Towards a Reference Model for the LifeWatch ICT Infrastructure

Vera Hernández Ernst[*], Axel Poigné[†], Jon Giddy[‡], Alex Hardisty[§], Angi Voss[¶], Hans Voss[‖]

Abstract: The LifeWatch Reference Model provides the basis for an interoperable ICT infrastructure for European biodiversity research building on standards whenever feasible. Distinguishing features will be support of workflow for scientific in-silico experiments, tracking of provenance, and semantic support for interoperability. This paper presents the key requirements and the key architectural concepts to satisfy these requirements.

1 Introduction

The European Strategy Forum on Research Infrastructures (ESFRI) identified the opportunity to strengthen biodiversity research in Europe by selecting LifeWatch for its 1st Roadmap. The Life Watch project[1], co-ordinated by the University of Amsterdam, has brought together eight European networks in biodiversity science, 19 national governments and, through its committees, further platform providers and users. The European Commission is funding the preparatory phase February (2008 - 2011) through the Framework Program (FP7) Infrastructures, were a roadmap as well as the legal, financial, and technical arrangements for an operational phase of 30 years is been prepared.

The LifeWatch ICT Infrastructure will be a distributed system of nodes that provide access to and processing of biodiversity data from a variety of sources through common open interfaces. Whilst it is difficult to predict future developments, even on the conceptual level, the infrastructure is based on common principles such as reusability, modularity, portability, interoperability, discoverability, and, in particular, compliance with standards. To meet these requirements, one of the technical arrangements of the preparatory phase is the development of a reference model for the specification of a service-oriented architecture.

This paper presents the current status of the Reference Model for the LifeWatch ICT infrastructure. The Reference Model provides guidelines for the specification and implementation of infrastructure as well as defining a number of generic information models

[*] Fraunhofer IAIS, vera.hernandez@iais.fraunhofer.de
[†] Fraunhofer IAIS, axel.poigne@iais.fraunhofer.de
[‡] Cardiff University, J.P.Giddy@wesc.ac.uk
[§] Cardiff University, hardistyar@Cardiff.ac.uk
[¶] Fraunhofer IAIS, angelika.voss@iais.fraunhofer.de
[‖] Fraunhofer IAIS, hans.voss@iais.fraunhofer.de
[1] `http://www.lifewatch.eu/`

and services.

The LifeWatch infrastructure is based on the assumptions that:

- Functionality is broken into component services based on the principles of *Service Oriented Architecture*
- *Workflows* are used for the chaining of operations from multiple distributed services in order to perform specific user tasks
- *Semantic Services* provide uniform semantically defined interfaces enabling syntactical and semantic interoperability between and substitution of components
- *Provenance* information about documents, data, and methods replaces the traditional "laboratory notebooks" where provenance comprises all the information about the source or origin of an object, its history and pedigree, and its derivation and passage through time.

The present paper only touches upon the first item providing a reference model for the information and services models to be constructed in LifeWatch.

The paper proceeds by referring to related work and standards. In Section 3, the requirements of LifeWatch are outlined from a user's perspective. Sections 4 and 5 sketches the architectural concepts and the structure of the reference model. Finally, we present ideas for the integration of existing source systems in Section 6 before discussing the present status of the reference model and pointing to ongoing work. Finally, we draw some preliminary conclusions.

2 Related Work and Relevant Standards

The LifeWatch Reference Model is based on the ORCHESTRA Reference Model (RM_OA) [1], a generic specification framework for geospatial service-oriented architectures and service networks that is endorsed as best practice by the Open Geospatial Consortium (OGC). In that LifeWatch builds on an established and proven framework with similar requirements and assumptions. While the LifeWatch Reference Model may at present be considered as an instantiation of the ORCHESTRA Reference Model, further development of the LifeWatch Reference Model will lead to deviations from and specialisation of the ORCHESTRA model, in particular with regard to information models and the engineering and technology viewpoint. Hence, given as well the longevity of the LifeWatch project, the label 'LifeWatch Reference Model' seems to be justified.

Biodiversity Information Standards[2] (TDWG) is an international not-for-profit group that develops standards and protocols for sharing biodiversity data. LifeWatch will adopt its recommendations where appropriate. For biodiversity data domains not yet covered by

[2] formerly known as the "Taxonomic Data Working Group" (`http://www.tdwg.org`)

TDWG, LifeWatch will be well positioned to contribute to the TDWG process in order to achieve internationally agreed specifications.

