaken as a PhD project, with the University of Southampton providing the matching funds during the second half of the PhD (as approved by the JISC policy department). The basic effort to establish the Materials Data Centre and extend its coverage to as many test types as possible has been planned for the first 18 months i.e. the formal term of the JISC project. Research into ontologies, rulesets, and new quality assurance and access models is scheduled for months 19 through 36 i.e. the term of the PhD project.

The support of the University of Southampton Microsoft HPC Institute offers the opportunity to develop a sophisticated and scalable data centre using 'Sharepoint Portal', which will itself be hosted by iSolutions (the University of Southampton computing and IT service). This high end application suite has already proved effective in establishing the new FLUXNET repository at <http://www.fluxdata.org/default.aspx>, which itself federates over 960 site-years of data from over 253 eddy covariance tower measurement sites for exchanges of carbon dioxide, water vapour, and energy between terrestrial ecosystem and atmosphere.

Critical measures of success will include the following:

- **Adoption by the SES Engineering Materials Group**—the group that develops the facility must be convinced of its value if there is to be any chance that the data centre is to find widespread adoption.
- **Development of a reliable quality measure**—without an indication of quality, data are unlikely to be reused because the third party (the data reuser) can have no confidence in data pedigree or provenance.
- **Sustained accumulation of experimental data**—data must regularly be posted (and retrieved).

4. Project Outputs

In an effort to promote the capture and curation of experimental data in a manner that is consistent with materials community initiatives, the proposed MDC project aims to establish an OAI-compliant repository for materials test data. This facility will be available to the materials research community, allowing the very significant resources invested in materials testing and analysis to be preserved for future research and development.

Besides the repository, the project will deliver the following:

- A best practices guide to the development of a materials data centre.
- A collection of standards-compliant schemas, ontologies, and rulesets for materials tests.
- An automated, schema-based quality measure.
- A data access model that promotes Open Access but accommodates the need for confidentiality where appropriate.

Of the less tangible outputs, the project will promote a culture of data capture and preservation in the R&D sector by demonstrating:

- How Web technologies can reduce the overhead of data capture.
- The benefits to researchers.

Since the work will be undertaken as a PhD project, with the University of Southampton providing the matching funds during the second half of the PhD (as approved by the JISC policy department), the work will be undertaken so that the Materials Data Centre will be fully operational by month 18 i.e. the formal term of the JISC project.

5. Project Outcomes

The project will deliver a facility to conserve and expose materials test data. The real value will become apparent as semantic technologies become established, at which stage the ever-increasing body of experimental data will provide new opportunities for knowledge discovery within and across disciplines.

6. Stakeholder Analysis

Several differing, but overlapping stakeholder groups are identified in the following table, and their needs and the approach to engaging each group further outlined below:

| Stakeholder | Interest / stake | Importance |
|------------------------------|-----------------------------------------------------------------------------------------|------------|
| Materials science researcher | A central location to store and manage project data. | High |
| Materials research groups | Data sharing provides a mechanism to promote their work and interact with other groups. | High |

| | | |
|--------------------------------------|---------------------------------------------------------------------------------|--------|
| Materials science data reuser | Easy access to data with of a predefined quality. | Medium |
| Data manager | Ability to manage and preserve data in a structured standards-compliant format. | Medium |
| Funder | Preservation of the research investment and a guaranteed auditing trail. | High |
| Systems developers | Data made available in a structured standard-compliant format | Medium |

Materials science researcher—a central location to store and manage project data. Although based at Southampton, the repository is intended be an open facility, designed to develop a culture of responsible data management and reuse, and researchers are expected to be primary beneficiaries. Not only will the preservation of annotated data be useful to the implementation of projects in which they were generated, but there is clearly the potential for individual researchers to benefit from the use of their data by external researchers (as named authors on publications based on their results, for example).

