Open Water Information Architecture

System Requirements Document

Version 2.0 Draft

OWIA Technical Working Group

November 26, 2017

Contents

1 Introduction **1**

2 Project Management Approach **3**

2.1 The Differences Between Objectives, Functional Requirements, Technical Requirements and Design Alternatives . . . 3

3 Concept of Operation **3**

3.1 Definition of the OWIA System . . . 4

3.2 Governance . . . 5

3.2.1 Technical Working Group (TWG) . . . 5

3.2.2 Stakeholder Working Group (SWG) . . . 5

3.3 Use Cases from Stakeholder Workshops and Follow-on Submissions . . . 5

3.4 Existing Reporting Requirements . . . 5

3.5 Information Gaps . . . 5

4 Functional Requirements **5**

4.1 FR-100-100: Data Acquisition . . . 5

4.1.1 FR-100-110: Manual . . . 5

4.1.2 FR-100-120: Automated . . . 6

4.2 FR-200-100: Quality Control . . . 6

4.2.1 FR-200-110: Verification . . . 6

4.2.1.1 FR-200-120: Documentation . . . 6

4.2.1.2 FR-200-130: Reproducibility . . . 6

4.2.1.3 FR-200-140: Data Traceability . . . 6

4.2.2 FR-200-150: Standardization . . . 6

4.2.2.1 FR-200-160: Metadata Conventions . . . 6

4.2.2.2 FR-200-160: File-naming Conventions . . . 6

4.2.3 FR-200-170: Interoperable Transformation . . . 6

4.2.3.1 FR-200-180: Separation of Data and Computation . . . 6

4.2.3.2 FR-200-190: Data Interoperability . . . 6

4.2.3.3 FR-200-200: Products or Resources . . . 7

4.3 FR-300-100: Publication . . . 7
4.3.1 FR-300-110: Cross-Referencing-Service . . . 7
4.3.1.1 FR-300-120: Assignment of Digital Object Identifiers . . . 7
4.3.2 FR-300-130: Packaging . . . 7
4.3.2.1 FR-300-140: Compression Methods . . . 7
4.3.2.2 FR-300-150: Archive File Formatting . . . 7
4.3.3 FR-300-160: Archival . . . 7
4.3.3.1 FR-300-170: Open Access Distribution . . . 7
4.4 FR-400-100: Data Traceability . . . 7
4.4.1 FR-400-110: Metadata Production . . . 7
4.4.2 FR-400-120: Intellectual Property Rights Management . . . 7
4.4.3 FR-400-130: Public Law Compliance . . . 7
4.4.4 FR-400-140: Licensing . . . 8
4.4.5 FR-400-150: Liability . . . 8
4.4.6 FR-400-160: Searching . . . 8
4.4.6.1 FR-400-170: Cross-referencing System Integration . . . 8
4.4.6.2 FR-400-180: Search Engine Optimization . . . 8
4.4.7 FR-400-190: Version Control . . . 8
4.4.7.1 FR-400-200: Binary Data . . . 8
4.4.7.2 FR-400-210: Non-Binary Data . . . 8
4.4.8 FR-400-220: Anomaly Reporting . . . 8
4.5 FR-500-100: System Portability . . . 8
4.5.1 FR-500-110: Backup and Restore . . . 8
4.5.2 FR-500-120: Platform Portability . . . 8
4.6 FR-600-100: External Interfaces . . . 9
4.6.1 FR-600-110: Data and Metadata Acquisition . . . 9
4.6.2 FR-600-120: Data and Metadata Distribution . . . 9

5 Technical Requirements **9**
5.1 TR-100-100-00100: Data Acquisition Methods . . . 10
5.1.1 TR-200-200-00100: Data Transformation Methods . . . 10
5.1.2 TR-200-300-00100: Programming Languages . . . 10
5.1.3 TR-200-400-00100: Data Interoperability . . . 10
5.1.4 TR-200-5000-00100: Products (List Derived Products traceable to Objectives) . . . 10
5.1.5 TR-300-100-00100: Data Traceability . . . 10
5.1.6 TR-300-100-00200: Reproducibility of Data Products . . . 10
5.2 TR-400-100-00100: Standardization . . . 10
5.2.1 TR-400-100-00200: Metadata . . . 10
5.2.1.1 TR-400-100-00300: Resolution of Metadata conflicts . . . 11
5.2.1.2 TR-400-100-00400: Controlled Vocabulary . . . 11
5.2.1.3 TR-400-100-00500: Climate and Forecasting Conventions . . . 11
5.2.1.4 TR-400-100-00600: Federal Geographic Data Committee (FGDC) Geospatial Metadata Standards And Guidelines . . . 11
5.2.1.5 TR-400-100-00700: Open-Geospatial Consortium Standards and Guidelines . . . 11
5.2.2 TR-400-100-00800: Mapping Standards . . . 11
5.2.3 TR-400-100-00900: Numerical Accuracy and Precision Standards . . . 11
5.2.4 TR-400-100-01000: Measures of Uncertainty . . . 11

5.2.5 TR-400-100-01100: File Naming Convention . . . 11
5.3 TR-600-100-00100: Data Publication . . . 11
5.3.1 TR-600-100-00200: Methods . . . 11
5.3.1.1 TR-600-100-00300: Assignment of Digital Object Identifiers . . . 11
5.3.2 TR-600-100-00400: Metadata Production . . . 11
5.3.3 TR-600-100-00500: Open Access Distribution . . . 11
5.3.4 TR-600-100-00600: Intellectual Property Rights Management . . . 11
5.3.4.1 TR-600-100-00700: Public Law Compliance . . . 11
5.3.4.2 TR-600-100-00800: Licensing . . . 11
5.3.5 TR-600-100-00900: Discovery . . . 11
5.3.5.1 TR-600-100-01000: Cross-referencing System Integration . . . 11
5.3.5.2 TR-600-100-01100: Search Engine Optimization . . . 11
5.3.6 TR-600-100-01200: Packaging . . . 11
5.3.6.1 TR-600-100-01300: Compression Methods . . . 11
5.3.6.2 TR-600-100-01400: Archive File Formatting . . . 12
5.3.7 TR-600-100-01500: Version Control . . . 12
5.3.7.1 TR-600-100-01600: Binary Data . . . 12
5.3.7.2 TR-600-100-01700: Non-Binary Data . . . 12
5.3.8 TR-600-100-01800: Anomaly Reporting . . . 12
5.4 TR-600-100-00100: System Interoperability . . . 12
5.4.1 TR-600-100-00200: Backup and Restore . . . 12
5.4.2 TR-600-100-00300: Platform Portability . . . 12

Appendix A Standards and Conventions **13**
A.1 Operating Systems (OS) . . . 13
A.2 Metadata Schema . . . 13
A.3 Controlled Vocabulary . . . 15
A.4 Trusted Archives . . . 15
A.5 Digital Objects . . . 16
A.6 File Formats and Data Encodings . . . 16
A.7 Cross-referencing Services . . . 18
A.8 Commercial Search Services . . . 18

Appendix B Support for AB1755 **18**

Appendix C Traceability Tables **20**

List of Figures

1 Relationship between system baseline documents and operations. . . . 1
2 Conceptual representation of the system engineering process. *Reproduced from [19].* . . . 3

3 Illustration of the OWIA federation concept with a triumvirate governance structure of partners (OWIA-GP) supported by interacting stakeholder working group (SWG) and the technical working group (TWG). The federation is comprised of dedicated OWIA system implementations to enable individual data providers to independently integrate the OWIA system into their existing methods and procedures within their operations. Shared OWIA system implementations provide the flexibility for the harvesting of non-compliant data sources into an OWIA system implementation that will support the OWIA federation without insisting that the producers be OWIA compliant. 4
4 OWIA system (cf. section 3.1) functional block diagram. Parenthetical references point to the governing functional requirements. 9

List of Tables

1 Definition of OWIA classes of datasets. 10
2 Classification of file formats, content and related interoperability features. References are supplied in brackets and listed in the bibliography. 17
3 Traceability of AB1755 objectives (columns) to OWIA SOPs by use case (rows). 19
4 Traceability Table: Objective O-1100-1000 to Functional Requirements. This is an example of what subordinate Technical Requirements might resolve to and is meant only to characterize what *Resolution* of Functional Requirements might look like in a Technical Proposal. . 21

1 Introduction

This document contains the functional and technical requirements for the *Open Water Information Architecture (OWIA)* and is called the *OWIA System Requirements Document (SRD)*. It has within it an Appendix:A Standards and Conventions that contains narrative explanations that are referred to within individual requirements where appropriate. This is done because the requirements are meant to be terse, declarative, testable statements that are not overloaded with narrative exposition. There are two companion documents to the SRD: (1) the subordinate document *OWIA Standard Operating Procedures (SOPs)* and the (2) parent document *California Council for Science and Technology (CCST) Stakeholder Use Case* document.

The SOPs are compliant with the requirements specified here yet written at a more detailed level of abstraction with examples of programming code or sometimes pseudo-code to exemplify the implementation details important to developers as well as precisely documenting the processing steps (i.e., procedures) used to operate on data. It is meant to be analogous to an *OWIA Programmer's Guide* and, as the OWIA implementation proceeds, there will be open-source code repositories with *minimal working examples (MWE)* for use in improvements and innovations to current procedures and applications implementing those procedures.