Since biodiversity data are, in general, geolocated, geospatial standards are of particular concern. The European Directive for Spatial Information[3] (INSPIRE) aims at providing the European spatial information infrastructure that delivers to the users integrated spatial information services. The LifeWatch Reference Model aims at complying with the INSPIRE Implementation Rules for data and metadata so that LifeWatch can access all INSPIRE-conformant data and services and can itself provide INSPIRE-conformant data and services.

GMES[4] and GEOSS[5] will be important sources of earth observation data for LifeWatch. GEOSS, in particular, has a focus on biodiversity data acquisition with the long-term aim of integrating a distributed biodiversity observation network with sectorial, crisis, health, and policy systems. It can be expected that all these data will be made available according to the INSPIRE Implementation Rules, hence will be accessible for LifeWatch.

The Open Geospatial Consortium (OGC) is an influential consortium for geospatial standards, and many spatial data are provided through OGC services. OGC has agreed to develop INSPIRE-conformant service specifications.

Geography Mark-up Language (GML) [3] is the data exchange format of OGC and INSPIRE. Ecological Metadata Language[6] (EML) is the metadata specification for the documentation and exchange of information about ecological data. LifeWatch will adopt GML, in particular the INSPIRE application schemas in GML, for processed spatial data and take EML into account as a candidate for metadata specification, besides ISO and INSPIRE.

The Open Grid Forum[7] (OGF) is an important source of standards and specifications for applied distributed computing fundamentals (i.e., Grid computing technologies, e-Infrastructure, etc.) that will be taken into account for defining the technological basis of the LifeWatch infrastructure.

3 Requirements for the LifeWatch Infrastructure

The primary user groups of the LifeWatch infrastructure will be biodiversity researchers. They will benefit from access to digital data sources and new digital services carrying out computational experiments more easily and effectively and being well documented.

But it is anticipated that the infrastructure will be capable of supporting a much wider range of uses for which supporting biodiversity science might be required. Such uses include, for example environmental management and control, conservation management, and support for other European research infrastructures and networks with interests overlapping the biodiversity domain (e.g., Aurora Borealis, ANAEE, EvolTree).

[3] http://inspire.jrc.ec.europa.eu/
[4] http://www.gmes.info/
[5] http://www.earthobservations.org/
[6] http://knb.ecoinformatics.org/software/eml/
[7] http://www.ogf.org/

The new infrastructure will, therefore, offer facilities for biodiversity information and analytical capabilities for specific markets and user group demands, and will develop and offer expert services and products that in return will also create new demands.

The communication between stakeholders and LifeWatch product management / development will be mediated through the LifeWatch Service Centre (e.g., through the User Platform and the Data Providers Platform). Research networks and research sites may act as data providers and/or service providers to LifeWatch. They may operate LifeWatch core functions on behalf of LifeWatch.

3.1 Functional Requirements

The functional requirements concern the types of operations that the users need in order to find, access, process and view data. They include:

- Searching and browsing mechanisms for distributed data and services.
- Access to existing data and services, distributed among multiple organisations. Data and service providers continue to manage their data (and services) independently as now, including control of the creation and modification of data/services. However, data can be accessed by authorised users located anywhere through a generic mechanism defined by LifeWatch.
- Mechanisms for source data preservation, i.e., access to past versions of data sets that have been used to produce secondary information.
- Capture data from users and lightweight devices, including field sensors and networks providing continuous streams of new data, and portable computing devices, often with intermittent connectivity.
- Mechanisms for data analysis as well as mapping and modelling tools, using standard ways to manipulate and view data.
- Mechanisms for data fusion, integrating different sources (such as sensor data, biodiversity parameters, geographic data, primary data, workflow execution), to allow fast retrieval at different levels of detail e.g. for analysis and visualisation.
- Support the understanding of results by the user, by providing tools and mechanisms to enhance knowledge extraction from discovery as well as from analysis results
- Provide control mechanisms for access to data and services, together with monitoring services for the support of service level agreements and, potentially, including charging mechanisms.