Materials research groups—data sharing provides a mechanism to promote their work and interact with other groups. Presently it is the case that procedural testing standards identify data to be recorded, but in the absence of corresponding formats, the data are rarely conserved, and their value diminishes as the material pedigree, test conditions and results become disassociated. The efforts to annotate data sets according to recognised procedural testing standards will result in a more nuanced and comprehensive set of metadata around the materials data. This will promote both the long-term value of experimental data and its relevance for research efforts beyond the project in which the data were produced, thereby ensuring not only that follow-on research is able to fully exploit previous work but also that data have meaning for external research groups not involved in producing the data. The opportunity for research groups to share data of high inherent worth and quality has the potential for a profound impact on the research process, not only making data available for the verification and development of new materials models, but promoting the exchange of knowledge, closer cooperation with other research groups and industry, greater motivation to publish and exploit data, and cross-referencing of publications and data

Materials science data reuser—easy access to data of a predefined quality. The arguments for the benefits to the data reuser extend those put forward for materials research groups to the broader materials community. The industrial community at large will generally value freer exchange of directly comparable/well-referenced materials data in standard formats (subject to the confidentiality restrictions noted for “Funders”) as this will directly enable lifing methodologies, structural assessments etc. to be tested and common reporting formats agreed. The wider academic materials community will also always value the availability of detailed well-evidenced materials data, not only for comparison and modelling purposes, but also auditing.

Data manager—ability to manage and preserve data in a structured standards-compliant format. The infrastructure that the MDC will deliver and its compliance procedural standards will promote the effective administration of the results of materials research programmes.

Funder—preservation of the research investment and a guaranteed auditing trail. The R&D community is coming under increasing pressure to demonstrate that curation and preservation of the research investment are an integral part of the research process. Funding bodies, and government agencies (both national and transnational) are quite reasonably beginning to demand a more responsible approach to curation and preservation of the data generated from the projects they fund. The MDC will provide the materials community a facility that contributes to the repository infrastructure that is needed if these demands are to be met. Direct industrial sponsors of materials research also have particular confidentiality issues about the identification of the materials used (e.g. Compositional or heat treatment information) or even the actual materials property data itself, but will want to ensure data integrity, useability and traceability internally within their organisation. The advantages then to them will be not only secure/useable data transmission from their university collaborations, or other external testing contractors, but also implementation approaches to collating and using their in-house data (although it should be noted that the amount of internal industrial materials testing capability within the UK is dramatically diminishing, hence data transmission and traceability issues are becoming more and more important)

Systems developers—data made available in a structured standard-compliant format. As various branches of the sciences strive towards a Semantic Web of data, identifying robust schemas, ontologies, and rulesets for presenting materials data is certainly a valid research domain in its own right. As the volume of online data increases, researchers and systems developers will be able to contribute to the development common tools and methodologies for the management and manipulation of online data. The opportunities for the sciences are very exciting, and there is great potential for new and improved research, and the MDC will represent a significant contribution from the materials community.

These can be seen as differing, but overlapping stakeholder groups, and the approaches to interacting with, and building up links between them will vary. For each stakeholder group therefore a more detailed requirements analysis will be worked up at the beginning of the project and delivery against this assessed as the project progresses. Representatives of each stakeholder group can and will be identified via existing contacts within the supervisory team. Southampton based workshops will be one way of raising the profile of the research, and should be able to engage across the stakeholder groups. We are well connected into various existing stakeholder groups via our industrial sponsorship and formal technical partnerships at Southampton, membership of various academic networks (e.g. UKCME, HEA engineering subject centre), professional societies (e.g. Institute of Materials, Minerals and Mining), and the 12-month ELSSI-EMD Workshop in which Dr Austin involved. As the primary intention of CEN Workshops is to engage a broad and representative group of stakeholders in a standards related development, there is significant scope for synergy with the proposed MDC. As an indication of the potential impact of the MDC, the ELSSI-EMD

proposal submitted by Dr Austin was supported originally by Alstom Power Systems, Elsevier, and the Digital Curation Centre, and invitations to participate in the year-long activity have been issued to upwards of 150 organizations, representing the research, business, and standardization communities. Dr Austin will be fully committed to ensuring that the two projects are aligned and complementary, and that MDC gets the maximum possible exposure amongst ELSSI-EMD participants. A direct link to the JISC data programme (with specific focus on research data) led by Simon Hodson will be beneficial and attendance by the PhD student (and supervisory team) at programme meetings is planned.

