BACHELOR THESIS

IMPEx - A Service-oriented Space Plasma Physics Virtual Observatory

carried out at

CAMPUS 02
GRAZ
FACHHOCHSCHULE DER WIRTSCHAFT

Course Programme
Information Technologies and IT-Marketing

By: Florian Topf
PID: 1010319048

Graz, June 20, 2012

....
Signature

Declaration of Authorship

I hereby declare, that I have written this thesis without any help from others and without the use of documents and aids other than those stated, that I have mentioned all used sources and that I have cited them correctly according to established academic citation rules.

....
Signature

Acknowledgements

The RTD work described in this thesis was accomplished within the EU FP7-SPACE project IMPEx (Integrated Medium for Planetary Exploration, project number 262863) with the support of technical expertise from the collaborating FP7 project EuroPlaNet. The author is thankful to all contributions made to the IMPEx project, in particular the services provided by the data analysis tool AMDA developed at IRAP, Toulouse and the HMM simulation environment established and maintained by FMI, Helsinki. The overall process of elaborating this thesis was assisted by the IMPEx project coordinator Dr. Maxim Khodachenko and project manager DI Tarek Al-Ubaidi from the IWF (Institut für Weltraumforschrung/Space Research Institute) of the Austrian Academy of Sciences, in Graz. Additionally, Prof. DI Manfred Steyer gave guideance with his didactic role at the Campus 02 and as supervisor of this thesis.

Florian Topf

Graz, June 2012

Abstract

The FP7-SPACE Project "Integrated Medium for Planetary Exploration" (IMPEx)[1] is funded by the EU to create an interactive framework, where existing archived space mission data is interconnected with numerical models. The main aims of this system are to test results from simulations of planetary plasma and magnetospheric environments against in-situ measurements, to improve the accuracy of the models, to simulate planetary environments in exceptional conditions and to help improving related instruments for future missions. The initial focus of the project is on increasing the efficiency of discovering and exploiting existing space mission data resources, simulation codes and their respective computational infrastructures. The currently designed service-oriented infrastructure of IMPEx is composed of existing web-based data analysis and visualisation tools and online access to interactive computational models including databases of archived simulation runs. The distribution of these resources and services requires a clear definition of a common communication protocol and similar interfaces between the different tools. This thesis will give an introduction to service-oriented infrastructures in general and the project structure of IMPEx including its workpackages. The proposed distributed system will also be compared with the well-known "Virtual Observatory" (VO) paradigm in order to investigate similarities. A selection of science cases for the envisaged IMPEx environment will be presented in order to extract the systems user requirements. Furthermore possible architectural concepts like authentication mechanisms, protocols, interfaces and simulation data models will be described and finally evaluated.

---

[1] IMPEx: `http://impex-fp7.oeaw.ac.at`

Contents

1. Introduction

The EU FP7-SPACE project "Integrated Medium for Planetary Exploration" (IMPEx) was initiated in course of scientific collaboration between European institutions working in the field of planetary sciences and space plasma physics. It is recognized as the successor of the work package "European Modelling and Data Analysis Facility" (EMDAF) of the FP7 project EuroPlaNet[2] which is defined in details in Blanc et al. (2008, pp. 38). The following sections will give an overview about the motivation for initiating the project IMPEx and the primary focus of the technical aspects presented in this thesis.

1.1. Present Status and Motivation

There is a measurable trend in the development of interoperable information systems in the scientific field of planetary sciences, which is stimulated with studies and evaluations in the EU FP7 Project EuroPlaNet. The transformation of big data centres which are archiving and preserving data into "Virtual Observatories" (VOs) which are providing added value services to the scientists as described in Topf (2012, pp. 1-4) is the main aspect of this trend. Given the fact that most recources are distributed over global networks, the "Service Oriented Architecture" (SOA) paradigm plays a significant role in scientific research. The separation of concerns regarding the handling of data coming from a variety of scientific observations was a logic step towards an infrastructure composed of services within the European science community. Furthermore the flexibilities of Web services as self-existing entities embedded in complex processes preserved the unique purpose of each functionality.

While in the field of observational astronomy with IVOA[3], the information systems are already well advanced, giving the user a lot of possibilities to harvest and manipulate data, the planetary science community is still engaged with defining the baseline of Web Services: An appropriate XML vocabulary which is able to describe the different types of data (see Cecconi et al., 2011). The studies in Topf (2012, pp. 38-42) were aimed at definition and implementation of simple Web services, which are

---

[2] EuroPlaNet: `http://www.europlanet-ri.eu/`

[3] IVOA: `http://www.ivoa.net/`

providing homeogenous access to data by describing it with XML metadata. The establishment of Web services within the work package “Integrated and Distributed Information System (IDIS)” of the EU FP7 project EuroPlaNet has shown, that the implementation of industry standards from the “Web Service Architecture Stack" by the W3C is of vital importance. With this approach, a sustaining environment and the re-usability of services can be achieved due to the “loose coupling” of the various components from different scientific fields.

In particular the “Automated Multi Dataset Analysis Tool” (AMDA) described in Topf (2012, pp. 38-44) provides a complete Web service infrastructure with all fundamental *semantics* and *vocabularies* needed for description of space plasma measurements. AMDA uses the “Space Physics Archive Search and Extract” (SPASE)[4] XML data model. Its flexible *AMDA-Registry* also provides access to the data from other web-based informationsystems. The service-oriented AMDA tool provides an integral part of the IMPEx project as one of the integrated data analysis tools, which will be interconnected in the developed IMPEx architecture. Besides of space mission measurements, the project will describe in a similar XML language simulation data coming from different magnetospheric models already used in the EMDAF work package of EuroPlaNet.

The IT framework of IMPEx will go a step beyond of providing a simple online data access and will transform complex research workflows into a automated Web service process. The scientific reason for combining simulation and space mission data is given in Khodachenko et al. (2009, p. 3) by the fact, that planetary phenomena can be interpreted much better by comparing numerical models based on theoretical principles with in-situ measurements. Furthermore the FP7-SPACE call provided an appropriate funding scheme for fostering the foundations of space sciences and technology by improving, modernising and streamlining existing scientific information systems (The European Commission, 2009, p. 6). The following chapters will introduce basic approaches of service-oriented infrastructures and present the general work package structure and actual progess in IMPEx project on the currently undergoing architectural design phase.

1.2. Conceptual Formulation

Based on the present status and the general motivation of the IMPEx project, the author has formulated a major research question, which will be answered in course of evaluation of user requirements and design aspects of distributed information systems:

---

[4] Space Physics Search and Extract consortium portal: `http://www.spase-group.org/`

"Which technical approaches do exist in the scientific community, to integrate existing web-based scientific tools into a distributed service-oriented infrastructure?"

1.3. Objectives of this Thesis

According to the major research question, the three main objectives of this thesis were established:

1. Presentation of potential scientific use cases for the EU FP7-SPACE project IMPEx and definition of user requirements for the foreseen service-oriented infrastructure.
2. Collection and evaluation of appropriate architectural concepts coming from the scientific community (e.g. IVOA) as well as correlating industry standards from the W3C.
3. Development of a deeper understanding of the "Virtual Observatory" paradigm and identification of possible critical issies with highly distributed software concepts and service-oriented infrastructures.