3.2 Advanced Capabilities

Beside these base functionalities, LifeWatch plans to provide further advanced capabilities that will distinguish LifeWatch from other Infrastructures currently being build by other biodiversity projects and organisations like GBIF, PESI. These are in particular concerned with the support of workflows, semantics, and provenance:

- *Service chaining as the basis for e-Science experiments* A biodiversity research task will usually involve a number of services that are executed sequentially, in parallel, or iteratively. LifeWatch will provide mechanisms for ”service chaining” and speaks of a ”workflow” when referring to a particular composition of service calls. Workflows are considered as the essential mechanism for formalising “in-silico” experiments that thus become reproducible and may be scrutinised by other researchers or research groups.

- *Semantic mediation for loosely coupled components* will assist to achieve interoperability on data and service discovery, data mediation, data fusion, data interpretation, service integration, and workflow identification

- *Provenance* To make the sharing of data, methods and experiments in LifeWatch attractive, all data shall be traceable back to their origin and usage of source data shall be documented. The (re-) examination of scientific results demands an exact recording of workflows in terms of data and services. LifeWatch will provide automatic tracking as a provenance mechanism for correct citations of all resources used by a research task; reproduction of experiments; and, for listing of all usages (citations) for a dataset or workflow (i.e. life cycle of a data set or workflow in LifeWatch).

- *Unambiguous identification and descriptive mechanisms* The LifeWatch infrastructure must be able to identify each component and resource unequivocally. This is not only important for granting access to the correct resource after it has been discovered, but also for reproducing workflows and creating provenance information for each transaction in the life cycle of a resource in LifeWatch. LifeWatch will attach a Globally Unique Identifier (GUID) to each resource from the first moment a resource is used. There are currently several mechanisms for constructing and attaching a GUID and some sources already have such a property.

- *Annotation of Data and Services* Allow users to add annotations to existing data and services, which may contribute to the quality assessment and feedback processes.

- *Integration of Source Systems* A source system is a container of unstructured, semi-structured, or structured data and/or a provider of functions in terms of services. The source systems that LifeWatch will interact with are of a very heterogeneous nature and contain information of a variety of types and in a variety of formats. Typical examples of source systems for LifeWatch are biodiversity data collections, provided by either a dedicated data providers (e.g., natural history collections, monitoring organisations, etc.) through clearing-houses such as GBIF and OBIS or new data (e.g., from electronic sensors or manual recordings).

There are data sources, systems, and services, referred to as *External Source Systems*, that do not comply with the architectural requirements imposed by LifeWatch since they pre-date the LifeWatch "era". To accommodate such systems, the general idea and challenge is to provide them with interfaces that conform to LifeWatch principles through source system integration services to become "LifeWatch Source Systems".

- *Virtual Organisations* Support collaboration between researchers for accessing remote resources, by forming virtual organisations, communicating in groups, sharing and discussing results, sharing knowledge in restricted groups, etc. Collaboration requirements include:

4 The Architectural Concept

The technical vision for LifeWatch is a network of services providing secure access across multiple organisations to biodiversity and related data and to relevant analytical and modelling tools to collaborative groups of researchers. The infrastructure will be implemented using the ideas of a "Service Networks", Spatial Data Infrastructures, and Grid Computing.

The definition of functional domains for a service network below provides a first classification of the infrastructure components according to interaction mechanisms. This description appeals to the user perspective.

- the *Resource Domain* contains resources such as data repositories, computational capacity, sensor devices, and modelling/analysis/visualisation tools that contribute to the LifeWatch system. In the RM-OA, this is called the Source System Domain.
- the *Infrastructure Domain* provides mechanisms for sharing the specific resources as generic services in a distributed environment spread across multiple administrative domains. This includes systems for identifying, accessing and processing resources located within multiple administrative domains, uniform security and access protocols, and ontologies for semantic integration. In the RM-OA, this domain is split in two: the Integration Domain and the Mediation and Processing Domain.
- the *Composition Domain* supports the intelligent selection and combination of services. This includes workflows, semantic metadata for the discovery of components, and the storage of additional attributes such as provenance and version information. This is a specialised part of the RM-OA Mediation and Processing Domain.
- the *User Domain* provides domain-specific presentation environments for the control and monitoring of tasks and tools to support community collaborations. This includes a primary LifeWatch portal incorporating workflow management tools, but also domain-specific portals. In the RM-OA, this is called the User Domain as well.