7. Risk Analysis

| Risk | Probability (1-5) | Severity (1-5) | Score (P x S) | Action to Prevent/Manage Risk |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Staffing —although the topic is very relevant to current research interests in the engineering and curation sectors, recruitment of PhD student is by 1st April is unlikely. | 5 | 5 | 25 | Delay project start, utilize internships where possible |
| Organisational —visiting status of the Principal Investigator (Tim Austin) means that the PI is unable to meet their supervisory and managerial duties | 2 | 2 | 4 | The role will be reassigned to another of the equally well-qualified individuals who are supporting the proposal. |
| Technical —procedural standards are not amenable to conversion to schemas and ontologies | 2 | 3 | 6 | The viability of standards-compliant schemas and ontologies is the subject of the EU ELSSI project. If there are issues that preclude against the development of standards-compliant schemas and ontologies, the ELSSI partnership is expected to propose alternatives, and the work of the MDC project will be aligned accordingly. |
| External suppliers | | | | N/A |
| Legal | | | | N/A |

8. Standards

| Name of standard or specification | Version | Notes |
|-----------------------------------|---------|-------------------------|
| XML | 1.0 | Markup-up for data sets |

| | | |
|----------------|-----|------------------------------------------|
| XML Namespaces | 1.0 | Aggregation of data definition languages |
| XSLT | 2.0 | Transformation of data sets |
| XML Schema | 1.0 | Data model definition language |

The listed specifications are those preferred at the time of writing. However, community preferences as expressed through the ELSSI-EMD project will have to be taken into consideration, and there is the possibility that proprietary standards may be preferred (such as ISO 10303).

9. Technical Development

Technical development will be undertaken according to industry best practices—adhering to an agile, iterative process. This is reflected in the workplan, where work packages deliver loosely coupled but highly cohesive components that can be developed, tested, and delivered independently. The project work packages will be ported to stories and tasks in XPlanner (a project management tool designed for the agile process). Defect tracking will be handled using Bugzilla. An infrastructure that supports team development and version control will be realized using VCS and Sharepoint.

10. Intellectual Property Rights

The model(s) for protecting intellectual property will be written into the consortium agreement before the project starts. The consortium agreement will include clauses that address the following:

- Detailed specification and descriptions of all background IP already developed and therefore owned by the University of Southampton.
- Foreground IP emerging from the project will be identified and the ownership of the IP agreed. Typically, foreground IP is freely available to all parties to the project and for a predetermined period from the conclusion of the project.
- Guidelines on the commercial exploitation of the IP generated by the project and the benefactors thereof.

Project Resources

11. Project Partners

None

12. Project Management

The work will be performed as a PhD project. The development of the Materials Data Centre will thus be subject to the scrutiny of the PhD supervisor(s), who will be senior researchers, well acquainted with project management procedures and corrective measures required to keep the project on track. The work will be under constant review by the PhD

supervisor(s). The work will be formerly assessed at quarterly face-to-face meetings, and issues addressed accordingly.

Dr Tim Austin (male) is a visiting researcher at the University of Southampton School of Engineering Sciences (SES) and the Principal Investigator for the proposed MDC project. He anticipates allocating 20% of his time to project management, including PhD supervision.