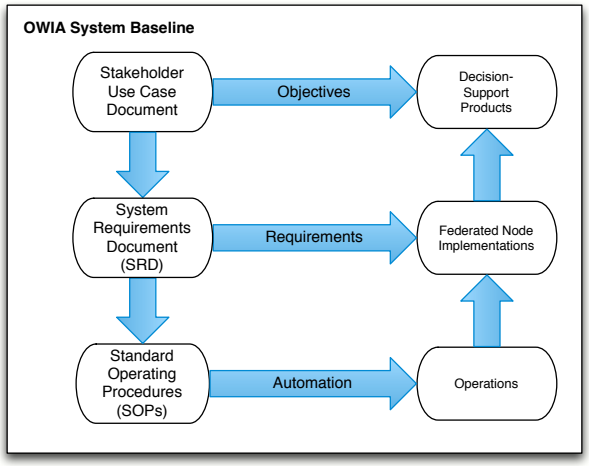

Figure 1: Relationship between system baseline documents and operations.

Each of these documents is intended for a technical audience although it is hoped that they are comprehensible to a motivated non-technical reader. There is a glossary in the back of the SRD to aid in navigating the technical language and as an effort to disambiguate some of the terms for which there may be competing and inconsistent definitions. In addition to these two, there is a third document that contains the stakeholder use cases used to develop the stakeholder objectives from each use case. These objectives are being used to define and constrain the requirements contained in the SRD and the procedures for satisfying them defined in the SOPs.

The SRD and SOPs are designed to provide a foundation for a community-based *OWIA* development of a federated set of cyberinfrastructure resources (i.e., computers, networks, data, metadata, and standards and conventions) that are interoperable and highly-automated to minimize labor as well as idiosyncratic anomalies. We therefore refer to them as the *baseline documents* (Figure 1). The objective of these baseline documents is to establish a framework for sustainable water resource management and to formalize that framework to a degree exemplified by other systems of standard methods such as those found in [7].

The federated nature of the OWIA extends to its (1) human governance structure as well as its (2) cyberinfrastructure (cf. Section 3 and Figure 3). Therefore we speak of the OWIA federation as including both these aspects and will differentiate the two parts contextually when using the term. The *open* aspect means open-access, open-source and open-architecture: encouraging innovation and automation while precluding the siloing and stove-piping that occurs when proprietary software and systems pose restrictive technology dependencies and requirements. The planning horizon is open-ended although intended to provide for a near-term operational system with an initial operating capability (IOC) within 1-2 years evolving to a final operating capability (FOC) over five (5) years that is operationally sustainable while responsive to technology innovation and risk minimization (i.e, cost, schedule, technical and operational) over its lifetime.

The approach is to follow standard system engineering practices [31] that: (1) define stakeholder objectives and, from these, (2) enumerate functional requirements in terms of functional components and

major interfaces both of which are implementation-independent, and (3) enumerate technical requirements which specify fundamental technical features such as network transfer rates, storage capacities, reliability, maintainability and availability (RMA), interface dependencies and contingencies and similar quantitative or qualitative requirements at a level of specificity (or abstraction) that is more detailed than the functional requirements on which they are based. It is also designed to present an initial evaluation of some of the obvious design trade-studies to explicate and focus on the key risk areas related to technical, schedule, cost and operational risks.

This is an interative and recursive, hierarchical design approach (Figure 2) which prioritizes *Stakeholder Objectives, Functional Requirements, and Technical Requirements* respectively and cross-correlates them to each other via a *traceability matrices* (Section 4) to ensure that there are no *widows or orphans* in the sense that there are no unsupported Objectives or Functional Requirements (i.e., widows) as well as no lower-level design features that are not specified in the Functional Requirements (i.e., orphans). As a development methodology, the system engineering method used here is sometimes contrasted with the agile development method. Every methodology has pros and cons and the reason we use this approach for the OWIA is because we already know a great deal about what is needed to improve access-to and reuse-of the collective set of water resource data and the OWIA focus is on the data content. This is not primarily a process of discovery and prototyping of software applications. For a broader discussion of the pros and cons of alternative software development approaches, the reader is encouraged to consider the discussions provided in [27] and [31].

Finally, some historical perspective is helpful. This document is meant to integrate the thinking on water resource information broadly and digital data about water resources specifically. The OWIA concept developed independently of the AB1755 legislation [1][29] that is currently, as of this writing, driving many efforts across the State of California to comply with its mandates and schedule. Fortuitously, the development of the OWIA and the activation of AB1755-related efforts overlap strongly such that AB1755 requirements are a subset of the broader OWIA requirements. The implementation of the OWIA will satisfy the requirements of AB1755 and support the Sustainable Groundwater Management Act (SGMA) in such a way that we can treat AB1755 as an OWIA use-case as described in Appendix B. The OWIA concept is a reflection and integration of a wide range of on-going efforts especially those in the UC WATER Security and Sustainability Research Initiative and CITRIS [9], California Council on Science and Technology (CCST), the Center for Western Weather and Water Extremes (CW3E)[36], the San Diego Supercomputer Center (SDSC) [4, 16, 12, 18, 15, 38, 33, 14, 30, 5, 11, 34, 17, 6, 2, 13, 20] and the UC Santa Barbara Bren School. We expect to grow this community to include private California universities, national laboratories and private sector partners as we go.

2 Project Management Approach

Figure 2 depicts the overall management approach used for this project and graphically summarizes it in the context of a *system engineering framework*. The system engineering framework is a set of methods and procedures for specifying design constraints to minimize the risk that a sought-after system implementation will successfully perform its intended functions. The output of this process is a system design and implementation that is *verifiably compliant with the functional and technical requirements* for the system and which can be *validated against the stakeholder objectives*.

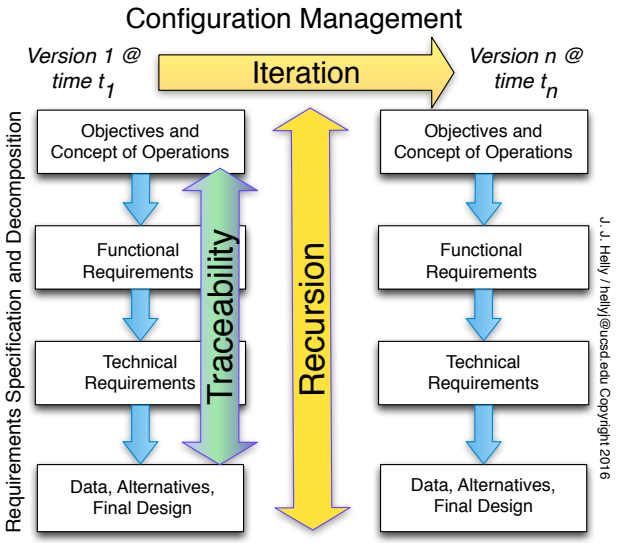

Figure 2: Conceptual representation of the system engineering process. *Reproduced from [19].*

2.1 The Differences Between Objectives, Functional Requirements, Technical Requirements and Design Alternatives

In designing and building any type of system the first step is to describe *what it is that you want to accomplish by building the system. These are the* ***objectives***. Sometimes these are called *stakeholder objectives*. They should be stated as simple declarative sentences focused on what the stakeholder wants the system to do. The language should be as non-technical as possible in order for the broadest understanding and consensus across the lay stakeholders who typically have diverse backgrounds and experience. On the other hand, the *functional requirements* are the translation of the objectives into engineering terms (i.e., functions using more precise technical language) describing how the objectives will be met. This is the first level of abstraction in specifying how the system will be implemented (Figure 2).

The articulation of objectives is often a stumbling block for stakeholders and developers alike because it poses a bit of chicken and egg or floor versus ceiling ambiguity. The way to get past this is focus on the description, often through workshops of stakeholders, of examples of usage without worrying very much whether something is an objective or a functional requirement. These can be re-factored by a *technical working group* once they are articulated. The most important thing is to articulate and document examples of how the system will be used. Once this process has an initial iteration, the system requirements document (SRD), can be maintained through configuration management of versions over time to provide accountability, via traceability tables, to the stakeholders as well as a path forward for developers and a contractually-applicable basis for acceptance testing for contract managers.

3 Concept of Operation

The OWIA system will be operated to produce standardized data of documented quality needed by stakeholders, as defined by the objectives, such as (1) figures, (2) tables, (3) statistics, and (4) analyses.

3.1 Definition of the OWIA System

The *OWIA system* is a federated collection of data systems, published using a shared set of data standards, standardized interfaces (i.e., APIs), and controlled vocabularies. The system should be designed adhering to open-data standards and principles and controlled by a set of functional requirements such as accessibility, interoperability, discoverability, and traceability. Open-data means that the data are free to use, re-use, and redistribute with no restrictions on their use [40, 37]. A federated system does not require a central catalogue or a single interface to such a catalogue. Rather, the federation permits the formation of one or more common catalogues (cf. Figure 4, Curated Collections) as an implementation dependent on the requirement for standardized, published metadata.