1.4. Approach and Methodology

The technological developments of this thesis are the result of the studies in the Research & Technical Development working groups of IMPEx in the fields of:

- Service-oriented infrastructures
- Web service processes and messaging protocols
- XML vocabularies for simulation data
- Relevant specifications of IVOA and W3C

1.5. Composition of this Thesis

This thesis is composed of the following theoretical chapters:

- Introduction to service-oriented infrastructures in general
- Presentation of the aims and goals of the EU FP7-SPACE project IMPEx.

- Description of the IMPEx work packages and the participating tools.
- Explaination of the "Virtual Observatory" paradigm and it's application to an environment of existing web-based tools.

The practice-oriented chapters include the results obtained within the work package "Data and Models Environment" (DaME) together with the project management team where the author is engaged as technical expert:

- Definition of scientific use cases and formulation of user requirements for the proposed IMPEx infrastructure.
- Suggestion of possible architectural concepts for IMPEx infrastructure with the focus on advanced Web service patterns (e.g. aggregation and composition of Web services)

1.6. Work done by the Author

The author of this thesis has been working on following content as part of the dedicated IMPEx Research & Technical Development working groups and the project management team:

1. Definition of scientific use cases according to the "Description of Work" (DoW) of IMPEx.
2. User requirements analysis (according to established standards within the scientific community)
3. Proposition and critical review of preliminary architectural design concepts.

2. Service-oriented Infrastructures

According to studies in Topf (2012) the scientific community is highly engaged in identifing data resources and analysis tools for their research needs. This work is done in order to improve the information systems needed for specific scientific applications. Based on the requirements for complex research processes, these tools and resources are to be transformed into web-based information systems which can be consumed via typical graphical user interfaces or via machine-readable interfaces. The web-based systems evolving within EuroPlaNet-IDIS already provide access to Web services with simple message exchange patterns such as *request/response* and *publish/subscribe*. The main idea behind the EuroPlaNet-IDIS architecture is to provide *registries* where single point-to-point communication with data exploitation systems can be conducted. The data provider publishes the provided Web services to the registry and the consumer may discover it by querying the registry interface.

Following the standards of the "Web Service Architecture Stack" illustrated in Booth et al. (2004) the cornerstones for realisation of a service-oriented architecture are given by four layers, which provide standards for Web service communication based on the XML metadata language. The *service contract* provided by EuroPlaNet-IDIS defines all interactions between consumers and providers via standardized interfaces and is integrating the four Web Service layers *communication*, *messages*, *descriptions* and *processes* (Rosen, Lublinsky, Smith, & Balcer, 2008, p. 50). *Service descriptions* made available through *registries* ensure, that a consumer gets the necessary information to interact with the service. This approach on one hand declares the service as "loosely coupled", dependencies are kept at a minimum between related services, but also provides a certain abstraction of the service, since only essential metadata is provided to the consumer (Erl, 2007, p. 155). Finally, the system establishes *service discoverability* of Web services, which provides the user an easy platform to look for needed resources and tools.

Beyond the realisation of *service contracts* between providers and consumers, the following two key functionalities of Web services are of importance in successive projects of EuroPlaNet-IDIS to make practical use of the elaborated service-oriented architecture:

- service reusability
- service composability

These two key funcionalities together with *service discoverability* ensure, that a service with a unique purpose, can be useful and integrated in more than one scientific process and potential service developers are able to search in *registries* if a certain functionality is already existing (Weerawarana, Curbera, Leymann, Storey, & Ferguson, 2005, chap. 8.2). Furthermore, *service composability* according to Erl (2007, p. 388), gives a possibility to connect Web services together and aggregate them into a particular task. This approach clearly goes beyond the simple point-to-point communication achieved within IDIS and is the main technological research topic for the IMPEx project as a part of the *processes layer* of the "Web Service Architecture Stack". It goes hand in hand with the previously achieved "separation of concerns" which encourage system designers to divide a complex scientific problem into several smaller tasks. The focus of this thesis is to identify different services implemented in course of the studies in Topf (2012) where repeated aggregation is necessary to solve complex scientific problems. Point-to-point communication alone is not flexible enough to maintain and handle the aggregation of services.

The *service composability* provides management capabilities for the coordination of smaller tasks within a process and gives the *service reusability* an additional purpose. The coordination of Web services within a composition is maintained by a coordinating service or controller, which is also exposing the whole aggregated Web service process to potential consumers (Erl, 2004, p. 358). Note, that the coordinating service itself almost always has only one particular process to manage and *service reusability* usually is not the primary goal of such a service. When developing service compositions, one has to define the workflow logic of the process which is to be automatised (Erl, 2007, p. 388). The flow of data and the impact of runtime conditions is also necessary to identify the different temporary roles a service can have within a composition. Furthermore each service has to be classified permanently according to their application logic provided (Topf, 2012, chap. 2.3).

The *composition controller* represents the top of the hierarchy in an aggregated Web service. It is identified as such, if its realized *service capability* is able to invoke other services (Booth et al., 2004). *designated controller*, which means that this is its only role within the process (Erl, 2007, p. 400). Web Services participating in a composition are usually referred to the *composition members* which need to have flexible *service contracts* by means of the data operations made available trough their *service capability*. According to Hass and Brown (2004) this also refers to the set of tasks identified by the service provider which implicitly assigns the service with a specific *service role*. Depending on the implementation of *composition members* this role can also permanently classify the service with a *service model*. In order to obtain an abstraction within the involved *service models* three types of services are introduced among *composition members* which are also common for primitive Web service implementations (Erl, 2005, chap. 5.2):

- Entity services
- Task services

- Utility services

Entity services build the fundamental part of compositions giving a possibility to describe and provide entity-centric services within a process (Erl, 2012a). These entities can be classified as objects in classical terms of programming. Services related to the objects may in particular include computational services related to a simulation model in the project IMPEx. A *task service* is usually more focused on a specific part of a whole process and is therefore more specialized and limited in reusability (Erl, 2005, chap. 9.4). On the contrary a *utility service* is highly resuable and not tied to one specific process. In relation to the created infrastructure within IMPEx such services may include the simple data access Web services provided by the AMDA system. More examples of *utility services* are provided in Erl (2007, p. 46). In context of a service composition, a specialised type of task service is represented by the *composition controller*. Figure 2.1 shows a possible configuration of an aggregated Web service environment where the controller is represented as an "orchestrated task service". The particular service layers are characterized by their associated *service model*.