In 2008 he led a partnership including Alstom Power, Elsevier and the DCC to submit a successful bid in response to the EC DG-Enterprise 'ICT Standardization' call for proposals. The project, entitled 'Economics and Logistics of Standards-compliant Schemas and Ontologies for Interoperability - Engineering Materials Data (ELSSI-EMD)' (and described in more detail at

<http://www.cen.eu/cenorm/businessdomains/businessdomains/iss/activity/ws-elssi.asp>) is a CEN Workshop that provides a forum for stakeholders in the engineering and standardization sectors to reach consensus on a solution to improved systems interoperability and data curation. ELSSI-EMD aims to determine (a) whether standard formats can provide the research community a solution to data conservation and systems interoperability, (b) whether there is a business case to justify the development of standards-compliant schemas and ontologies, and (c) the implications of standards-compliant schemas and ontologies for their corresponding procedural standards. Through his role in the ELSSI-EMD project, Dr Austin will provide valuable insight into the preferences expressed by the engineering and standardization communities in respect of interoperability and curation, and he will ensure that the ELSSI-EMD project and the proposed JISC project remain complementary to one another and fully informed of one another's activities.

Professor Philippa Reed (female) is Head of the Materials Research Group and Director of Graduate School in SES, she has extensive experience of materials testing (her research interests centre on failure analysis, fatigue testing and micromechanistic evaluations in a range of engineering alloys) and has recent and ongoing collaborations with Rolls Royce, QinetiQ, Mahle, Luxfer Gas Cylinders, Airbus, DSTL, Alstom and E.On Engineering (UK). She therefore has a large research group (currently 8 PhD students) who will act as a primary data source and user base for the proposed Materials Data Centre.

Dr Kenji Takeda (male) is a senior lecturer in the Aerodynamics and Fluid mechanics research group in SES. He has extensive experience in high performance computing and in the development of commodity supercomputing systems and applications, and Grid and Web Service computing. Current research includes new IT infrastructures for handling and analysing large amounts of disparate data. Of note is the development of a grid-based data management and analysis system for wind tunnel data processing. He will therefore be the lead academic supervisor for the PhD programme. The Microsoft HPC Institute at the University of Southampton, led by Simon Cox and Kenji Takeda, is working to demonstrate why, where, and how we can exploit current and future tools and technologies to make the engineering design process faster, cheaper and better. A current activity is as the lead university for the IT workstream of the £17.4m Centre for Fluid Mechanics Simulation cross-industry project led by Airbus. As part of this the Institute has developed data management solutions for Airbus, and workflow and knowledge management technology demonstrators

for BAE SYSTEMS and Rolls-Royce. The proposed Materials Data Centre is an ideal project for the Microsoft HPC Institute to facilitate. Postdoctoral researchers from the HPC institute will be able to support the PhD student on system delivery and hence several weeks of staff time will be available.

13. Programme Support

Where necessary, evaluation of technical decisions from the perspective of complementary JISC initiatives would be appreciated. For example, mechanisms to guard against replication of data sets are likely to have been developed in the context of other projects, and while the project team will make every effort to , verification and guidance from the project manager will ensure that MDC is aligned with and complementary to JISC interests.

14. Budget

Note—please refer to Appendix A for a budget that adheres to the JISC template.

For the benefit of the project, it has been decided that a small proportion of the funds would be better allocated to address dissemination and project management objectives.

Detailed Project Planning

15. Workpackages

Note—please refer to Appendix B for a workplan that adheres to the JISC template.

WP1: Data Center Phase 1 — Deployment

Objectives:

- To deliver an extensible Web-enabled materials data centre that is structured according to the existing MatML schema for materials property data.
- To integrate new standards-compliant schemas

Activities:

A1.1. Configure facility (M01-M03)

A1.2. Deploy MatML-compliant data store (M01-M06)

A1.3. Develop data entry and data retrieval application (M07-M12)

A1.4. Add support for tensile, fatigue, fatigue crack growth, and impact test data (M13-M18)

A1.5. Draft a best practices guide on designing and developing a materials data repository (M15-M18).

WP2: Quality Assurance

Objectives:

- To formulate an empirical the quality measure based on the standards-compliant XSDs, including the extent, quality, and importance of the parameters featured in the XSDs.