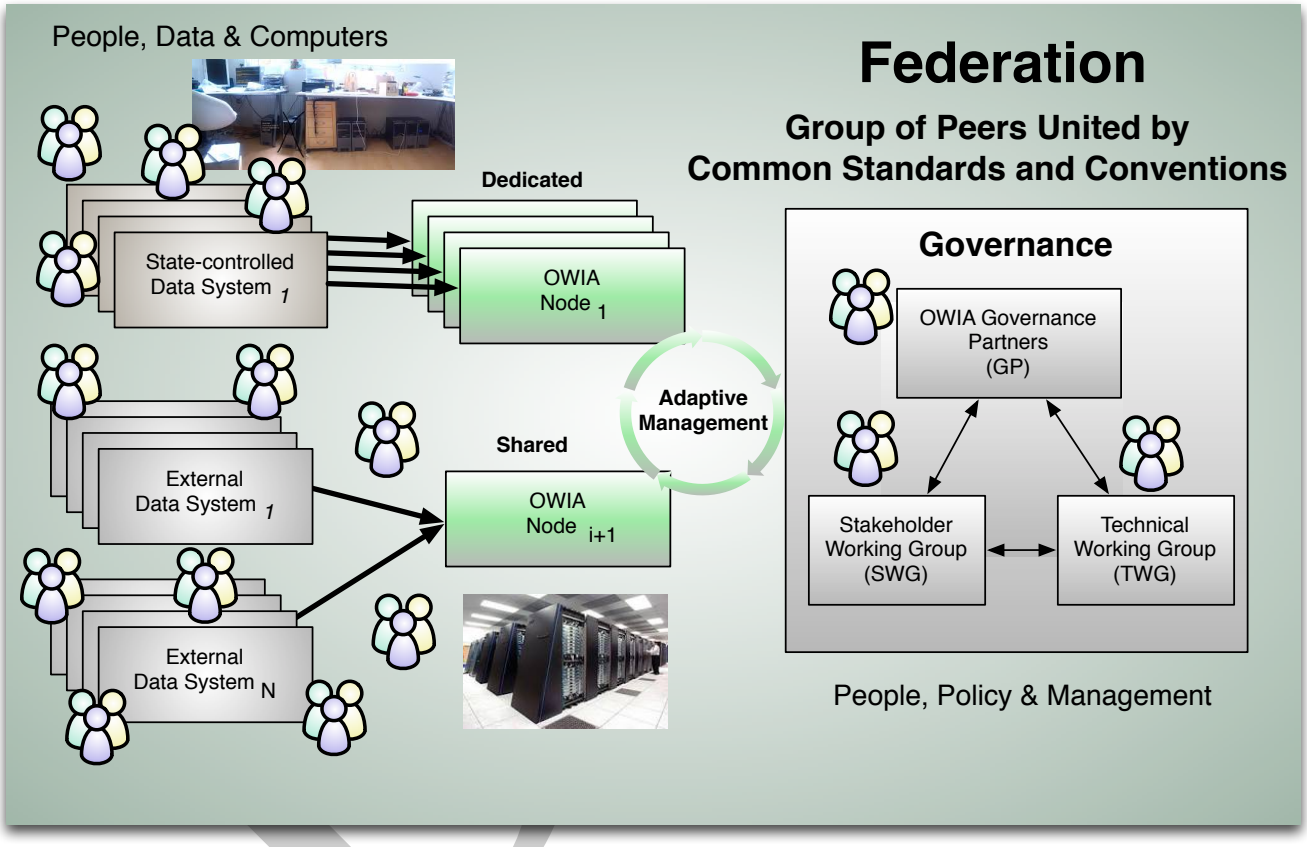

Figure 3: Illustration of the OWIA federation concept with a triumvirate governance structure of partners (OWIA-GP) supported by interacting stakeholder working group (SWG) and the technical working group (TWG). The federation is comprised of dedicated OWIA system implementations to enable individual data providers to independently integrate the OWIA system into their existing methods and procedures within their operations. Shared OWIA system implementations provide the flexibility for the harvesting of non-compliant data sources into an OWIA system implementation that will support the OWIA federation without insisting that the producers be OWIA compliant.

3.2 Governance

OWIA Governance Partners (OWIA-GP) are responsible for the direction of the system, policy, prioritization and resourcing of work efforts, and curation of functional requirements. This group must also meet with the Stakeholder Working Group to review feedback and hear use cases.

3.2.1 Technical Working Group (TWG)

The *Technical Working Group (TWG)* is responsible for the identifying, adopting, approving data standards, data publication approaches, and controlling functional and technical requirements.

3.2.2 Stakeholder Working Group (SWG)

The *Stakeholder Working Group (SWG)* is responsible for informing the OWIA Governance Partners of the intended uses of the OWIA system, and providing iterative feedback on the effectiveness of the system to meet their requirements.

3.3 Use Cases from Stakeholder Workshops and Follow-on Submissions

The use cases are documented in a separate report. They have been translated into *Stakeholder Objectives* (cf. Appendix **??**) for use in this document.

3.4 Existing Reporting Requirements

Figures and tables from existing reports implicitly define data requirements. As such, they can be harvested to be used to tightly define functional and technical requirements. Table **??** identifies and summarizes the current set of reports supported by the OWIA system requirements.

3.5 Information Gaps

Existing workflows are predicated on the legacy methods of mobilizing information and injecting it into the regulatory, producer and consumer facets of the information architecture. Consequently, there are likely to be a considerable range of opportunities for mobilizing and integrating relatively new sources of information, results of earth science research and technological methods of data integration at new spatial and temporal scales of measurement and analysis. These newer results may help to fill gaps in our ability to better manage California's water resources and may require a long-term commitment to validating them as acceptable sources of information in the regulatory and legal framework that currently exists.

4 Functional Requirements

4.1 FR-100-100: Data Acquisition

Manual and automated methods shall be provided for data acquisition. Data at the acquisition stage of OWIA processing shall be referred to as *Level 0* data.

4.1.1 FR-100-110: Manual

Manual data acquisition methods shall provide metadata conforming the the OWIA minimum metadata standard.

4.1.2 FR-100-120: Automated

Automated data acquisition methods shall provide metadata conforming to the OWIA minimum metadata standard.

4.2 FR-200-100: Quality Control

No data transformation shall require the use of proprietary software, methods or special-purpose computing platforms for data processing and transportation. Data that has received quality control processing according to OWIA standards and conventions shall be referred to as *Level 1* data.

4.2.1 FR-200-110: Verification

Data verification shall be accomplished according to OWIA standards and conventions (cf. Appendix A).

4.2.1.1 FR-200-120: Documentation Documentation shall be provided according to OWIA standards and conventions (cf. Appendix A).

4.2.1.2 FR-200-130: Reproducibility All data products shall be verifiably reproducible by an anonymous second-party from the input data, metadata and the processing methods used to produce the data product.

4.2.1.3 FR-200-140: Data Traceability All data products shall be traceable to their parent data sources to the extent that a data product composed of multiple input data sources shall be decomposable and traceable to its parents.

4.2.2 FR-200-150: Standardization

All data products shall conform to the OWIA standards and conventions (cf. Appendix A).

4.2.2.1 FR-200-160: Metadata Conventions Metadata shall be provided according to OWIA standards and conventions (cf. Appendix A).

4.2.2.2 FR-200-160: File-naming Conventions File-name shall be performed according to OWIA standards and conventions (cf. Appendix A).

4.2.3 FR-200-170: Interoperable Transformation

All data transformations shall be achievable with open-source, non-proprietary software, non-proprietary data formats and commodity computers.

4.2.3.1 FR-200-180: Separation of Data and Computation Data and computation shall be separated between data files and stored procedures.

4.2.3.2 FR-200-190: Data Interoperability All data products shall be interoperable across OWIA-supported computing platforms and be able to be operated on using non-proprietary, open-source software and commodity computers and communications systems to operate on them or transport them.

328 **4.2.3.3 FR-200-200: Products or Resources** Data products shall be developed in accordance with the
329 objectives as per section **??**.

330 4.3 FR-300-100: Publication

331 Data shall be published according to OWIA standards and conventions (cf. Appendix A).

332 4.3.1 FR-300-110: Cross-Referencing-Service

333 Data objects shall be registered with a cross-referencing service.

334 **4.3.1.1 FR-300-120: Assignment of Digital Object Identifiers** A digital object identifier (DOI) shall
335 be acquired for each Level 1 digital object according to the OWIA standards and conventions (cf. Appendix
336 A).

337 4.3.2 FR-300-130: Packaging

338 Packaging shall conform to OWIA standards and conventions (cf. Appendix A).

339 **4.3.2.1 FR-300-140: Compression Methods** Compression methds shall be non-lossy and conform to
340 OWIA standards and conventions (cf. Appendix A).

341 **4.3.2.2 FR-300-150: Archive File Formatting** Archive file formats shall be only those conforming with
342 OWIA standards and conventions.

343 4.3.3 FR-300-160: Archival

344 Data shall be archived in trusted data archives with external interfaces to provide for data access and trans-
345 portation to end-users and applications.

346 **4.3.3.1 FR-300-170: Open Access Distribution** All data products shall be accessible using OWIA stan-
347 dard protocols or transportable external media where network transport is impossible or impractical.