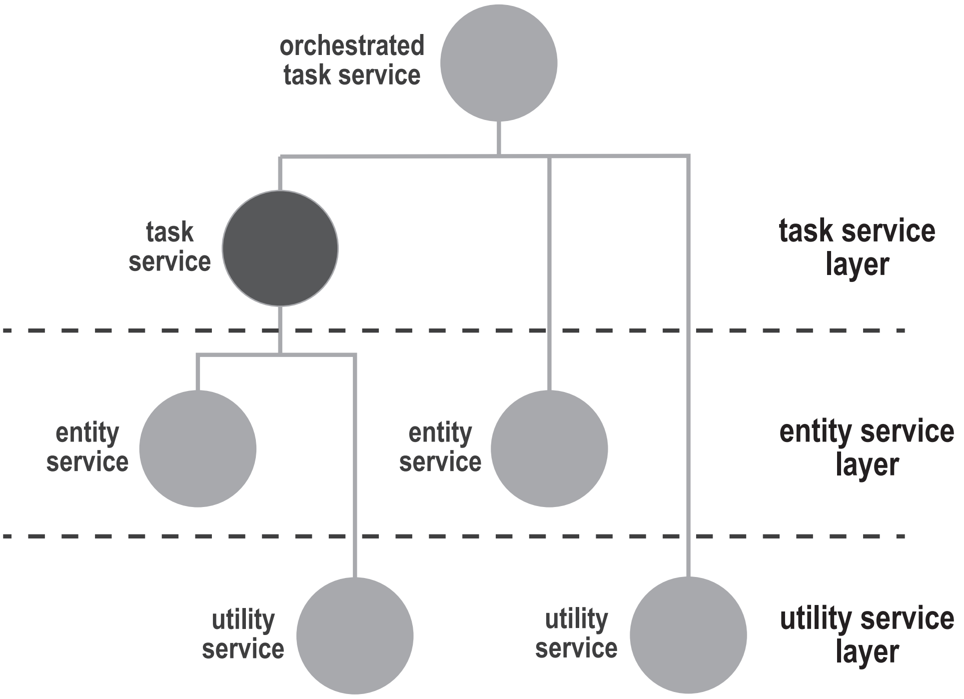


Figure 2.1.: Service models within an aggregated Web service environment (Erl, 2007, p. 429)

After having identified the flow of data between Web services within a process and classification based on the purpose of these participating Web services, a distinct message-path is formed through the planned service-oriented infrastructure. This message-path defines the interaction among all contributing services. It represents a *service activity*. Since all Web services described in Topf (2012) already provide primitive service activities with message-exchange patterns, the focus task for IMPEx is to set up complex service activities with more than two collaborating service providers and/or consumers within one process.

The following sections will present a selection of design principles to achieve such a composition with a coordinating entity and participating services assigned to one particular role resulting in a so-called *composite service* (Rosen et al., 2008, p. 274). Since "tight coupling" is not the aim of a complex service-oriented infrastructure (Rosen et al., 2008, p. 142), the *interaction processes* are described which are able to act as a *service intermediary* to transform messages and delegate them through a service composition. These design principles are driven by a set of so-called *WS-* extensions* which are based on the technologies of the "Web Service Architecture Stack" in Booth et al. (2004) such as the "Simple Object Access Protocol (SOAP), the "Web Service Description Language" (WSDL) and "Universal Description Discovery and Integration" (UDDI) (see also Topf, 2012, pp. 19-25).

2.1. WS-Coordination

In order to organize *service activities* within a composed Web service infrastructure, the *WS-Coordination* framework specified by OASIS[5] provides a *coordinator service* handling the commonly assigned task with a collection of coordination protocols. The *coordinator service* model consists of two component services, namingly the activation service and the registration service (Newcomer & Robinson, 2007). The framework enables to manage, save or update and distribute context information, which is shared among the registered participating services. The context information itself identifies the operations and messages needed to achieve a particular task and are defined using common WSDL interfaces (Weerawarana et al., 2005, chap. 11.3). The complete hierarchy of all XML elements and associated XML schemas forming the protocol can be found in Newcomer and Robinson (2007).

A *WS-Coordination* composition is usually called by an initiating task service asking the activation service to provide the coordination context. Each Web service owning the context information is either able to register itself at the registration service or to invite services for participation in the coordinated task. The process of task execution can be controlled by asking the coordinator service whether the participating services are completed or still running (Erl, 2012b). There are two commonly used protocols within the *WS-Coordination framework* to standardize the control of complex activities: *WS-Atomic Transaction* and *WS-Business Actitity*. The first protocol is implementing ACID[6] transaction principles. It provides a possiblity to conduct atomic operations, which means that every participating service must succeed to complete the whole operation sucessfully. The data flow must also be consistent during the whole process and multiple parallel transaction should not interfere each other (Erl, 2005, chap. 6.4). A successful transaction is also considered durable, since it can not be changed due

---

[5] OASIS - Organization for the Advancement of Structured Information Standards `http://www.oasis-open.org/`

[6] Atomic Consistent Isolated Durable (ACID) transactions (Rosen et al., 2008, p. 382)

to consecutive errors. The *WS-Business Activity* protocol is more "loosely coupled" by means of atomicity and isolation. Participating services directly expose their completed work to the *coordinator service* and some results might not be included in the overall process outcome delivered to the initiating task service. Figure 2.2 shows the complete environment of the *WS-coordination* management framework available for service-oriented infrastructures.

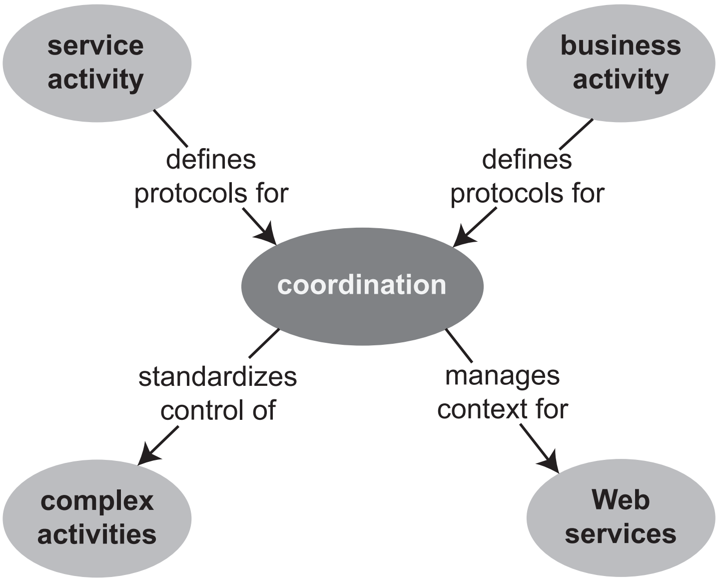


Figure 2.2.: Coordination within a service-oriented infrastructure (Erl, 2005, chap. 6.3)

2.2. WS-Orchestration

The definitions within *WS-Coordination* ensure that the existing service interfaces can manage a coordination context among a set of composed service activities. The separate actions of creation, registration and initiation of the coordination context are fully integrated and centralized with the *WS-Orchestration* framework (Erl, 2004, p. 362).

An *orchestration engine* is implemented in a service composition if there is the need to encapsulate *service activities* with specific rules into one distinct workflow (Rosen et al., 2008, p. 276). As the name already indicates, this is accomplished by a coordinating entity which acts like a conductor of an orchestra. The main idea with this approach is to easily connect different processes without redeveloping or changing the existing implementations. The workflow logic of the composed processes is completely independent from the components and therefore can be maintained much more easier than being directly incoporated within in the participating services (Erl, 2005, chap. 6.6). *WS-Orchestration* provides all the common context management functionalities like data transformation and transaction management from traditionaly message-oriented middleware like *WS-Coordination*. As indicated in Erl

(2007, p. 431) the "Web Services Business Process Exectution Language" (WS-BPEL) provides a standardized way to describe workflow logic similar to WSDL. This standard follows the rules of the *WS-Business Activity* protocol and hence is not tied to atomic transactions. Figure 2.3 shows, that *WS-Orchestration* builds up the heart of service-oriented infrastructure patterns and orchestrated services are a key ingredient for *WS-Choreography* described in the next section, where they act together as one aggregated service.