The classification is refined in Figure 1 to capture service layers in accordance with ISO 19119. The Infrastructure Domain spreads over several layers, the *Communication Services*, the *Information Management Services*, the *Processing Services*, and the *System Management Services*, while the Composition Domain corresponds to the *Workflow Services*. The service bus represents the standardised mechanisms for connecting applications to the infrastructure as well as the services to each other. The boxes inside the service

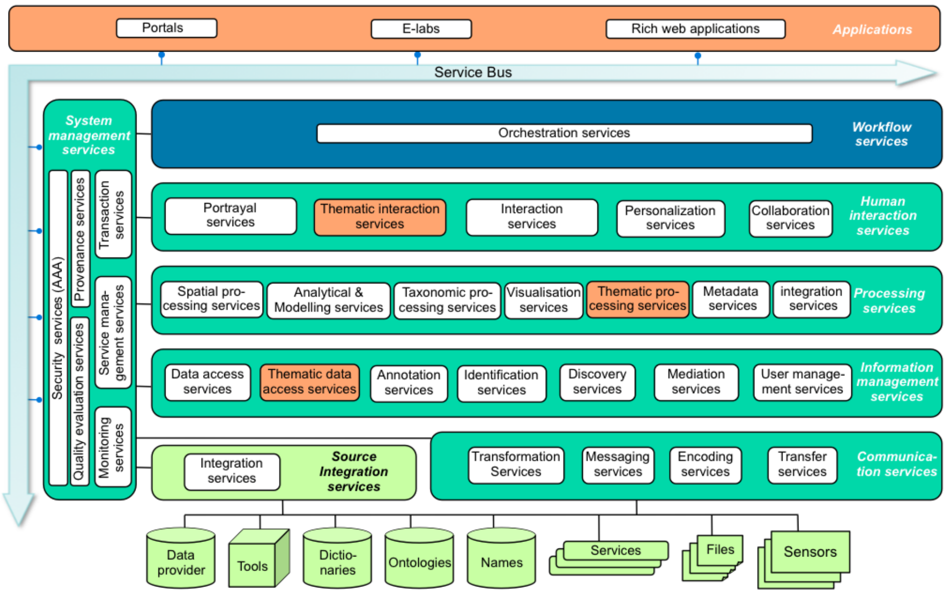

Figure 1: Layering LifeWatch Service Types according to ISO 19119

layers represent a sub-categorisation of layers within the service taxonomy. Thematic extension service groups or services represent particular application domains as provided by users.

An additional service group, the *Source Integration Services*, has been added to those defined by ISO 19119 due to the important role these play for the infrastructure. This group is settled directly over the resources to indicate that resources have two interfaces: one conforming to the LifeWatch models to be used by other services within Life-Watch and one representing the proprietary characteristics of the resources outside the scope of LifeWatch (e.g. for use by non-LifeWatch services).

The layered service groupings indicate broad dependencies between services, with services from the higher layers using services from the same layer or from the layers beneath. For example, a processing service may call an access service to obtain input data, which may in turn call a transformation service to translate the data to the required for-mat. The System Management Services are orthogonal to the other service groups, as they are used at all horizontal layers and use or are used by the other services.

Services will be provided by many sources. Services such as the System Management Services are likely to be provided by the LifeWatch Service Centre while the stakeholders of LifeWatch will contribute others. Considering the heterogeneity of sources and activities, the LifeWatch Reference Model is meant to provide the normative reference for defining conformance with LifeWatch and for achieving interoperability between services.

5 Structure of the LifeWatch Reference Model

The LifeWatch Reference Model is a cornerstone for syntactical and semantical interoperability between all kinds of data and services.

The Reference Model lays down the rules for the specification of information models and services. Certain generic information models and meta-information models that are of general applicability for LifeWatch, and a number of basic services that are generic for many application areas will be specified. These may be further extended by additional (meta-) information models and services for particular ”thematic” application areas within LifeWatch. The choice and the definition of the Reference Model has been a key design decision in LifeWatch.

5.1 Design Principles

The Reference Model provides one framework for the specification of content (data and information models to be used by the network) and one for the functional operations occurring within the network in term of services.