Activities:

A2.1. Develop quality measure (M13-M24)

WP3: Data Access Management

Objectives:

- To formulate a data access management that promotes Open Access but that accommodates data ownership and confidentiality issues. Where possible, due attention will be given to integrating with data access management mechanisms promoted by the JISC e-infrastructure programme, such as promoted by the JISC Core Middleware initiative.

Activities:

A3.1. Develop data access model (M13-M24)

WP4: Materials test data DDLs

Objectives:

- To develop schemas and rulesets for the most common materials tests.

Activities:

A4.1. Schema and ruleset for tensile testing at ambient temperature (M04-M06)

A4.2. Schema and ruleset for tensile testing at elevated temperature (M07-M09)

A4.3. Schema and ruleset for fatigue testing (M10-M12)

A4.4. Schema and ruleset for fatigue crack growth testing (M13-M15)

A4.5. Schema and ruleset for impact testing at ambient temperature (M16-M18)

WP5: Materials test data ontologies

Objectives:

- To extend the schemas for common materials tests to equivalent ontologies.

Activities:

A5.1. Standards-compliant ontology for tensile testing at ambient temperature (M19-M21)

A5.2. Standards-compliant ontology for tensile testing at elevated temperature (M22-M24)

A5.3. Standards-compliant ontology for fatigue testing (M25-M27)

A5.4. Standards-compliant ontology for fatigue crack growth testing (M28-M30)

A5.5. Standards-compliant ontology for impact testing at ambient temperature (M31-M33)

WP6: Dissemination and sustainability

Objectives:

- To disseminate knowledge and deliverables, and to promote the long-term viability of the data centre.

Activities:

A6.1. Deploy project web site and blog (M01)

A6.2. JISC rapid innovation proposal for enhanced publishing process (M01)

A6.3. EC Enterprise and Industry DG 2009 ICT Standardisation Work Programme proposals for data collection prototypes (M01-M04)

A6.4. Contribute documentation to the IE Demonstrator project and post regularly to the IE Demonstrator blog (M03-M18)

WP7: Management and evaluation

Objectives:

- Meet both the terms and conditions of the JISC grant and the academic excellence demanded of a PhD project

Activities:

- A7.1.** Project plan (M01-M02)
- A7.2.** Evaluation plan (M01-M02)
- A7.3.** JISC reporting (M01-M18)
- A7.4.** Recruitment—internships (M01-M02, M13-M14, M25-M26)
- A7.5.** Recruitment—studentship (M01-M03)
- A7.6.** Supervision and mentoring (M02-M36)
- A7.7.** Scientific publications (M12-M36)
- A7.8.** PhD thesis (M34-M36)

16. Evaluation Plan

Means to measure the impact of the project, as follows:

- The number of test specific DDLs and ontologies that are developed.
- The number of data sets a user conserved prior to and after registration to use the facility.
- The publication rate before and after the introduction of the Materials Data Centre. The expectation here is that the data access model will likely promote a policy of limited period restricted access to data from publicly funded projects (exactly as implemented by FLUXNET at <http://www.fluxdata.org/default.aspx>).
- The number of publications that acknowledge use of content at the Materials Data Centre.

| Timing | Factor to Evaluate | Questions to Address | Method(s) | Measure of Success |
|--------|--------------------|----------------------|-----------|--------------------|
| | | | | |

17. Quality Plan

Quality assurance practices will rely on utilization of recognized practices and tooling for managing software development projects and consultation with subject matter experts for alignment with JISC best practices (Project Manager) and code review (MSHPCI research team members).

| Output | Quality criteria | QA method(s) | Evidence of compliance | Quality responsibilities | Quality tools (if applicable) |
|---------|------------------|-----------------|------------------------|--------------------------|-------------------------------|
| M01-M36 | Code | Defect tracking | Tooling available | | Bugzilla |

| | | | | | |
|----------------|-----------------|-------------------------------------------|----------------------------------------------------|--|-----------------|
| | | | and relevant data entered | | |
| M01-M36 | Code | Team development | Tooling available and relevant data entered | | SVN |
| M01-M36 | Planning | Change control development process | Documented change controls | | |
| M03-M36 | Code | Software development process | Tooling available and workplan entered | | XPlanner |