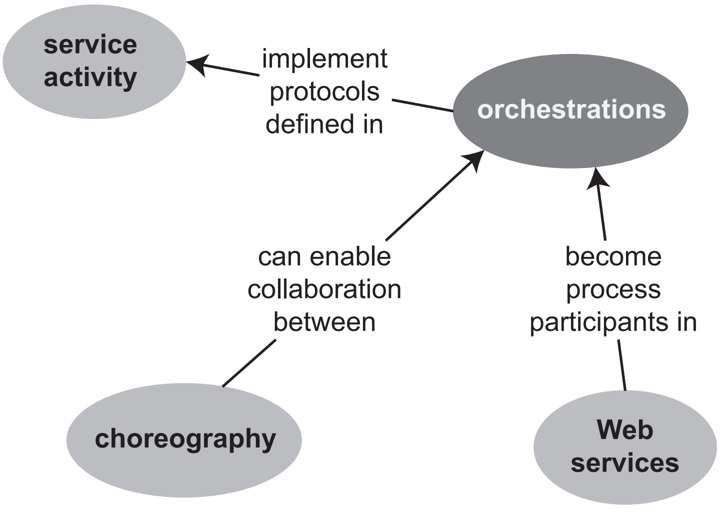


Figure 2.3.: Orchestration within a service-oriented infrastructure (Erl, 2005, chap. 6.6)

A simplified approach of *WS-Orchestration* with its centralized processing middleware is represented by the *hub and spoke* design pattern (Erl, 2004, p. 372). It provides a central hub which is able to implement all message-exchange patterns used within the EuroPlaNet-IDIS activity and the AMDA tool. As a step further, a possible implementation called "Simple Application Messaging Protocol" (SAMP) defined by the "International Virtual Observatory Alliance" (IVOA) is evaluated in the frame of IMPEx which will be described in chapter 7. It is basically designed for multi-directional messaging following the *publish/subscribe* pattern.

2.3. WS-Choreography

In contrast to the *WS-Orchestration* definition which focuses on the flow of data among participating services from a coordinating entity's point of view, the *WS-Choreography* is seen as an external observer concentrating on the exchange of messages between services. *WS-Choreography* is intended for inter-process collaboration, where each participating process must be aware of the common task to achieve, which operations have to be executed and how message interaction is organized (Rosen et al., 2008, p. 277). None of the collaborating *service activities* controls the overall process and therefore the process itself is not instanitated like with the implementation of atomic transactions in *WS-Orchestration*. All the collaborating services are assigned to a specific *service role* within *WS-Choreography* which is specified in their according WSDL descriptions. Each possible message exchange scenario between two *service roles* within this composition is defined as *service relationship* and *service channels* are providing the characteristics on where and how the messages are sent by a *service role* (Kavantzas & Burdett, 2004). One has to mention, that with these characteristics within a composition it is possible to decentralize the ownership of the collaborative process. This can be of particular interest if the participating entities are providing complex service activities which are reused for a variety of other tasks. Ultimately the *service interaction* in *WS-Choreography* is representing the complete logic implementing the actual message exchange.

Figure 2.4 shows, how the collaboration between two or more *WS-Orchestration* infrastructures, which are autonomous and encapsulated *service activities* within different organizations, is achieved with *WS-Choreography* (Erl, 2005, chap. 6.7). A practical implementation of this service infrastructure and its inter-process communication concept will be presented with the "Enterprise Service Bus" (ESB) pattern in chapter 7, which is of potential interest for the IMPEx project.

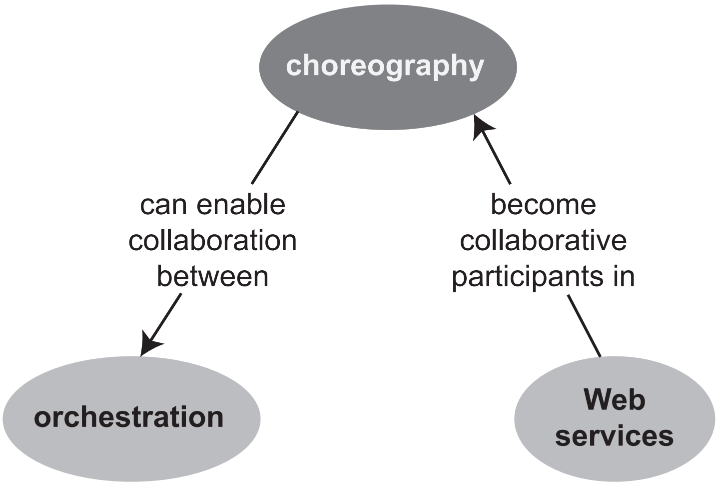


Figure 2.4.: Choreography within a service-oriented infrastructure (Erl, 2005, chap. 6.7)

In order to build up an infrastructure for IMPEx, with its distributed services and resources, one has to look at the information processes accopagning the scientific research. It its crucial to get an exact knowledge on the capabilities of each tool and database to be able to choose on appropriate infrastructural service technology as the best solution for the IMPEx system. Before going into the details of the project environment and considerating potential use cases to identify repeating processes in the related scientific research, the already well-known "Virtual Observatory" (VO) software paradigm will be introduced and compared with the general service-orientation concept.

3. The Virtual Observatory Paradigm

Recent studies in Topf (2012) have already revealed a distributed software architecture paradigm evolving within the planetary science community, called as "Virtual Observatory" (VO). The goal within these studies was to identify the low level requirements for development and deploying Web services within the frame of EuroPlaNet-IDIS according to specifications from IVOA and IPDA. These low level requirements included basic mechanisms to characterize services according to their purpose and a clear distinction between *service providers* and *service consumers* interconnected through a VO-middleware messaging layer (Arviset, Gaudet, & the IVOA Technical Coordination Group, 2010, pp. 5). Before focusing on how to connect the composed VO-*service activities* or VO-tools together within a service-oriented infrastructure and the corresponding high level requirements, the cornerstones of a VO are presented. According to Djorgovski and Williams (2005, pp. 2), the VO paradigm aims at assembling distributed data archives and services for both harvesting and processing scientific data. It is enabled by technological aspects from the "Web Service Architecture Stack" such as WSDL and driven by scientific problems which require an inter-disciplinary research. The foundation of such a service-oriented distributed environment is built with *registries* providing discovery mechanisms, which according to Topf (2012, p. 34) become a popular way to publish Web services within the planetary science community.

The increasing importance of adapting VO-standards and strategies to make scientific resources and tools available and interoperable for a wide spectrum of applications were recognized in the IMPEx project during the design of the overall infrastructure. The appropriate use of XML standards in previous projects such as EuroPlaNet-IDIS has defined a general path for the reuse, extension and composition of the available Web services. The "International Virtual Observatory Alliance" (IVOA) already has already investigated potential ways to implement *service composability* within their enhanced service-oriented framework for astronomical research. Following the principles provided by the *WS-* extension* specifications, different *service activities* were classified according to their specific role within the scientific community. The *VO-architecture* specification of IVOA described in Djorgovski and Williams (2005) distinguishes the services by defining the characteristics of their interaction with the consumer. Similar to the existing industry standards provided in chapter 2, three service types provide the middleware for communication between *service provider* and *service consumer* in the *VO layer*, which is illustrated in figure 3.1.

The three types of Web services in the IVOA architecture are as follows (Djorgovski & Williams, 2005):

- **Data Services** – provide the most fundamental service in space research, the access to datasets. These services are highly resuable and have their characteristics similar to those of *utility services*.
- **Registry Services** – establish *publish/subcribe* scenarios and *service discovery* processes.
- **Compute Services** – expose capabilities for processing and federation of data. This type is comparable with the functions of a *task service*, e.g. having the role of a *coordinator* or an *observer* within a composition.