348 4.4 FR-400-100: Data Traceability

349 Data traceability shall be provided according to OWIA standards and conventions (cf. Appendix A).

350 4.4.1 FR-400-110: Metadata Production

351 All data products shall have metadata provided with them sufficient to meet the OWIA minimum metadata
352 standard.

353 4.4.2 FR-400-120: Intellectual Property Rights Management

354 Metadata shall be produced according to the OWIA standards and conventions (cf. Appendix A).

355 4.4.3 FR-400-130: Public Law Compliance

356 All data and metadata products shall comply with relevant public law requirements.

4.4.4 FR-400-140: Licensing

Licensing of data and metadata shall conform to OWIA standards and conventions (cf. Appendix A).

4.4.5 FR-400-150: Liability

Liability limitations shall be declared with each data object through the metadata in conformance with OWIA standards and conventions (cf. Appendix A).

4.4.6 FR-400-160: Searching

Minimal metadata shall be provided to meet OWIA standards and conventions for search and discovery.

4.4.6.1 FR-400-170: Cross-referencing System Integration Cross-referencing system integration shall be based on a digital objects DOI.

4.4.6.2 FR-400-180: Search Engine Optimization Search-engine optimization shall be based on the metadata associated with the DOI as a minimum.

4.4.7 FR-400-190: Version Control

Source code, data and metadata shall be version-controlled in conformance with OWIA standards and conventions (cf. Appendix A).

4.4.7.1 FR-400-200: Binary Data A version control method shall be provided for binary data products in accordance with OWIA standards and conventions (cf. Appendix A).

4.4.7.2 FR-400-210: Non-Binary Data A version control method shall be provided for non-binary data products in accordance with OWIA standards and convetions.

4.4.8 FR-400-220: Anomaly Reporting

There shall be a method for reporting of anomalies detected in the data products and there shall be a method of tracking the anomalies for resolution and notification of those that have obtained the anomalous data products that an anomaly has been reported and a method for determining the resolution of the anomaly.

4.5 FR-500-100: System Portability

System portability shall be provided such that any implementation of the OWIA system is portable to other platforms in conformance with OWIA standards and conventions (cf. Appendix A).

4.5.1 FR-500-110: Backup and Restore

Backup and restore capabiltiy shall be provided using interoperable procedures and systems according to OWIA standards and conventions (cf. Appendix A).

4.5.2 FR-500-120: Platform Portability

Platform portability shall be provided in conformance with OWIA standards and conventions (cf. Appendix A).

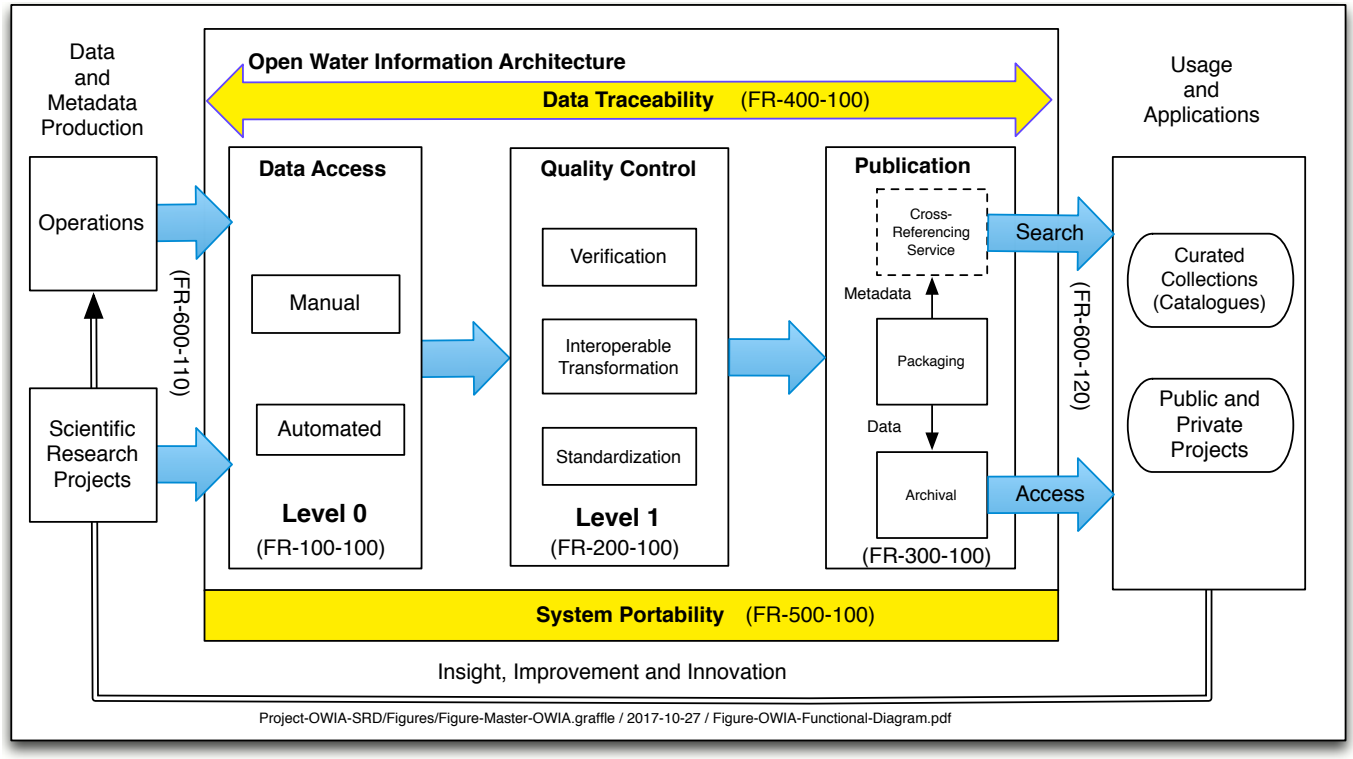

Figure 4: OWIA system (cf. section 3.1) functional block diagram. Parenthetical references point to the governing functional requirements.

4.6 FR-600-100: External Interfaces

External interfaces shall be provided for data acquisition and open-access to data products.

4.6.1 FR-600-110: Data and Metadata Acquisition

External interfaces shall be provided supporting the data sources in Appendix **??**.

4.6.2 FR-600-120: Data and Metadata Distribution

External interfaces to end-users and applications shall be provided supporting those listed in **??**.

5 Technical Requirements

The technical requirements are subordinate and traceable to the functional requirements above and have a higher-level of engineering detail and more precise, more technical language. This is the first level sufficiently specific to require decision-making about engineering trade-offs and what types of hardware, software and data representations qualify to be included in an OWIA-node.

Table 1: Definition of OWIA classes of datasets.

Dataset Class	Provenance	Transformation	Quality Control	Metadata	Published
Level 0	cf. Table 2	N/A	N/A	N/A	N/A
Level 1	Traceable to Level 0 parent	Interoperable	SOPs	Listing 1	DOI, Archived
Level 2	Traceable to multiple Level 1 parents (composite, decomposable)	Interoperable	SOPs	Listing 1	DOI, Archived

5.1 TR-100-100-00100: Data Acquisition Methods

5.1.1 TR-200-200-00100: Data Transformation Methods

All data transformations shall be accomplished with open-source, non-proprietary software and commodity computers for data processing and transportation.

5.1.2 TR-200-300-00100: Programming Languages

Data processing shall be realized through the use of stored procedures written in the GNU programming languages.

5.1.3 TR-200-400-00100: Data Interoperability

All data products shall be interoperable across OWIA-supported computing platforms and be able to be operated on using non-proprietary, open-source software and commodity computers and communications systems to operate on them or transport them.

5.1.4 TR-200-5000-00100: Products (List Derived Products traceable to Objectives)

5.1.5 TR-300-100-00100: Data Traceability

All data products shall be traceable to their parent data sources to the extent that a data product composed of multiple input data sources shall be decomposable and traceable to its parents.

5.1.6 TR-300-100-00200: Reproducibility of Data Products

All data products shall be verifiably reproducible by an anonymous second-party from the input data and the method used to produce the data product.

5.2 TR-400-100-00100: Standardization

5.2.1 TR-400-100-00200: Metadata

All data products shall have metadata provided with them sufficient to meet the OWIA minimum metadata standard.

5.2.1.1 TR-400-100-00300: Resolution of Metadata conflicts

All conflicts in metadata standards and conventions shall be subject to the determination of the TWG.

5.2.1.2 TR-400-100-00400: Controlled Vocabulary

All metadata shall be populated with controlled vocabulary drawn from the following standards.:

5.2.1.3 TR-400-100-00500: Climate and Forecasting Conventions

Climate and Forecasting conventions shall be used when suitable.

5.2.1.4 TR-400-100-00600: Federal Geographic Data Committee (FGDC) Geospatial Metadata Standards And Guidelines

FGDC standards and guidelines shall be used when suitable.

5.2.1.5 TR-400-100-00700: Open-Geospatial Consortium Standards and Guidelines

Open-Geospatial Consortium standards and guidelines shall be used when suitable.