18. Dissemination Plan

An MDC project Web site (compliant with the JISC Information Environment requirements posted at <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/>) will be maintained together with the Materials Data Centre. It will provide access to the Materials Data Centre, provide background information, resources, and details of news and events relevant to users of the facility. The Engineering Materials Group, and the School of Engineering Sciences of which it is a part, have excellent industrial and academic contacts. The availability of the Materials Data Centre will be made known to all contacts, and the use of the facility to serve their needs and those of the projects in which they are involved will be actively promoted. Through his involvement in the ELSSI project, Dr Austin will ensure that news of the facility diffuses into the wider community, not only promoting use of the Materials Data Centre by remote parties with no direct affiliation to the University of Southampton, but also promoting similar initiatives at other institutions. To facilitate the development of a network of repositories that have the potential for federation, lessons learnt during the design and development phases of the MDC project will be published as a best practice guide to materials data centre development.

The IE Demonstrator project will be used as the primary means for promoting dissemination of MDC know-how and deliverables to the JISC community. There will be a commitment to register design and development documentation to IE Demonstrator project. IE Demonstrator aims to capture and share both the eventual output of projects contributing to the repository landscape as well as interim outputs—especially those relating to the design/development process, such as deployed or packaged prototypes, screen shots of same, designs/paper-prototypes, documentation relating to the development process (e.g. evidence of user-testing, requirements etc.), and a roadmap for future development. The IE Demonstrator wiki holds information about developments in progress, and MDC will contribute accordingly

| Timing | Dissemination Activity | Audience | Purpose | Key Message |
|----------------|-------------------------------|-----------------|----------------------------|-------------------------------------------------------------|
| M01 | Project website | All | Portal for the data centre | Publicize the project and provide access to the data centre |
| M01-M36 | IE Demonstrator input | JISC community | Disseminate | Technologies |

| | | | | |
|--|--|--|----------------------------------|-------------------------------------------|
| | | | data centre development know-how | and processes employed in the MDC project |
|--|--|--|----------------------------------|-------------------------------------------|

19. Exit and Sustainability Plans

The project delivers a data centre that is expected to be sustainable for the long-term, and there is no expectation for the need for exit plans.

| Project Outputs | Why Sustainable | Scenarios for Taking Forward | Issues to Address |
|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| The Materials Data Centre | An integral part of computing infrastructure at the School of Engineering Sciences, ensuring that the significant investments in materials testing are preserved for future research | The School will be able to satisfy the requirements of any funding agency that mandates a data curation strategy. There will also be the added responsibility of continuing the service for the materials community at large. | Data collection mechanisms, which are planned to be addressed by collaborations with stakeholders in the publishing and data generations sectors. |