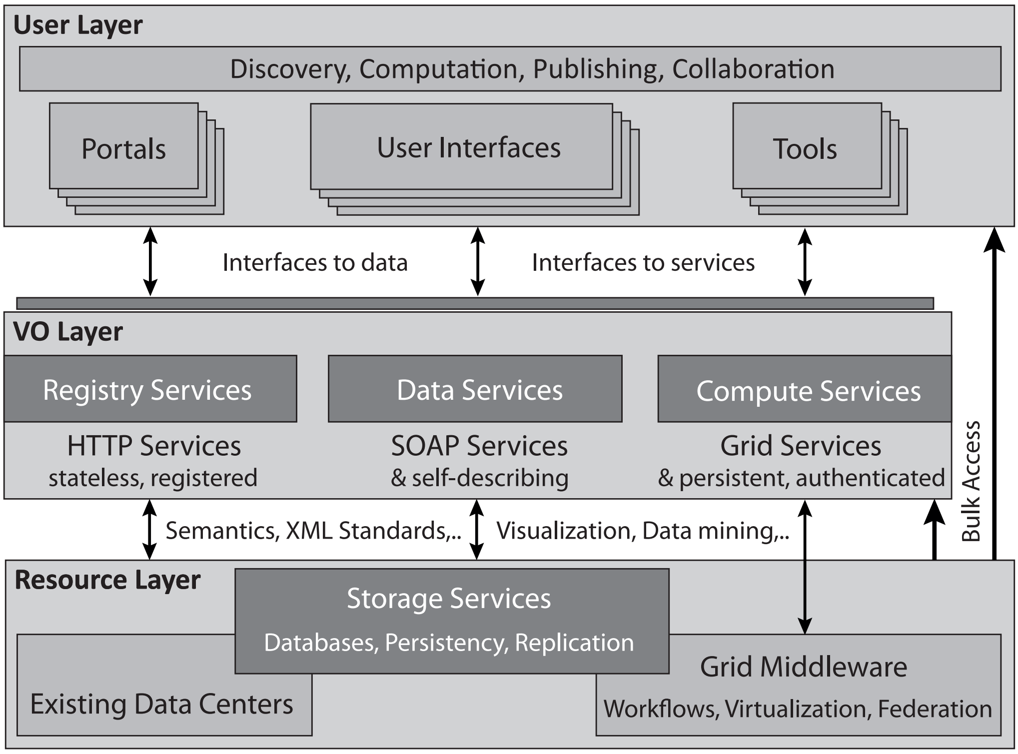


Figure 3.1.: Components of the Virtual Observatory architecture (Williams et al., 2004)

According to the characterization in Williams et al. (2004) *compute services* are part of so-called *grid services* which enable composability and reusability by standardizing repeated workflows, making them instantiable such as the processes in *WS-Orchestration*. Furthermore an appropriate *grid middleware*, as seen in figure 3.1, provides mechanisms for federating message interaction within complex service activities, such as *WS-Choreography*. It is important to note, that the IVOA-architecture also provides governance by means of security and authentication, which is needed when dealing with restricted and personalised data within interoperable research infrastructures. The following chapter will explain all the particular tools and resources which will collaborate within the IMPEx infrastruture to identify potential correlations with definitions in the VO paradigm and service-oriented infrastructures.

4. The EU FP7-SPACE Project IMPEx

The infrastructure of the "Integrated Medium for Planetary Exploration" project belongs to a type of multi-disciplinary "Virtual Observatory" (VO) build up around a set of numerical models for planetary magnetic field and plasma environments and related observational data. The four participants of IMPEx agreed upon one concrete goal, which is to be fulfilled by the end of the project according to the FP7 proposal (Khodachenko et al., 2009, p. 3): The integration of the simulated planetary data and related computational modelling services within existing space mission data analysis and visualization tools. The resulting framework must be standardized, easily maintainable, extendable and be made available to a broad scientific community.

4.1. IMPEx Project Structure

The IMPEx Project is organized according to the FP7 funding schemes as a "Collaborative project". Following the FP7 definitions, this type of project is classfied as a research project organized and maintained by participants from different countries, with the goal of establishing new products and technologies (The European Commission, 2007). The primary focus is on implementing demonstration interfaces and aggregation of common research resources already established within the relevant interest groups. The principle work programme of IMPEx is listed in Khodachenko et al. (2009, pp. 4-6) and is comprised of five work packages, each having a corresponding set of milestones and deliverables defined according to the goals of the project:

- **WP1 - Project Management**: Establishment of an operational management structure, communcation policies and document handling; Reporting of the project's progress to the EU and creation of internal controlling mechanisms.
- **WP2 - Data and Models environment (DaME)**: Development of standard interfaces and communication protocols between the participating services and resources; Extension of the resulting IMPEx toolset and their integrated observational databases with visualization & data analysis capabilities; Provision of interfaces to external infrastructures.
- **WP3 - Hybrid and MHD models (HMM) and WP4 - Paraboloid Magnetospheric Models (PMM)**: Development of interfaces and protocols for compu-

tational modelling services and simulation catalogues in order to interconnect with DaME; Computational modelling support of the infrastructure.

- **WP5 - Outreach and Dissemination (ODis)**: Development of the IMPEx Web site[7] and establishment of public outreach activities.

The following three work packages are representing the RTD branches and are the most relevant for the technical realization of the IMPEx infrastructure.

4.2. Data & Models Environment (DaME)

The work package 2 "Data & Models Environment" (DaME) as seen in figure 4.1 was created for providing all the required data standards, interfaces and communication protocols supporting visualization and analysis of data from observational databases as well as data generated by simulation codes provided by work packages 3 and 4. Furthermore DaME is designed for the establishment of discovery and search mechanisms with Web services for both types of data, in order assist the user in finding needed simulation model and experimental data (Khodachenko et al., 2009, p. 13). The data analysis and visualization services interconnected within DaME are 3DView Multimission[8], AMDA and CLWeb[9].

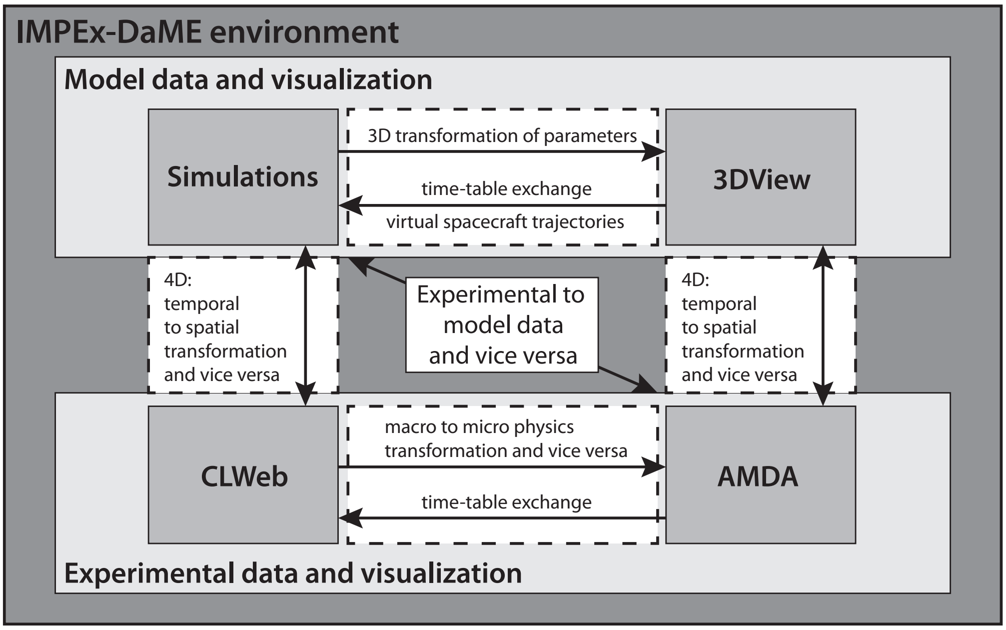


Figure 4.1.: Conceptional structure of the IMPEx-DaME environment, adapted from The IMPEx consortium (2012)