5.2.2 TR-400-100-00800: Mapping Standards

5.2.3 TR-400-100-00900: Numerical Accuracy and Precision Standards

5.2.4 TR-400-100-01000: Measures of Uncertainty

5.2.5 TR-400-100-01100: File Naming Convention

5.3 TR-600-100-00100: Data Publication

5.3.1 TR-600-100-00200: Methods

5.3.1.1 TR-600-100-00300: Assignment of Digital Object Identifiers

5.3.2 TR-600-100-00400: Metadata Production

5.3.3 TR-600-100-00500: Open Access Distribution

5.3.4 TR-600-100-00600: Intellectual Property Rights Management

All data products shall be accessible using OWIA standard protocols or transportable external media where network transport is impossible or impractical.

5.3.4.1 TR-600-100-00700: Public Law Compliance

5.3.4.2 TR-600-100-00800: Licensing

5.3.5 TR-600-100-00900: Discovery

5.3.5.1 TR-600-100-01000: Cross-referencing System Integration

5.3.5.2 TR-600-100-01100: Search Engine Optimization

5.3.6 TR-600-100-01200: Packaging

5.3.6.1 TR-600-100-01300: Compression Methods

5.3.6.2 TR-600-100-01400: Archive File Formatting

5.3.7 TR-600-100-01500: Version Control

5.3.7.1 TR-600-100-01600: Binary Data

5.3.7.2 TR-600-100-01700: Non-Binary Data

5.3.8 TR-600-100-01800: Anomaly Reporting

5.4 TR-600-100-00100: System Interoperability

5.4.1 TR-600-100-00200: Backup and Restore

5.4.2 TR-600-100-00300: Platform Portability

Appendix A Standards and Conventions

This document narrative descriptions of the standards and conventions referred to in the functional and technical requirements. The Internet Engineering Task Force is used as a reference and as the default for standards and conventions that are not otherwise superseded by those identified here. For the subset of standards and conventions that pertain only to web-related services and interfaces, we refer to the World-Wide Web Consortium (W3C) unless explicity superseded in this Appendix.

A.1 Operating Systems (OS)

All conventional operating systems may be employed. If there is an exceptional consideration or doubt about a particular OS or version, it should be submitted as a review item for explicit consideration by the TWG.

A.2 Metadata Schema

There are many important metadata standards that bear consideration. Examples include:

1. Ecological Metadata Language ([22], [30]),
2. Open Geospatial Metadata
3. FGDC

For that reason, the metadata schema recommended here is referred to as a multilateral metadata convention that supports the production of arbitrary metadata files to support compliance with the current and future complement of metadata interfaces: it is designed to be independent of any particular standard but compatible with most.

The metadata schema in Listing 1 is the standard schema. It is intended to be the basis of all metadata interoperability schemas derived from it for integration purposes as required to integrate with other systems and catalogues. This metadata schema is dependent on the controlled vocabulary standard described in section A.3.