Appendixes

Appendix A. Project Budget

| | | | | |
|--------------------------------------------------------------------------------------------|-------------------------|--------------------------|--------------------------------------|-----------------------|
| Directly Incurred Staff | Apr09– Mar10 | Apr10 – Mar11 | Apr11 – Mar12¹ | TOTAL £ |
| PhD Student, 40 h/w | £ 13 250 | £ 13 550 | £ 13 850 | £ 40 650 |
| Etc. | £ | £ | £ | £ |
| Total Directly Incurred Staff (A) | £ 13 250 | £ 13 550 | £ 13 850 | £ 40 650 |
| Non-Staff | Apr09– Mar10 | Apr10 – Mar11 | Apr11 – Mar12 | TOTAL £ |
| Travel and expenses | £ 1 000 | £ 1 000 | £ 1 000 | £ 3 000 |
| Hardware/software | £ 1 000 | £ 1 000 | £ 500 | £ 3 000 |
| Dissemination | £ | £ | £ | £ |
| Evaluation | £ | £ | £ | £ |
| Other (tuition and bench fees) | £ 3 965 | £ 5 092 | £ 4 820 | £ 13 923 |
| Total Directly Incurred Non-Staff (B) | £ 5 965 | £ 7 092 | £ 6 320 | £ 19 923 |
| Directly Incurred Total (C) (A+B=C) | £ 19 215 | £ 20 188 | £ 20 170 | £ 60 573 |
| Directly Allocated | Apr09– Mar10 | Apr10 – Mar11 | Apr11 – Mar12 | TOTAL £ |
| Staff | £ | £ | £ | £ |
| Estates | £ | £ | £ | £ |
| Other | £ | £ | £ | £ |
| Directly Allocated Total (D) | £0 | £0 | £0 | £0 |
| Indirect Costs (E) | £0 | £0 | £0 | £0 |
| Total Project Cost (C+D+E) | £ 19 215 | £ 20 642 | £ 20 170 | £ 60 027 |
| Amount Requested from JISC | £ 19 215 | £ 10 321 | £ 0 | £ 29 536 |
| Institutional Contributions | £ 0 | £ 10 321 | £ 20 170 | £ 30 491 |
| Percentage Contributions over the life of the project | Partners 0 % | | | Total 100% |
| No. FTEs used to calculate indirect and estates charges, and staff included | No FTEs | Which Staff | | |

¹ Not compulsory for projects of shorter duration.

Appendix B. Workpackages

| WORKPACKAGES | Month | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|------------------------------------------------|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|
| 1: Deployment | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| 2: Quality Assurance | | | | | | | | | | | | | | X | X | X | X | X | X |
| 3: Data Access Management | | | | | | | | | | | | | | X | X | X | X | X | X |
| 4: Materials test data DLLs | | | | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| 5: Materials test data ontologies and rulesets | | | | | | | | | | | | | | | | | | | |
| 6: Dissemination and sustainability | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |
| 7: Management and evaluation | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |

| WORKPACKAGES | Month | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 |
|------------------------------------------------|-------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1: Deployment | | | | | | | | | | | | | | | | | | | |
| 2: Quality Assurance | | X | X | X | X | X | X | | | | | | | | | | | | |
| 3: Data Access Management | | X | X | X | X | X | X | | | | | | | | | | | | |
| 4: Materials test data DLLs | | | | | | | | | | | | | | | | | | | |
| 5: Materials test data ontologies and rulesets | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | | | |
| 6: Dissemination and sustainability | | | | | | | | | | | | | | | | | | | |
| 7: Management and evaluation | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X | X |

Project start date: 1st April 2009

Project completion date: 31st March 2012

Duration: 36 months

| | Start | End | Outputs | Milestone | Responsibility |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|---------|-----------|----------------|
| Materials Data Centre | | | | | |
| WORKPACKAGE 1: Data Center Phase 1 – Deployment Objective: To deliver an extensible Web-enabled materials data centre that is structured according to the existing MatML schema for materials property data and integrate test-type specific schemas. | M01 | M18 | | | |
| A1.1 Configure facility | M01 | M03 | | | |
| A1.2 Deploy MatML-compliant data store | M01 | M06 | | | |
| A1.3 Develop data entry and data retrieval | M07 | M12 | | | |
| A1.4 Add support for test types | M13 | M18 | | | |
| A1.5 Best practices guide | M16 | M18 | | | |
| | | | | | |
| WORKPACKAGE 2: Quality assurance Objective: To formulate an empirical quality measure based on the standards-compliant XSDs, including the extent, quality, and importance of the parameters featured in the XSDs. | M13 | M24 | | | |
| A2.1 Develop quality measure | M13 | M24 | | | |
| | | | | | |

| | | | | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|------------|----------------------------------------------------------------|--|--|
| WORKPACKAGE 3: Data access management Objective: To formulate a data access management that promotes Open Access but that accommodates data ownership and confidentiality issues. Where possible, due attention will be given to integrating with data access management mechanisms promoted by the JISC e-infrastructure programme, such as promoted by the JISC Core Middleware initiative.. | M13 | M24 | | | |
| A3.1 Develop data access model | M13 | M24 | | | |
| | | | | | |
| WORKPACKAGE 4: Materials test data DDLs Objective: To develop schemas and rulesets for the most common materials tests. | M04 | M18 | | | |
| A4.1 Tensile testing at ambient temperature | M04 | M06 | Schema and ruleset for tensile testing at ambient temperature | | |
| A4.2 Tensile testing at elevated temperature | M07 | M09 | Schema and ruleset for tensile testing at elevated temperature | | |
| A4.3 Fatigue testing | M10 | M12 | Schema and ruleset for fatigue testing | | |
| A4.4 Fatigue crack growth testing | M13 | M15 | Schema and ruleset for FCG testing | | |
| A4.5 Impact testing at ambient temperature | M16 | M18 | Schema and ruleset for impact testing at ambient temperature | | |
| | | | | | |

| | | | | | |
|-------------------------------------------------------------------------------------------------------------------------|------------|------------|--------------------------------------------------------------------------------------------------------------|--|--|
| WORKPACKAGE 5: Materials test data ontologies | M19 | M33 | | | |
| Objective: To extend the schemas for common materials tests to equivalent ontologies. | | | | | |
| A5.1 Tensile testing at ambient temperature | M19 | M21 | Ontology for tensile testing at ambient temperature | | |
| A5.2 Tensile testing at elevated temperature | M22 | M24 | Ontology for tensile testing at elevated temperature | | |
| A5.3 Fatigue testing | M25 | M27 | Ontology for fatigue testing | | |
| A5.4 Fatigue crack growth testing | M28 | M30 | Ontology for FCG testing | | |
| A5.5 Impact testing at ambient temperature | M31 | M33 | Ontology for impact testing at ambient temperature | | |
| | | | | | |
| WORKPACKAGE 6: Dissemination and sustainability | | | | | |
| Objective: To disseminate knowledge and deliverables, and to promote the long term viability of the data centre. | | | | | |
| A6.1 Deploy project site and blog | M01 | M01 | Web site with blog support | | |
| A6.2 JISC rapid innovation proposal for enhanced publishing process | M01 | M01 | EP2MDC proposal | | |
| A6.3 EC Enterprise and Industry DG 2009 ICT Standardisation Work Programme proposals. | M01 | M04 | Proposals for data collection prototypes that engage industrial and research sectors | | |
| A6.4 IE Demonstrator input | M03 | M18 | A6.4. Contribute documentation to the IE Demonstrator project and post regularly to the IE Demonstrator blog | | |

| | | | | | |
|-------------------------------------------------------------------------------------------------------------------------------|-----|-----|-----------------------------------------------------------------------------------------------------------------------------------------------|--|--|
| WORKPACKAGE 7: Management and evaluation | | | | | |
| Objective: Meet both the terms and conditions of the JISC grant and the academic excellence demanded of a PhD project. | | | | | |
| A7.1 Project plan | M01 | M02 | A robust project plan developed over several iterations according to the recommendations of the JISC Project Management Guidelines (May 2008) | | |
| A7.2 Evaluation plan | M01 | M02 | An evaluation plan developed over several iterations according to the recommendations of the JISC Project Management Guidelines (May 2008) | | |
| A7.3 JISC reporting | M01 | M18 | Timely delivery of reports in accordance with the terms and conditions of the JSC grant | | |
| A7.4 Recruitment—internships | M01 | M26 | Annual recruitment of interns to be employed for an 8-week term to fulfil specific tasks complementary to the project. | | |
| A7.5 Recruitment—studentship | M01 | M03 | Recruitment of PhD student | | |
| A7.6 Supervision and mentoring | M02 | M36 | | | |
| A7.7 Scientific publications | M12 | M36 | | | |
| A7.8 PhD thesis | M34 | M36 | | | |

Members of Project Team:
Tim Austin
Philippa Reed
Kenji Takeda
Charlotte Rimer (summer intern)