---

[7] IMPEx Web site: `http://impex-fp7.oeaw.ac.at/`

[8] 3DView Multimission: `http://3dview.cesr.fr/`

[9] CLWeb: `http://clweb.cesr.fr/`

3DView Multimission, as seen in figure 4.1 provides capabilities to visualize both simulated parameters from the modelling environment of IMPEx and observational data coming from AMDA in a 3D environment. AMDA and CLWeb on the other hand are integrating data analysis and 2D plotting for all types of data in IMPEx. There will be an interface within AMDA which is able to compare both types of data by searching observed data close to the input conditions of a simulation . An additional task of DaME is dedicated to the extension of the AMDA observational database, following the established deployment process described in Topf (2012, pp. 42-44). Further potential examples of data mining Web services can be found in André et al. (2009). Ultimately the IMPEx-DaME environment will establish a global IMPEx authentication concept which enables the user to exchange user created information such as time-tables among all tools in a dedicated user session.

4.3. Hybrid & MHD Models (HMM)

The work package 3 “Hybrid & MHD Models" (HMM) is part of the simulation environment of IMPEx. The solar system objects covered with the 3D quasi-neutral hybrid models in HMM are matching the observational data accessible within the IMPEx infrastructure. They are comprised of Mercury, Venus, the Earth, Mars and Titan (see science cases for Venus in Kallio et al. (2008) and Mercury in Kallio and Janhumen (2003)). The scientific focus of the models is on “plasma environments around solar system objects” (Khodachenko et al., 2009, p. 19) and their capabilities range from computation of individual magnetospheric environments to exploitation and processing of the simulation results. The model consists of a numerical code with the ability to simulate ions kinetically as singular particles and electrons magnetohydrodynamically as a fluid under specific solar and planetary conditions over time (Topf et al., 2012). These capabilities will be incoporated into the tools AMDA, CLWeb and 3DView with Web service technologies, since all codes and result catalogues remain accessible only remotely at the responsible project participants to preserve original ownership of each contributing science product.

4.4. Paraboloid Magnetospheric Models (PMM)

The work package 4 “Paraboloid Magnetospheric Models" (PMM) is the second part of the simulation environment in IMPEx and is comprised of a set of analytical magnetospheric models and their respective numerical codes for relevant solar system objects in this project. An application of the PMM to the solar systems magnetic planets has been decribed in plenty of publications, such as Alexeev et al. (2010) for the Mercurian environment.

5. Science Cases for IMPEx

Science cases represent specific research workflows, which form a basis for the conceptional design of infrastructures like EuroPlaNet-IDIS and IMPEx or analysis tools such as AMDA. They can be seen as a collection of resources in a specific thematic field each providing a certain functionality needed to solve the underlying scientific problem. The first step in elaborating science cases is dedicated to deriving involved key procedures which are mandatory in its chain of data exploitation, processing and delegation. Usually, a science case is assembling generic use cases out of this composed procedures, which then can help to define the user requirements of the planned infrastructure (see chapter 6). This process is also contributing to the formal descriptions of the workflow logic needed to automatize science cases with Web service technologies. One of the authors tasks is to study and establish the below considered science case for the IMPEx infrastructure with the focus on the magnetospheric environment of Venus. A complete list of IMPEx science cases is provided at the IMPEx Web site[10] in this project.

5.1. Venus Magnetosphere Studies

The "Venus Magnetosphere Studies" science case is aimed at the characterization of the Venusian magnetic environment, which is created under the influence of solar wind and its interaction with the planets environment. According to studies in Zhang et al. (2008, p. 1), Venus has no "global intrinsic magnetic field, so the magnetosphere is only induced by interaction of solar wind with the planet's conductive ionosphere". In order to understand this process and to identify the borders of the created bowshock, the Venus Express (VEX) spacecraft is currently in orbit, measuring the magnetic field around Venus with the MAG instrument and the solar wind parameters with the IMA and ELS instrument (ESA, 2010). The global magnetic environment of Venus is also studied with an adaption of the 3D quasi-neutral hybrid model (HMM) whose results are presented in Kallio et al. (2008). To verify the accuracy of the numerical model, simulation runs with specific input conditions for the solar activity around the planet are conducted and compared with actual observations by Venus Express. The procedurce of validating a model run against experimental data and vice-versa builds up the basement of this science case. Furthermore, the comparison

---

[10] Science cases of IMPEx: `http://impex-fp7.oeaw.ac.at/cases.html`

of observed extreme solar events against model runs with the same input conditions is envisaged. The final goal of the science case is to provide a structural and procedural overview of each step of the scientific workflow.

5.1.1. Preconditions

The science case assumes a collection of preconditions with regard to the availability and a standardized accessability of observational and simulation data products as well consumation of related computing services. From the spacecraft data sector, the major resources are the MAG, IMA and ELS instrument as well as spacecraft orbit data available via AMDA. These data products provide the input parameters of the numerical model such as solar wind parameters (density, velocity, solar flux inputs) and the direction of the interplanetary magnetic field near Venus Orbit. Furthermore the MAG data acts as a frame of reference for the simulation results of the 3D quasi-neutral hybrid model, available from the modelling data sector. In addition to the resources, a unified inter-tool communication and user data exchange is also considered as a precondition for this science case.

5.1.2. Use cases

Figure 5.1 represents the first generic use case, which was extracted by analyzing the common methodology aimed at meeting the scientific goals of this science case (see Kallio et al., 2008). It shows the way how a user may extract observational and simulation data for further processing. Within the foreseen IMPEx infrastructure the user will enter the system by accessing the data analysis tool AMDA (or any other tool within IMPEx) and work through the following steps:

1. The user will be able to select an appropriate archived simulation run from the HMM environment suited for Venus, either by searching trough a “simulation tree” with keywords or by visually selecting a run while browsing through metadata descriptions.
2. The interplanetary magnetic field and solar wind input parameters of the selected simulation run are saved and similar conditions are searched by the user in the available observational data of the used analysis tool for a given time frame. In case of inconsistencies between the two data types, transformation mechanisms must be executed to properly match the units and scales.
3. The results of the search are saved as a time-table containing time stamps, the matching parameters of the relevant instruments for Venus and the orbital position of Venus Express available in AMDA.