Listing 1: The current metadata schema.

```
# ##############################################################################################
# Canonical Collection
# ##############################################################################################
OWIA_CanonicalCollection_ArchivistEmail,"VARCHAR(50000)","Required","1","OWIA","Manager Email"
OWIA_CanonicalCollection_ArchivistInstitution,"VARCHAR(50000)","Required","1","OWIA","Manager Institution"
OWIA_CanonicalCollection_ArchivistName,"VARCHAR(50000)","Required","1","OWIA","Manager Name"
OWIA_CanonicalCollection_ArchivistPhone,"VARCHAR(50000)","Required","1","OWIA","Manager Phone"
OWIA_CanonicalCollection_CollectionIdentifier,"VARCHAR(50000)","Required","1","OWIA","Collection Identifier"
OWIA_CanonicalCollection_ControlledVocabulary,"VARCHAR(50000)","Required","1","OWIA","Controlled Vocabulary"
OWIA_CanonicalCollection_Creator,"VARCHAR(50000)","Required","1","OWIA","Creator of collection"
OWIA_CanonicalCollection_Description,"VARCHAR(50000)","Required","1","OWIA","Thorough Description of collection"
OWIA_CanonicalCollection_Language,"VARCHAR(50000)","Required","1","OWIA","Language"
OWIA_CanonicalCollection_MTFVersion,"VARCHAR(50000)","Required","1","OWIA","MTFVersion"
OWIA_CanonicalCollection_Ontology,"VARCHAR(50000)","Required","1","OWIA","Ontology"
OWIA_CanonicalCollection_Publisher,"VARCHAR(50000)","Required","1","OWIA","Publisher of collection"
OWIA_CanonicalCollection_Subject,"VARCHAR(50000)","Required","1","OWIA","General Subject area of collection"
OWIA_CanonicalCollection_Title,"VARCHAR(50000)","Required","1","OWIA","Specific Title of collection"
# ##############################################################################################
# Canonical ADO
# ##############################################################################################
OWIA_CanonicalADO_ADOIdentifier,"VARCHAR(50000)","Required","1","OWIA","ADO Identifier"
OWIA_CanonicalADO_ADOVersion,"VARCHAR(50000)","Required","1","OWIA","ADO Version"
OWIA_CanonicalADO_AccessControl,"VARCHAR(50000)","Required","1","OWIA","Access Control for this object"
OWIA_CanonicalADO_Author,"VARCHAR(50000)","Required","1","OWIA","Author"
OWIA_CanonicalADO_BlockTypes,"VARCHAR(50000)","Required","1","OWIA","Major data block types"
OWIA_CanonicalADO_Children,"VARCHAR(50000)","Required","1","OWIA","Children"
OWIA_CanonicalADO_CollectionIdentifier,"VARCHAR(50000)","Required","1","OWIA","Collection Identifier"
OWIA_CanonicalADO_ContentFilenames,"VARCHAR(50000)","Required","1","OWIA","Content Filenames"
```

```
OWIA_CanonicalADO_Contributor,"VARCHAR(50000)","Required","1","OWIA","Contributor of this upload"
OWIA_CanonicalADO_ControlledVocabulary,"VARCHAR(50000)","Required","1","OWIA","Controlled Vocabulary"
OWIA_CanonicalADO_Coverage,"VARCHAR(50000)","Required","1","OWIA","Coverage min max lat lon"
OWIA_CanonicalADO_Creator,"VARCHAR(50000)","Required","1","OWIA","Original creator of object"
OWIA_CanonicalADO_DOI,"VARCHAR(50000)","Required","1","OWIA","Digital Object Identifier"
OWIA_CanonicalADO_Date,"DATE","Required","1","OWIA","Date of publication into collection"
OWIA_CanonicalADO_Description,"VARCHAR(50000)","Required","1","OWIA","Description including importance"
OWIA_CanonicalADO_ExpertLevel,"VARCHAR(50000)","Required","1","OWIA","Expert Level"
OWIA_CanonicalADO_Filesize,"NUMERIC","Required","1","OWIA","Filesize"
OWIA_CanonicalADO_Format,"VARCHAR(50000)","Required","1","OWIA","Format MIME type"
OWIA_CanonicalADO_Keywords,"VARCHAR(50000)","Required","1","OWIA","Keywords"
OWIA_CanonicalADO_Language,"VARCHAR(50000)","Required","1","OWIA","Language"
OWIA_CanonicalADO_LatitudeEnd,"REAL","Required","1","OWIA","Latitude End"
OWIA_CanonicalADO_LatitudeNorth,"REAL","Required","1","OWIA","Latitude Northernmost"
OWIA_CanonicalADO_LatitudeSouth,"REAL","Required","1","OWIA","Latitude Southernmost"
OWIA_CanonicalADO_LatitudeStart,"REAL","Required","1","OWIA","Latitude at Start of object"
OWIA_CanonicalADO_LongitudeEast,"REAL","Required","1","OWIA","Longitude Easternmost"
OWIA_CanonicalADO_LongitudeEnd,"REAL","Required","1","OWIA","Longitude at End of object"
OWIA_CanonicalADO_LongitudeStart,"REAL","Required","1","OWIA","Longitude at Start of object"
OWIA_CanonicalADO_LongitudeWest,"REAL","Required","1","OWIA","Longitude Westernmost"
OWIA_CanonicalADO_MD5SUM,"VARCHAR(50000)","Required","1","OWIA","Verifier (md5sum)"
OWIA_CanonicalADO_MIFVersion,"VARCHAR(50000)","Required","1","OWIA","MetaData Content Version"
OWIA_CanonicalADO_MTFVersion,"VARCHAR(50000)","Required","1","OWIA","MetaData Template File Version"
OWIA_CanonicalADO_Ontology,"VARCHAR(50000)","Required","1","OWIA","Ontology"
OWIA_CanonicalADO_Parent,"VARCHAR(50000)","Required","1","OWIA","Parent"
OWIA_CanonicalADO_PhysicalStorageLocation,"VARCHAR(50000)","Required","1","OWIA","Physical storage location"
OWIA_CanonicalADO_Publisher,"VARCHAR(50000)","Required","1","OWIA","Publisher"
OWIA_CanonicalADO_Relation,"VARCHAR(50000)","Required","1","OWIA","Relation"
OWIA_CanonicalADO_Rights,"VARCHAR(50000)","Required","1","OWIA","Link to rights statement"
OWIA_CanonicalADO_Siblings,"VARCHAR(50000)","Required","1","OWIA","Siblings"
OWIA_CanonicalADO_Source,"VARCHAR(50000)","Required","1","OWIA","Source of object for Dublin Core"
OWIA_CanonicalADO_SourceFileName,"VARCHAR(50000)","Required","1","OWIA","Source File Name"
OWIA_CanonicalADO_Subject,"VARCHAR(50000)","Required","1","OWIA","Subject area of object"
OWIA_CanonicalADO_TimeEnd,"DATE","Required","1","OWIA","End Date Time of object"
OWIA_CanonicalADO_TimeStart,"DATE","Required","1","OWIA","Start Date Time of object"
OWIA_CanonicalADO_Title,"VARCHAR(50000)","Required","1","OWIA","Title to identify object in specific detail"
OWIA_CanonicalADO_Type,"VARCHAR(50000)","Required","1","OWIA","Type of Dublin Core resource"
OWIA_CanonicalADO_URL,"VARCHAR(50000)","Required","1","OWIA","Universal Resource Locator"
# ########################################################################
#  Documentation
# ########################################################################
OWIA_Documentation_MTFVersion,"VARCHAR(50000)","Required","1","OWIA","MTFVersion"
OWIA_Documentation_ADOIdentifier,"VARCHAR(50000)","Required","1","OWIA","ADOIdentifier"
OWIA_Documentation_ControlledVocabulary,"VARCHAR(50000)","Required","1","OWIA","ControlledVocabulary"
OWIA_Documentation_Ontology,"VARCHAR(50000)","Required","1","OWIA","Ontology"
OWIA_Documentation_Description,"VARCHAR(50000)","Arbitrary","1","OWIA","Document Description or Title"
OWIA_Documentation_Format,"VARCHAR(50000)","Arbitrary","1","OWIA","Format"
OWIA_Documentation_ObjectType,"VARCHAR(50000)","Required","1","OWIA","Data Object Type"
OWIA_Documentation_Type,"VARCHAR(50000)","Arbitrary","1","OWIA","Document Type"
# ########################################################################
#  Products
# ########################################################################
OWIA_Products_MTFVersion,"VARCHAR(50000)","Required","1","OWIA","MTFVersion"
OWIA_Products_ADOIdentifier,"VARCHAR(50000)","Required","1","OWIA","ADOIdentifier"
OWIA_Products_ControlledVocabulary,"VARCHAR(50000)","Required","1","OWIA","ControlledVocabulary"
OWIA_Products_Ontology,"VARCHAR(50000)","Required","1","OWIA","Ontology"
OWIA_Products_Description,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Product Description"
OWIA_Products_Format,"VARCHAR(50000)","Arbitrary","1","OWIA","Format"
OWIA_Products_Method,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Production Method"
OWIA_Products_ObjectType,"VARCHAR(50000)","Required","1","OWIA","Data Object Type"
OWIA_Products_Type,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Product Type"
# ########################################################################
#  CKAN per Greg Smith 2017-11-01
# ########################################################################
OWIA_CKAN_MTFVersion,"VARCHAR(50000)","Required","1","OWIA","MTFVersion"
OWIA_CKAN_ADOIdentifier,"VARCHAR(50000)","Required","1","OWIA","ADOIdentifier"
OWIA_CKAN_ControlledVocabulary,"VARCHAR(50000)","Required","1","OWIA","ControlledVocabulary"
OWIA_CKAN_Ontology,"VARCHAR(50000)","Required","1","OWIA","Ontology"
```

```
OWIA_CKAN_Description,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Product Description"
OWIA_CKAN_Format,"VARCHAR(50000)","Arbitrary","1","OWIA","Format"
OWIA_CKAN_Method,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Production Method"
OWIA_CKAN_ObjectType,"VARCHAR(50000)","Required","1","OWIA","Data Object Type"
OWIA_CKAN_Type,"VARCHAR(50000)","Arbitrary","1","OWIA","Data Product Type"
OWIA_CKAN_title,"VARCHAR(50000)","Required","OWIA","Title"
OWIA_CKAN_description,"VARCHAR(50000)","Required","OWIA","Description"
OWIA_CKAN_keyword,"VARCHAR(50000)","Required","OWIA","Tags"
OWIA_CKAN_modified,"VARCHAR(50000)","Required","OWIA","Last Update"
OWIA_CKAN_publisher,"VARCHAR(50000)","Required","OWIA","Publisher"
OWIA_CKAN_contactPoint,"VARCHAR(50000)","Required","OWIA","Contact Name and Email"
OWIA_CKAN_identifier,"VARCHAR(50000)","NA","OWIA","Unique Identifier"
OWIA_CKAN_accessLevel,"VARCHAR(50000)","Required","OWIA","Public Access Level"
OWIA_CKAN_bureauCodeUSG,"VARCHAR(50000)","NA","OWIA","Bureau Code"
OWIA_CKAN_programCodeUSG,"VARCHAR(50000)","NA","OWIA","Program Code"
OWIA_CKAN_license,"VARCHAR(50000)","NA","OWIA","License"
OWIA_CKAN_rights,"VARCHAR(50000)","NA","OWIA","Rights"
OWIA_CKAN_spatial,"VARCHAR(50000)","Required","OWIA","Spatial"
OWIA_CKAN_temporal,"VARCHAR(50000)","Required","OWIA","Temporal"
OWIA_CKAN_distribution,"VARCHAR(50000)","NA","OWIA","Distribution"
OWIA_CKAN_@type,"VARCHAR(50000)","NA","OWIA","Metadata Type"
OWIA_CKAN_accrualPeriodicity,"VARCHAR(50000)","NA","OWIA","Frequency"
OWIA_CKAN_conformsTo,"VARCHAR(50000)","NA","OWIA","Data Standard"
OWIA_CKAN_dataQualityUSG,"VARCHAR(50000)","NA","OWIA","Data Quality"
OWIA_CKAN_describedBy,"VARCHAR(50000)","NA","OWIA","Data Dictionary"
OWIA_CKAN_describedByType,"VARCHAR(50000)","NA","OWIA","Data Dictionary Type"
OWIA_CKAN_isPartOf,"VARCHAR(50000)","NA","OWIA","Collection"
OWIA_CKAN_issued,"VARCHAR(50000)","NA","OWIA","Release Date"
OWIA_CKAN_language,"VARCHAR(50000)","NA","OWIA","Language"
OWIA_CKAN_landingPage,"VARCHAR(50000)","NA","OWIA","Homepage URL"
OWIA_CKAN_primaryITInvestmentUIIUSG,"VARCHAR(50000)","NA","OWIA","Primary IT Investment UII"
OWIA_CKAN_references,"VARCHAR(50000)","NA","OWIA","Related Documents"
OWIA_CKAN_systemOfRecordsUSG,"VARCHAR(50000)","NA","OWIA","System of Records"
OWIA_CKAN_theme,"VARCHAR(50000)","NA","OWIA","Category"
```

A.3 Controlled Vocabulary

Controlled vocabulary is referenced to existing community standards where they exist. The OWIA will conform to the community standards listed here and these will be superseded by the OWIA Controlled Vocabulary when there is a conflict.

(1) CF Conventions and Metadata: Standard Names,
(2) World Meteorological Organization Practices,
(3) Open Geospatial Consortium WaterML 2.0,
(4) OWIA Standard Names (TBD).

A.4 Trusted Archives

Trusted archives are digital object repositories where data published within the OWIA community can be reliably found. These may not be the only locations but they are considered to be the primary authoritative sources of copies of digtial objects. The criteria used to determine trusted archive status are those of the

(1) USGS Acceptable Digital Repositories for USGS Scientific Publications and Data,
(2) CoreTrustSeal,
(3) DIN 31644 Information and documentation - Criteria for trustworthy digital archives, and
(4) ISO 16363:2012 Space data and information transfer systems - Audit and certification of trustworthy digital repositories.

OWIA trusted archives include:

(1) University of California (in discussion),
(2) California Department of Water Resources (TBD).

A.5 Digital Objects

Digital objects are anything that can be stored and retrieved from within the file system of an operating system. Streaming data presented to a display device are not considered to be digital objects since the data contained in the stream is not stored and cannot be directly used in reproducible analyses or unambiguously referred to or re-used. A suggested approach to employing data of this type is to checkpoint the stream into a file which can be used as a stored digital object.

A.6 File Formats and Data Encodings

Recommended file formats and data encodings are summarized in Table 2. The default standard for character encodings is UTF-8 [25] with extended ASCII as a secondary alternative.

Table 2: Classification of file formats, content and related interoperability features. References are supplied in brackets and listed in the bibliography.