Information models and service specifications come in two flavours - *platform-neutral* and *platform-dependent* specifications. The platform-dependent specifications and their implementations must conform with the respective platform-neutral specification. This is a basic requirement for interoperability, at least for all information models and services that are of general interest for all of LifeWatch. We shall speak of *Basic* or *Architectural Services* if referring to such services. LifeWatch will take advantage of the work done by ORCHESTRA in that the platform-neutral specification of many architectural services defined by ORCHESTRA will be (re-) used or modified according to specific needs of LifeWatch when these will evolve in the preparation or construction phase.

Nowadays services are often identified with Web Services but this may not remain the case. New paradigms, such as cloud computing, will emerging in future. LifeWatch will define platform-independent interaction models as a reference for service interaction. This includes abstract models of how to organise service networks. At present, LifeWatch will consider a combination of two platforms as the preferred technological basis: Web and Grid. LifeWatch will take advantage of W3C, ORCHESTRA, and of the Open Grid Services Architecture (OGSA) by establishing itself on top of emerging European e-Infrastructures (e.g. EGI).

LifeWatch will adapt and extend the ORCHESTRA Reference Model in two directions by

- introducing new generic information models, services and rules, if required, and by
- defining a Base Application Architecture with common biodiversity-oriented information models and services.

The LifeWatch Reference Model will define the level of conformance that all LifeWatch approved services and tools will have to comply with. A set of test cases and tools will be provided to support (semi-) automatic conformance testing. The specification and design process began with the preparatory phase of the LifeWatch programme. Significant analysis is taking place in the preparatory phase. This will continue into the construction phase of the LifeWatch programme.

5.2 Information Framework

LifeWatch is about biodiversity and biodiversity related data. Since biodiversity includes four levels of organisation (genetic, species, ecosystems and landscape) and related data can range from meteorological parameters to measurements of human impact, no single data model is sufficient to organise this complex information. Several, possibly overlapping, models are needed to cover multiple application areas. Nevertheless the LifeWatch Reference Model promotes standardised data and attribute types in its models. The following data types can be regarded as generic, since they can often be used by the different biodiversity application models.

Geospatial data contains information about the physical location of the data subject. Geospatial data may have attributes, whose types are standardised by the ISO 19100 series. Temporal ranges add an additional dimension to geospatially referenced data, which is particularly true when geospatial values change over time.

Taxonomic data When working with taxonomic data, it is important to know which model of taxonomy was used by the data creator. Taxon Concepts should provide unique specification of a species, restricting the identity of a species to a single published definition.

Abiotic data Much biodiversity data is used in conjunction with a wide range of abiotic data types describing the chemical and physical characteristics of the envionment in which organisms are situated. Such data includes: topological data, climatalogical data, soil chemistry data, hydrology data, temperature data (soil, air, sea), etc. to name but a few examples.

Metadata ISO 19115 and INSPIRE propose models for metadata that help to describe resources and provide quality information about those resources. ORCHESTRA does not specify meta-information attributes, but gives rules for the construction of meta-information models for specific purposes. Services should use these information models in their interfaces.

Following ORCHESTRA, the Reference Model differentiates between the terms data and information: Data represents measurements, facts, concepts or instructions in a formalised manner suitable for communication, interpretation or processing by humans or by automatic means, while when speaking about Information, a meaning has been assigned to data by means of the conventions specified by LifeWatch to be used in their representation.

The LifeWatch Information Framework further distinguishes between two aspects of information:

- *Primary and derived information* (including metadata) related to biodiversity research
- *Meta-information*, that is: descriptive information about available information and resources with regard to a particular purpose (i.e. a particular mode of usage)

The definition of meta-information stresses that the need for meta-information arises from particular tasks or purposes (like catalogue organisation or data storage), where many different services and data objects must be handled by common methods and therefore should have common attributes and descriptions. For example, for the purpose "Discovery", related to catalogue organisation, attributes such as: `taxonomic_species`, `spatial_location`, and `time_frame` are the only attributes used, while other properties of a feature of type `occurrence_record` like `date_of_creation` may be irrelevant for that purpose. In that meta-information reflects only a part of or may be extracted from the information of a data object. Meta-information is purpose or context dependent in that what is meta-information in one context may be information in another context and vice versa.