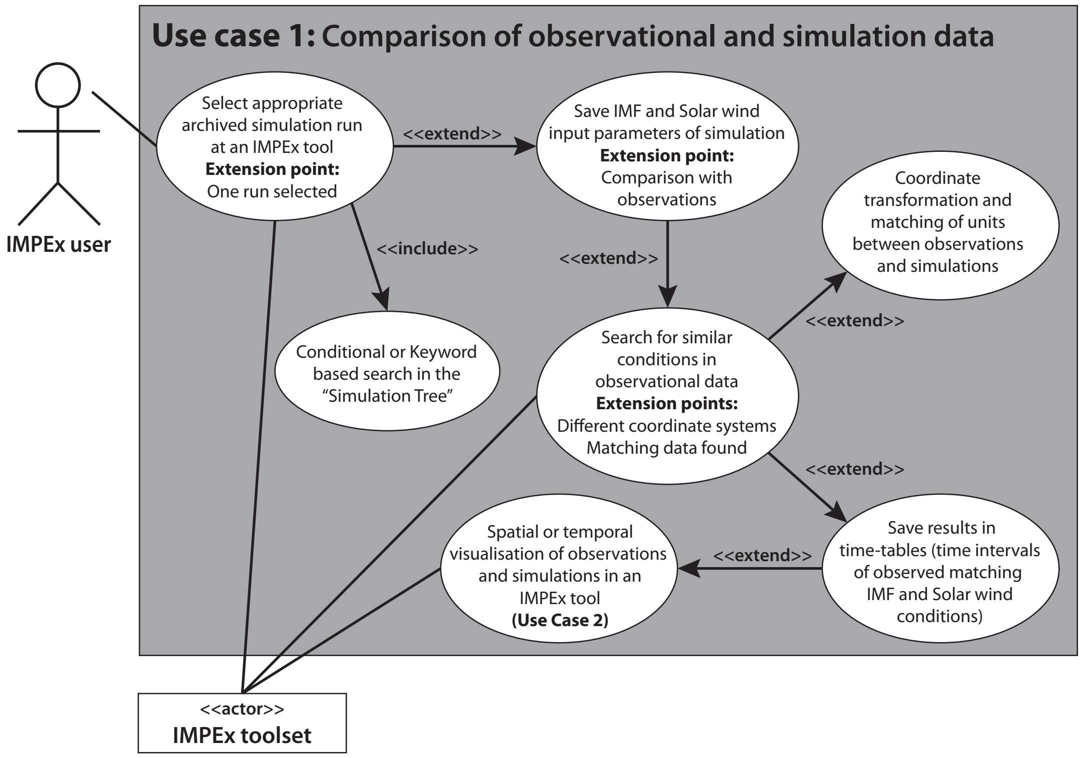


Figure 5.1.: Use case 1: comparison of observational and simulation data

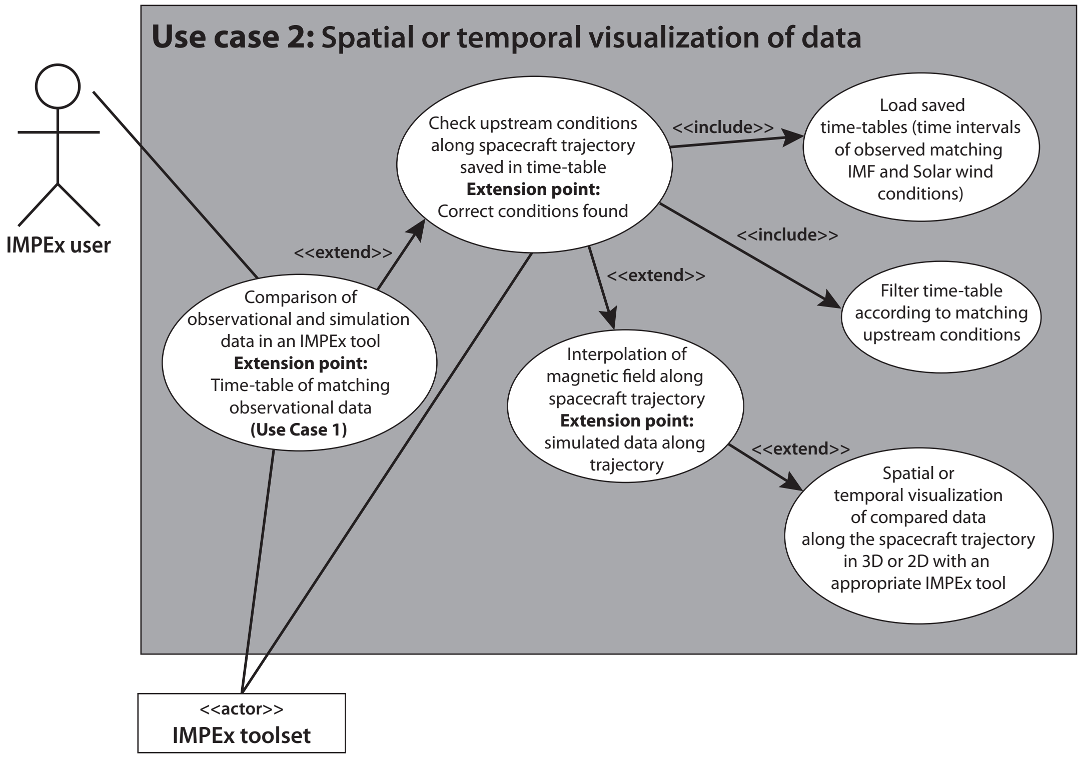


Figure 5.2.: Use Case 2: Spatial or temporal visualisation of Data

The second generic use case in this example is illustrated in figure 5.2. Usually it is directly executed after the first use case. It provides a procedure for using the previously created time-table for visualization of both observational and simulation data aiming at the validation of the numerical model and proofing the accuracy of its results. The major steps are as follows:

1. The user will at first be able to check if the spacecraft was upstream of the bowshock generated by the interaction of the solar wind with the Venusian environment. This is done by conducting a conditional search following the identification procedure in Zhang et al. (2008) within the saved time intervals by analyzing and filtering the positional data of the spacecraft and instrument measurements.
2. The result is a modified version of the input time-table where only the time stamps, corresponding to a position of the spacecraft inside of the bowshock, are kept. In these positions the spacecraft measures only the induced magnetic field of Venus.
3. The spacecraft position data of the time-table will be provided to the HMM computational services which will interpolate the simulated magnetic field components at the provided trajectory points. The results will be delivered to the user and saved accordingly in a particular analysis tool used.
4. Finally, the user will be able to visualize the compared data along the spacecraft trajectory corresponding to the used time-table. There will be two ways of data representation in the IMPEx infrastructure: Either with AMDA to produce 2D plots of the compared simulation and observation (see Fig. 2 in Kallio et al., 2008, p. 799) or with the JAVA tool 3DView to create 3D virtual spacecraft visualisations around Venus. In this case it is also forseen to include animations of the planetary magnetic field in different views and representations.

5.1.3. Post-conditions

The science case is considered as sucessfully accomplished when each obove described step is properly executed. There may be situations where one of the functionalities is not working for specific cases. In each scenario the participating services and resources must be self-consistent and available for interaction with the user. It is necessary to identify the users touchpoints with the IMPEx infrastructure in order to give the user the exact instructions needed to accomplish the task. Particular services or resources of this science case may be also reused, modified and composed within other science cases or other solar system objects. The analysis and visualization tools AMDA, CLWeb and 3DView may offer the possibility to download the science products in appropriate data formats.