OWIA Class	File Type	Structure	Encoding	Georeferencing	Controlled Vocabulary	Interoperability Tools
Level 1	Comma-separated Values (*.csv)	Record-oriented, Scalar (Integer, Float), Text	ASCII, UTF-8	EPSG [21], DAU-County, HUC10/12 [43]	OWIA, CF [3]	Any
	Spread-sheet (*.ods [42])	Record-oriented, Scalar (Integer, Float), Text	ASCII, UTF-8	EPSG	OWIA, CF	Any
	Geospatial (GDAL-supported)	Raster	Binary	EPSG	OWIA, CF	GDAL
		Vector	Any	EPSG	OWIA, CF	GDAL [8], ogr2ogr, QGIS [35], GRASS [10], GMT [44]
	NetCDF [41]	Multi-dimensional, self-documenting	Binary	EPSG [21], DAU-County, HUC10/12	OWIA, CF	NetCDF API, NCL, NCO [45], GMT
	Text-processing	Rich Text Format (rtf), free-text	Binary	N/A	OWIA, CF	OpenOffice, LibreOffice, rtf2latex, rtf2html
		TeX [39]	ASCII, UTF-8, human-readable	N/A	OWIA, CF	latex2rtf, latex2html, tex4ht
Level 0	Microsoft Excel Spreadsheets (*.xlsx, *.xls)	Cell, Worksheet	Binary	No	No	OpenOffice [32], LibreOffice [28]
	Microsoft Word Documents (*.docx, *.doc)	Free-text	Binary	No	No	OpenOffice, LibreOffice
	DBMS Export	Human-readable	ASCII, UTF-8 (*.txt)	No	No	None
	ESRI Geodatabase	Proprietary	Binary	No	No	QGIS (GDAL-enabled)
	NetCDF	Multi-dimensional, self-documenting	Binary	Any	Any	NetCDF API, NCL, NCO, GMT
Inadmissible	DBMS	Database Structure and Schema	Any	Any	Any	None
	Proprietary w/o Interoperability Tools	Proprietary	Any	Any	Any	None

A.7 Cross-referencing Services

Cross-referencing services are used to support global searching for digital objects published using the OWIA standards and conventions. The default system is the University of California's EZID service.

A.8 Commercial Search Services

Commercial search services using the WWW are typified by Google.

Appendix B Support for AB1755

The OWIA provides complete support for the open-data and transparency requirement of the AB1755 legislation. Table 4 summarizes the relationship between the functional requirments and the objectives stated in the AB1755 bill.

Table 3: Traceability of AB1755 objectives (columns) to OWIA SOPs by use case (rows).

Identifier	Name	Data Sharing	Documentation	Quality Control	Public Access	Open-source platforms and decision support tools
FR-100-100	Data Acquisition	X				
FR-100-110	*-Manual-	X				
FR-100-120	*-Automated-	X				
FR-200-100	Quality Control-*-	X		X		
FR-200-110	*-Verification-	X		X		
FR-200-120	*-*-Documentation	X	X	X		
FR-200-130	*-*-Reproducibility	X		X		
FR-200-140	*-*-Data Traceability	X		X		
FR-200-150	*-Standardization-	X	X	X		X
FR-200-160	*-*-File-naming Conventions	X	X	X		X
FR-200-170	*-Interoperable Transformation-	X		X		X
FR-200-180	*-*-Separation of Data and Computation	X		X		X
FR-200-190	*-*-Data Interoperability	X	X	X		X
FR-200-200	*-*-Products or Resources	X		X		X
FR-300-100	Publication-*-	X	X		X	
FR-300-110	*-Cross-Referencing-Service-	X	X		X	
FR-300-120	*-*-Assignment of Digital Object Identifiers	X	X		X	
FR-300-130	*-Packaging-	X			X	
FR-300-140	*-*-Compression Methods	X			X	
FR-300-150	*-*-Archive File Formatting	X			X	
FR-300-160	*-Archival-	X			X	
FR-300-170	*-*-Open Access Distribution	X			X	
FR-400-100	Data Traceability-*-	X	X		X	
FR-400-110	*-Metadata Production-	X	X		X	
FR-400-120	*-Intellectual Property Rights Management-	X	X		X	
FR-400-130	*-Public Law Compliance-	X	X		X	
FR-400-140	*-Licensing-	X	X		X	
FR-400-150	*-Liability-	X	X		X	
FR-400-160	*-Searching-	X			X	
FR-400-170	*-*-Cross-referencing System Integration	X			X	
FR-400-180	*-*-Search Engine Optimization	X			X	
FR-400-190	*-Version Control-	X	X			
FR-400-200	*-*-Binary Data	X	X			
FR-400-210	*-*-Non-Binary Data	X	X			
FR-400-220	*-Anomaly Reporting-	X	X			
FR-500-100	System Portability-*-					X
FR-500-110	*-Backup and Restore-					X
FR-500-120	*-Platform Portability-					X
FR-600-100	External Interfaces-*-	X				X
FR-600-110	*-Data and Metadata Acquisition-	X				X
FR-600-120	*-Data and Metadata Distribution-	X				X

Appendix C Traceability Tables

The traceability tables for *stakeholder objectives to functional requirements* and functional requirements to technical requirements are listed below in Tables 4 and **??**, respectively. These tables are provided to assist in the evaluation of change proposals and design approaches in order to understand more conveniently how any proposed change may ripple through the OWIA in unintended way and to provide a sound basis for engineering analysis of the interdependencies of the requirements both functional and technical as they bear upon project management and design decisions.

Table 4: Traceability Table: Objective O-1100-1000 to Functional Requirements. This is an example of what subordinate Technical Requirements might resolve to and is meant only to characterize what *Resolution* of Functional Requirements might look like in a Technical Proposal.

Functional Requirement	Label	Resolution
FR-100-100	Data Acquisition-*-NULL	
FR-100-110	*-Manual-NULL	Level 0: HTTP scraping (cf. Table **??** for UC001)
FR-100-120	*-Automated-NULL	Level 0: Stored procedures for updating
FR-200-100	Quality Control-*-NULL	
FR-200-110	*-Verification-NULL	Stored programs and transformation of Level 0 sources to OWIA standards, Compute checksums and version control a list of the checksums.
FR-200-120	*-*-Documentation	OWIA Standard Formats
FR-200-130	*-*-Reproducibility	Stored procedures and input data with descriptive metadata.
FR-200-140	*-*-Data Traceability	OWIA Level 0 metadata generation, OWIA standard Level 0 processing
FR-200-150	*-Standardization-NULL	OWIA Level 0 standard processing (verification of contents, anomaly detection, missing value coding)
FR-200-160	*-*-File-naming Conventions	OWIA Level 0 naming convention
FR-200-160	*-*-File-naming Conventions	Level 0 verification of data access and reproduction of quality control and standardization
FR-200-170	*-Interoperable Transformation-NULL	Level 0 metadata verification
FR-200-180	*-*-Separation of Data and Computation	
FR-200-190	*-*-Data Interoperability	EZID (External Interface)
FR-200-200	*-*-Products or Resources	(1) The water manager must identify potential source(s) of water, and for each determine the quantity and timing of water available for recharge and its cost. (2) To determine where the project should be located, the water manager must examine different options based on basin capacity and suitability of recharge areas; parcel data indicating available land and land values; and water quality implications based on current or past land use and the design of the project. (3) To determine the best method for recharge, basin characteristics such as subsurface characteristics, soil types, topography, current and planned land use, and basin capacity must be taken into account.
FR-300-100	Publication-*-NULL	Level 0 Metadata Production
FR-300-110	*-Cross-Referencing-Service-NULL	Transfer to Trusted Archive with public facing HTTPS server
FR-300-120	*-*-Assignment of Digital Object Identifiers	Identifier Assignment (e.g., EZID) via External Interface
FR-300-130	*-Packaging-NULL	AB1755
FR-300-140	*-*-Compression Methods	Lossless
FR-300-150	*-*-Archive File Formatting	tar.gz, zip
FR-300-160	*-Archival-NULL	Data shall be placed in a trusted archive for access and delivery using OWIA-compliant external interfaces.
FR-300-170	*-*-Open Access Distribution	ftp, http, rsync, scp, sftp, export
FR-400-100	Data Traceability-*-NULL	Via DOIs for parents and siblings.
FR-400-110	*-Metadata Production-NULL	Lossless
FR-400-120	*-Intellectual Property Rights Management-NULL	Attribution 4.0 International (CC BY 4.0), Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)
FR-400-130	*-Public Law Compliance-NULL	AB1755
FR-400-140	*-Licensing-NULL	Compute checksums and version control a list of the checksums.
FR-400-150	*-Liability-NULL	OWIA-standard version control system
FR-400-160	*-Searching-NULL	OWIA bug tracking system
FR-400-170	*-*-Cross-referencing System Integration	Crossref, DataCite
FR-400-180	*-*-Search Engine Optimization	Google bots
FR-400-190	*-Version Control-NULL	Open-source systems verified on Linux, Windows, OSX

FR-400-200	*-*-Binary Data	Naming convention.
FR-400-210	*-*-Non-Binary Data	ASCII-based version control systems (e.g., git, svn, mercurial)
FR-400-220	*-Anomaly Reporting-NULL	Curatorial email address
FR-500-100	System Portability-*-NULL	Open-source operation on major operating systems.
FR-500-110	*-Backup and Restore-NULL	Rsync-based
FR-500-120	*-Platform Portability-NULL	Demonstrated operation across major platforms: Linux, OSX, Windows
FR-600-100	External Interfaces-*-NULL	Uniquely identified per the Interface Control Appendix.
FR-600-110	*-Data and Metadata Acquisition-NULL	Compliant with OWIA standards and conventions
FR-600-120	*-Data and Metadata Distribution-NULL	Compliant with OWIA standards and conventions

Glossary

federated See Federation 1

federation A federation is a group of data providers and users using jointly agreed-upon standards of operation in a collective fashion to ensure the interoperability of the resources they collectively hold and employ. The term may be used, for example, when describing the interoperation of distinct cyberinfrastructure networks with different internal structures. The term may also be used when human groups agree to collectively manage cyberinfrastructure development and operation using commonly held, and managed, requirements, standards and conventions, and operating procedures to ensure the interoperability of distinct cyberinfrastructure resources (cf. Wikipedia Definition). 1, 23

Federation See federation 23

interoperability The ability of computer systems or software to exchange and make use of data (adapted from the Oxford English Dictionary). 23

procedures An established or official way of doing something (Oxford English Dictionary). 1, 23

protocol Protocols are methods of implementing a set of objectives and requirements in a systematic way. In computing, protocols mean both specific implementations of methods such as HTTP [26] and FTP [23] and, more generally as described by the *Internet Engineering Task Force*, protocols are sequences of processing steps that are also referred to as procedures [24]. 7

References

[1] California Legislature, *AB-1755 The Open and Transparent Water Data Act. (2015-2016)*, Assembly Bill 1755, California Legislature, 2015.