The Reference Model will provide meta-information models for purposes such as "Discovery", "Identification", "Access", "Storage and Invocation", "Integration", "Orchestration", "Mediation", "Personalisation", "Provenance", "Collaboration", "Data capture", and "Authentication, Authorisation, and Accounting" (AAA) at a generic level that may be extended with domain specific attributes. This approach allows providers the flexibility of restricting information about resources to the particular scope of their usage.

The notion of meta-information is a crucial concept borrowed from ORCHESTRA. It avoids being forced to define one monolithic meta-model but allows to define a number of small meta-models reflecting the particular context or "purpose" thus increasing flexibility and relieving in particular data providers from the burden of generating sophisticated meta-data not needed for their purposes.

The components of the LifeWatch Information Framework and their reference to the different levels and aspects are captured in Figure 2 that is based on the Orchestra Information Model Framework (RM-OA V2, Section 8.6.2).

The picture is divided into several parts.

The *information aspect* is concerned with the information related to the biodiversity domain. The domain ontologies provide semantics to the information specified on the other levels. Several ontologies may coexist. They are defined and shared by the user communities. The Biodiversity feature set subsumes all the biodiversity information represented as features according to the application schemas.

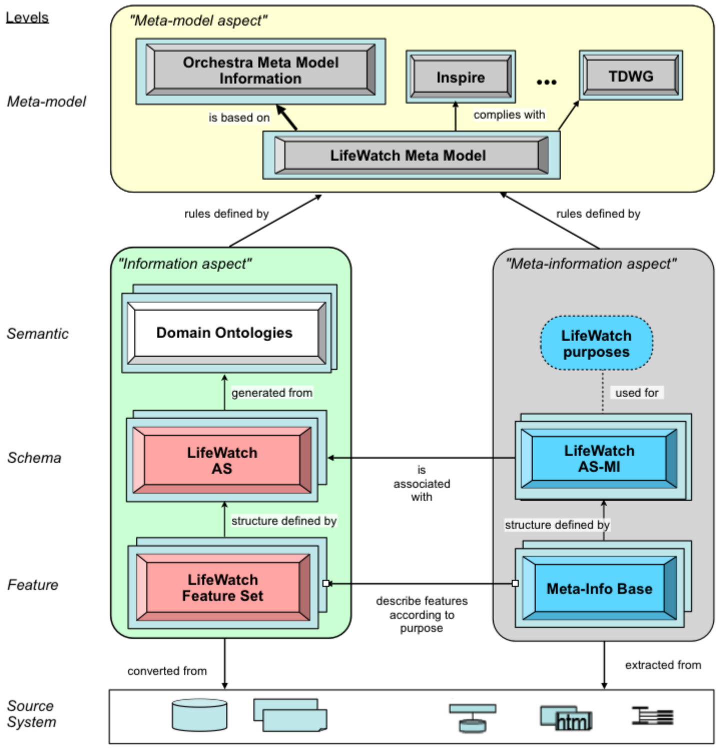

Figure 2: The LifeWatch Information Framework

The *meta-information aspect* is concerned with the meta-information models reflecting each a particular purpose. The structure of meta-information is defined by Application Schemas for Meta-Information Models (LifeWatch AS-MI). The Meta-Info Base is the store for meta-information. It contains information that describes features in the form of the Biodiversity Feature Set according to a well-defined purpose.

It is worth noting that purposes, hence meta-information models, reflect more or less an ICT perspective on the LifeWatch Architecture while the "information aspect" reflects the biodiversity domain, hence the user perspective.

The *meta-model aspect* captures all the normative rules to be applied when constructing the entities on the other levels.

5.3 The Service Framework

The Services Framework provides the basis for the specification of LifeWatch services, the service model, and a model for service chaining.

The service model is based on the fundamentals of Service Oriented Architecture defined by the interaction between service consumers and service providers within a distributed system. The interaction between consumer and provider is performed by service requests, service responses and service exceptions. This leads to the following definition of a service and its constituent parts.

Services are determined by their interfaces that consist of operations, according to ISO/DIS 19119. A *Service Type* is defined by the specification of the externally visible behaviour of a service through its interfaces. The service model provides rules for the specification of service types with the target of providing the syntactic and semantic interoperability between services, source systems and applications. The service model considers therefore two levels of specification of service types:

- Platform neutral: an abstract description of the services and an abstract specification of their interfaces (in UML).
- Platform specific: an implementation of the services and interfaces in a platform specific paradigm e.g., Web Services.

The ORCHESTRA Reference Model needs little modification with regard to the service framework in order to derive the LifeWatch service model. Only some general biodiversity oriented services need to be added to it to extend it into the LifeWatch domain. As for information models, the service models is model is adapted from the ORCHESTRA Framework for Services and is organised into the different levels and components (similar to Figure 2).

A mandatory requirement for LifeWatch Services is that they provide the *ServiceCapabilities* interface. The *ServiceCapabilities* interface has one operation *getCapabilities* that informs the client about the capabilities of a Service instance.

Architectual services are generic services that cover general infra-structural needs of LifeWatch, as opposed to services that emerge from the needs of particular applications or application areas. *LifeWatch Basic Services*, provide the "back-bone" of LifeWatch.

Particular application areas may develop services of particular interest for the respective areas. These are referred to as *LifeWatch Thematic Services*.

The list of LifeWatch Basic Services shall be specified as platform independent services in the preparatory phase while LifeWatch Thematic Services will in general be defined and implemented during the construction phase.

Abstract and implementation specifications for existing Architectural Services will be used where possible in order to reduce the effort required for development of new LifeWatch specifications, and to improve interoperability with external systems. The Open Grid Services Architecture (OGSA), for example, provides services for Authentication and User Management, in addition to other useful services such as Job Execution, Data Access, and Accounting. Additional Basic Services will be derived from the biodiversity domain and added to this set.

6 Towards Integration of External Source Systems

As indicated above, source system integration refers to the transformation of a non Life-watch source system into a LifeWatch Service Instance (within the LifeWatch service network, which means conform to the meta-models and interfaces). Due to the heterogeneity of External Source Systems, one cannot expect to define a service type with predefined interfaces that covers all these systems. Hence the term *Source System Integration Service* is only used as a generic name for the class of services serving this purpose. Source System Integration Services are to be implemented by the providers of the resources, as part of the admission procedures for joining the LifeWatch infrastructure. ORCHESTRA proposes (and LifeWatch adopts) the following strategy.

- If an External Source System can be treated as an instance of an existing LifeWatch service type, the interfaces of this service type shall be implemented. Otherwise a new service type shall be specified using existing LifeWatch interface types.
- If the operations of the External Source System do not comply with those of any given interface, either an available interface type can be extended or a new interface type can be created.
- Provide a mechanism to extract the meta-information needed by LifeWatch and publish it in the appropriate LifeWatch catalogue.
- Supply /ServiceCapabilities /interface.

Typical external source systems are those providing a TAPIR interface, which is a TDWG protocol for the exchange of structured data widely used for biodiversity data (e.g. used by GBIF providers). Figure 3 depicts the scheme for using a OccurrenceFeatureService as

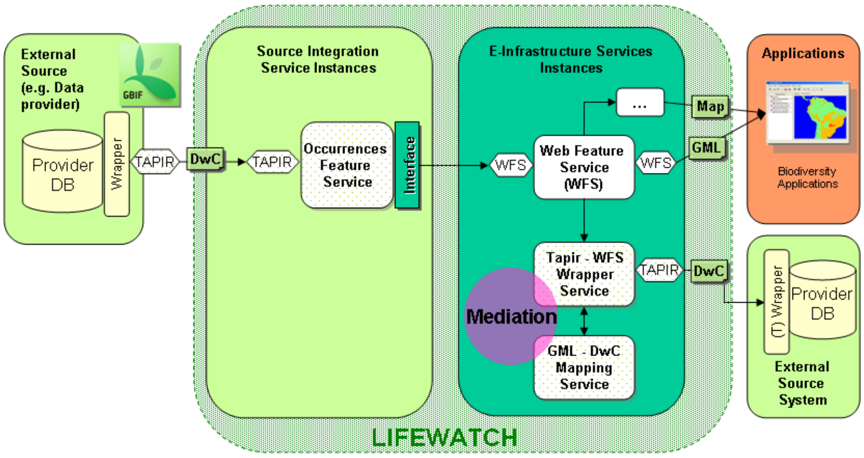

Figure 3: Integration of External Source systems

a source integration service.