[2] Dan Cayan and John Helly, *The Wireless Watershed at the Santa Margarita Ecological Reserve*, Southwest Hydrology **September/October** (2003).

[3] CF Governance Conventions Committee, *CF Conventions and Metadata*.

[4] B. Chadwick, P.F. P.F. Wang, J. Helly, M. Brand, R. Flick, A. Young, W. O'Reilly, P. Bromirski, W. Crampton, R. Guza, J. Helly, T. Nishikawa, S. Boyce, M. Landon, M. Martinez, I. Canner, and B. Leslie, *A Methodology for Assessing the Impact of Sea Level Rise on Representative Military Installations in the Southwestern United States*, Technical Report RC-1703, SPAWAR Systems Center Pacific, 2014.

[5] CODATA-ICSTI Task Group on Data Citation Standards and Practices, *Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data*, Data Science Journal **12** (2013), 1–75.

[6] B. W. Eakins, S.P. Miller, J. Helly, and B. Zelt, *The fully electronic IODP Site Survey Data Bank*, Scientific Drilling **2** (2006), 40–42.

[7] Eugene W. Rice and Roger B. Baird and Andrew D. Eaton and Lenore S. Clesceri (ed.), *Standard Methods for the Examination of Water and Wastewater*, vol. 22nd Edition, American Public Health Association, American Water Works Association, Water Environment Federation, 2012.

[8] Geospatial Data Abstraction Library, *Geospatial Data Abstraction Library*.

[9] Steven Glaser, Martha Conklin, and Roger Bales, *Building an Intelligent Water Information System – American River Prototype*, Technical report, CITRIS, 2013.

[10] GRASS Development Team, *Geographic Resources Analysis Support System (GRASS GIS) Software*, Open Source Geospatial Foundation, 2012.

[11] Kay Gross, Edie Allen, Caroline Bledsoe, Robert Colwell, Paul Dayton, Megan Dethier, John Helly, Robert Holt, Nancy Morin, William Michener, Steward T. Pickett, and Susan Stafford, *Report of the Committee on the Future of Long-term Ecological Data (FLED)*, Tech. report, Ecological Society of America, Washington, D. C., 1995.

[12] J. Helly, T. T. Elvins, D. Sutton, D. Martinez, S Miller, S. Pickett, and A. M. Ellison, *Controlled Publication of Digital Scientific Data*, CACM **45** (2002), no. 5, 97–101.

[13] J. J. Helly, *Visualization of ecological and environmental data*, LTER Network Office, University of New Mexico, Albuquerque, New Mexico, 1998.

[14] John Helly, *New concepts of publication*, Nature **393** (1998).

[15] _____, *Digital libraries in hydrology*, Hydroinformatics: Data Integrative Approaches in Computation, Analysis, and Modeling (Praveen Kumar, Mike Folk, Momcilo Markus, Jay C Alameda, and Peter Bajcsy, eds.), Taylor and Francis, London, England, 2005, p. 552.

[16] John Helly and Maggi Kelly, *Collaborative Management of Natural Resources in San Diego Bay*, Coastal Management **29** (2001), 16.

[17] John Helly, Hubert Staudigel, and Anthony Koppers, *Scalable models of data sharing in Earth sciences*, Geochemistry Geophysics Geosystems **4** (2003).

[18] John J. Helly, *Cyberinfrastructure for data authorship, publication and application interoperability*, no. 1501928, AGU, AGU, December 2012.

[19] _____, *Systems engineering for scalable cyberinfrastructure: Data, software, systems*, DRAFT, 2017.

[20] John J. Helly and Kevin T. Herbinson, *Visualization of the salinity plume from a coastal ocean water desalination plant*, Water Environment Research **66** (1994), no. 5, 753–758.

[21] International Association of Oil & Gas Producers, *International Association of Oil & Gas Producers*.

[22] International Standards Organization (ISO), *EML - Ecological Metadata Language*, http://www.dcc.ac.uk/resources/metadata-standards/eml-ecological-metadata-language (2017).

[23] Internet Engineering Task Force (IETF), *The File Transfer Protocol*, RFC 354, IETF, 1972.

[24] _____, *The Internet Standards Process – Revision 3*, Tech. Report RFC 2026, IETF, October 1996.

[25] _____, *UTF-8, a transformation format of ISO 10646*, Internet standard, IETF, November 2003.

[26] _____, *Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing*, RFC 7230 ISSN: 2070-1721, IETF, June 2014.

[27] ISTQB EXAM CERTIFICATION, *What is Agile model – advantages, disadvantages and when to use it?*, http://istqbexamcertification.com/what-is-agile-model-advantages-disadvantages-and-when-to-use-it (2018).

[28] LibreOffice, *LibreOffice*.

[29] J. Gage Marchini, *Connecting the "Drops" of California Water Data: Chapter 506: The Open and Transparent Water Data Act*, The University of the Pacific Law Review (2017), 785–799.

[30] William K. Michener, James W. Brunt, John J. Helly, Thomas B. Kirchner, and Susan G. Stafford, *Nongeospatial metadata for the ecological sciences*, Ecological Applications **7** (1997), no. 1, 330–242.

[31] NASA, *NASA Systems Engineering Handbook*, Tech. Report NASA-SP-2007-6105-Rev-1-Final-31Dec2007, NASA, 2007.

[32] OpenOffice, *OpenOffice*.

[33] Paul F. Uhlir Rapprteur, *For Attribution – Developing Data Attribution and Citation Practices and Standards: Summary of an International Workshop*, Report ISBN 978-0-309-26728-1, National Acadeies Press, 2012.

[34] J. Pundsack, R. Belland, D. Broderson, G.C. Fox, J. Dozier, J. Helly, W. Li, P. Morin, M. Parsons, A. Roberts, C. Tweedie, and C. Yang., *Report on Workshop on Cyberinfrastructure for Polar Sciences. St. Paul, Minnesota. University of Minnesota Polar Geospatial Center.*, Report, National Science Foundation, 2013.

[35] Quantum GIS Development Team, *Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project*.

[36] F. M. Ralph, K. A. Prather, D. Cayan, J. R. Spackman, P. DeMott, M. Dettinger, C. Fairall, R. Leung, D. Rosenfeld, S. Rutledge, D. Waliser, A. B. White, J. Cordeira, A. Martin, J. Helly, and J. Intrieri, *CalWater Field Studies Designed to Quantify the Roles of Atmospheric Rivers and Aerosols in Modulating U.S. West Coast Precipitation in a Changing Climate.*, Bulletin of the American Meteorological Society (2015).

[37] SmartH20, *Standards and open data report*, Tech. Report smarth2o.D82.SUPSI.WP8.V1.1, Coordinator: SUPSI Contractors: POLIMI, UoM, SETMOB, EIPCM, TWUL, SES, MOONSUB, July 31 2015.

[38] Hubert Staudigel, John Helly, Anthony A. P. Koppers, Henry F. Shaw, William F. McDonough, Albrecht W. Hofmann, Charles H. Langmuir, Kerstin Lehnert, Baerbel Sarbas, Louis A. Derry, and Alan Zindler, *Electronic data publication in geochemistry*, Geochem. Geophys. Geosyst. **4** (2003), no. 3.

[39] TeX Users Group, *TeX Users Group*.

[40] The Aspen Institute, *Internet Of Water: Sharing And Integrating Water Data For Sustainability: A Report From The Aspen Institute Dialogue Series On Water Data*, https://assets.aspeninstitute.org/content/uploads/2017/05/Internet-of-Water-Report-May-2017.pdf ISBN: 0-89843-665-6, The Aspen Institute, One Dupont Circle, NW Suite 700, Washington, DC 20036, 2017.

[41] Unidata Community Programs, *NetCDF*.

[42] United States Library of Congress, *Sustainability of Digital Formats: Planning for Library of Congress Collections*.

[43] USGS, *Hydrologic Unit Maps*.

[44] P. Wessel and W. H. F. Smith, *Free software helps map and display data*, EOS Trans. AGU **72** (1991), 441.

[45] Charlie Zender, *netCDF Operator (NCO)*.