The OccurrencesFeatureService which implementes the interfaces of a FeatureAccess service type, providing the infrastructure with a standardised OGC compliant mechanism, namely the WFS protocol, to access data through this service. At the other end, the OccurrenceFeatureService implements methods to extract the relevant meta-information and the data from the data provider using the TAPIR protocol and the DarwinCore data format. The service implements a wrapping mechanism, that allows to offer the data as a feature model using the GML data format, or, as an alternative offeres the context or meta-information as a GML document, providing also a reference to the original DarwinCore file.

An important issue for LifeWatch is that service consumers also want to be served using external systems; for instance data providers, which wants to enhance the quality of their data through using the value adding services of the infrastructure. Those can use either a semantic based transformation service, which wrapps LifeWatch information models into a specific format (e.g. DarwinCore) or the original resources, if available in the format. For those clients also a TAPIR interface should available through LifeWatch.

7 Status and Ongoing Work

The LifeWatch Reference Model is available in Version 0.2 [4]. The presentation follows that of the ORCHESTRA[8] in that it is structured in terms of viewpoints. At present, the "enterprise", "information", and "computational" viewpoints are elaborated.

The enterprise viewpoint (on which the material in Section 3 is based) has been finalised. It reflects the planning objectives, strategies and policy issues and it reflects the biodiversity research viewpoint of the Infrastructure.

Sections 5.2 and 5.2 relate to the information and computational viewpoint. The information viewpoint specifies the modeling approach and a core ontologyto all categories of information of the LifeWatch Architecture while the computational viewpoint specifies the LifeWatch Interface and Service Types that aim at improving the syntactic and semantic interoperability within the LifeWatch Architecture.

In both cases, the general syntactic formats and rules are established and, in case of the information framework, candidates for Meta-data elements are listed that are derived from standards. The next step will be to derive meta-information models from these standard for the LifeWatch purposes as defined above. LifeWatch will adopt the platform-neutral and the platform-dependent service type specifications, thus avoiding a considerable amount of work. However, careful inspection of these specification, to be done in the next couple of month, will reveal whether modifications are needed. Further, it will be investigated whether specific LifeWatch Architectural Services will be needed and, if so, they will be defined.

The elaboration of the "engineering" and the "technology viewpoint" is still missing. The

[8] that is based on the Reference Model of Open Distributed Processing (ISO/IEC 10746) [5]

engineering viewpoint specifies the mapping of the LifeWatch service specifications and information models to the chosen platforms and it considers the characteristics and principles for service networks while and the technology viewpoint specifies the technological options for platforms, their characteristics and its operational issues. Both viewpoints will be added in future versions of the LifeWatch Reference Model document.

Further, we have not touched upon the advanced topics of workflows, provenance, semantic mediation, etc. . All these subjects are far from being settled in terms that no commonly accepted standards are available. LifeWatch will take a pragmatic stand in that we will rely on reasonably well established approaches but this needs careful inspection and some experimentation. We expect that our findings will be presented as part of the engineering and the technology viewpoint.

8 Conclusion

The definition of a reference model is a major step in the definition of the ICT infrastructure of LifeWatch. The ORCHESTRA Reference Model has provided an excellent starting point on which to base our work. Though originally targeting risk management, the model is generic and we expect that it will (and recommend that it should) be used by other European initiatives.

Acknowledgements. We acknowledge the ORCHESTRA team for the bulk of work accomplished that provides a solid basis for our efforts.

References

[1] Reference Model for the ORCHESTRA Architecture (RM-OA) V2 (Rev 2.1), `http://www.eu-orchestra.org`

[2] ISO/IEC 10746-1:1998 (E). Information technology - Open Distributed Processing - Reference model `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=20696`

[3] ISO 19136:2007. Geographic Information – Geography Markup Language (GML) `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=32554`

[4] LifeWatch Reference Model (Version 0.2), for a copy `http://lifewatch.eu/images/stories/PDFs/lw_referencemodel.pdf`

[5] ISO/IEC 10746-3:1996 Information technology – Open Distributed Processing – Reference Model: Architecture `http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=20697`