6. Key User Requirements for IMPEx

The key user requirements of the science case “Venus Magnetosphere Studies” represent the mandatory functionalities which appear general enough and remain applicable to other science applications. Therefore they have to be implemented in order to successfully integrate further science cases within IMPEx. The scientific research to be conducted with the IMPEx infrastructure is based on complex problems and use scenarios, so the key user requirements have to focus on the essence of what a generic scientific workflow within IMPEx must establish. The use cases extracted in section 5.1.2 provide the blueprint for all service activities being part of this generic composition. According to Hull, Jackson, and Dick (2011, p. 99), a set of goals is defined from the scenario (the science case) which then have to be transformed into new system capabilities during the engineering process. In other words, the user requirements are requesting specific capabilities from the given scenarios, which have to be implemented. In the distributed IMPEx infrastructure, four types of capabilities are needed to sucessfully execute a generic scientific workflow. These capabilities and the assigned requirements for achieving the “Venus Magnetosphere Studies” science case are listed in the following sections.

6.1. Search capabilities

The *search capabilities* build up the core of the workflow, since they are primarily responsible for the standardized and unified access to observational and simulation data. The user will be able to browse all types of data in a hierarchical tree structure within a tool like AMDA, CLWeb or 3DView. For observational data the tree shall be structured mission based, down to the parameter level of each instrument. Simulation data shall be classified according to the involved solar system object, down to the parameter level of each model and its according archived runs. As a quick filter for the available data, a set of keywords will be defined, which will be a part of the standardized descriptions of each data resource, made available within IMPEx. The HMM and other model environments will make all needed physical parameters and a minimum set of standardized metadata from a selected simulation run available within the IMPEx infrastructure. The physical parameters range from scalars, such as electron densities, to vectors, such as the magnetic field vector (B).

For the exploitation of simulation data capabilities, IMPEx tools will implement a mechanism for searching conditions close to the input parameters (e.g. interplanetary magnetic field and solar wind parameters) of a given run within appropriate observational data. The search algorithm will be able to provide a relevance for each search result, so the user can decide on the matching degree of the comparison. Furthermore the user will be able to limit the search parameter range.

6.2. Computation capabilities

The *computation capabilities* represent highly specialized functionalities within the scientific workflows of IMPEx. They are tailored to specific use cases in the overall architecture and may not be as reusable (and composable) as other capabilities of IMPEx.

In particular for the here explained “Venus Magnetosphere Studies" science case, the data analysis tools of IMPEx must be able to process a selected observational dataset of Venus Express to identify the boundary conditions of the Venusian bowshock according to elaborated analytical functions and provide a filtered result where the spacecrafts position was inside the magnetosphere of the planet.

IMPEx shall be able to understand a variety of different coordinate systems and contain their respective transformation rules to related metrics. All the tools and resources participating in a scientific workflow must expose information about their format and the units used for each included parameter in both machine-readable and human-readable format.

The user will be able to obtain the provided model parameters of a selected simulation run and observational data along the trajectory of a given spacecraft. Therefore, a user will be able to transfer a selection of spacecraft coordinates to the modelling environment from an IMPEx data analysis tool (see also section 6.4). This service will deliver the interpolated values in an appropriate manner to the user for further processing (see also section 6.3).

6.3. Visualization capabilities

The AMDA tool will provide the capability to plot observational and simulation data, after selection of a particular dataset with the according IMPEx *search capabilities*. AMDA will display all types of data in time series which include both scalar and vector quantities. The *visualisation capabilities* of AMDA will include an appropriate

2D representation of measured and simulated physical quantities along a spacecraft trajectory. Similar data analysis and 2D visualisation capbilities are available in other participating tools in IMPEx such as CLWeb.

The 3DView tool will offer the visualization of observations and results of simulations after an appropriate selection with similar *search capabilities* as those implemented in the AMDA tool. The 3DView displays the data within a 3D environment of selected celestial bodies and spacecraft orbits. The visualization of time series with scalar and vector quantities coming from any kind of data will be performed in a way which is optimized for the 3D environment. The 3D results of an interpolation of simulated physical parameters within the HMM or PMM environment will be displayed as arrows which can be controlled in number, size and thickness along a spacecraft trajectory.

6.4. Communication and Authentication capabilities

The IMPEx system as an integrated set of autonomous tools will not duplicate data and services of its components. All products used in the infrastructure will remain by the tool of their origin. To fullfil this requirement, *interoperable and standardized messaging protocols* will be used for communcation between the data analysis tools and computational services as well as for the access to observational and simulation data resources participating in IMPEx.

All participating IMPEx tools must be able to handle time-tables, which are generated via the *search capabilties*. Time-tables consist of a list of time spans corresponding to searched physical phenomena or instrument events. At the moment all participating data analysis and visualization tools of IMPEx possess proprietary user authentication system with a dedicated user workspace, so time-tables are saved and loaded within a user session at one of the tools. The infrastructure should provide a possibility to interchange this user data among AMDA, 3DView and CLWeb within a shared user session.

For example, any time-table generated with the *search capabilties* of AMDA or loaded within an user session will be delegated to 3DView with the same parameters selected in AMDA, when the user is active at both systems. Starting from a plot layout at AMDA the user will have the possibility to visualize these data in 3DView with the same time interval and for the same spacecraft and/or simulation result.

The IMPEx infrastructure should finally provide a possibility to connect exsting tool user accounts to simplify the process of sharing the same user session among participating tools in a scientific workflow.

7. Architectural Concepts for IMPEx

According to the key user requirements presented, the distributed infrastructure of IMPEx requires a certain degree of standardization to become interoperable. Each autonomous system activity integrated in a scientific workflow of IMPEx has an assigned role, which needs to be known to all other collaborating services and resources. Following the standards of the VO paradigm there is a distinction between different types of participating functionalities, such as data services, registry services and compute services. In particular, the system capabilities of IMPEx needed to solve the “Venus Magnetosphere Studies" science case can be mapped on these VO functionalities. Similar approaches with regard to collaboration and inter-connection of analysis tools and resources for observational and simulation data are needed also in other applied science cases. The implementation of *registry services* providing queryable interfaces for discovery of observational data available in AMDA is seen as a best-practice example with regard to standardization of metadata and Web service integration (Topf, 2012, p. 40). Since there is actually no such capability available at the HMM environment, a similar approach was proposed with adaption of the SPASE data model. After integration of the defined DaME environment into the established infrastructure, all services will become *reusable* and *composable* for a variety of science cases. In order to sucessfully accomplish a scientific workflow in IMPEx a certain degree of process management must be established to control the whole operational infrastructure. The interaction procedures must be defined and organized for monitoring purposes and the system should be extendable for new tools and services. All the necessary context information must be shared among the collaborating partners including service roles and their relevant relationships between services.

7.1. Overall Infrastructure

The design of the overall infrastrucure of IMPEx is defined by the way how specific interactions within a scientific workflow are managed. The exchange of time-tables generated during the execution of a science case within user sessions at the particular tools must be generally accessible by all other collaborating tools. The IMPEx tools will share user generated data through a “virtual workspace” by inter-connecting the different authentication and user management systems existing in AMDA, 3DView and CLWeb. As one can see in figure 7.1 the “virtual workspace” spans over the server architectures of these tools and complements individual workspaces of the tools by providing a Web service interface, which is called upon authentication at another tool, generating a shared user session on demand. For example, the user will see e.g. in the AMDA user data tree "time-tables from 3DView" and will be able to use a selected time-table for data analysis or visualization at any other available tool. The exchange of the needed user attributes, related security aspects, and the association of user accounts from different tools will be described in details in section 7.4.

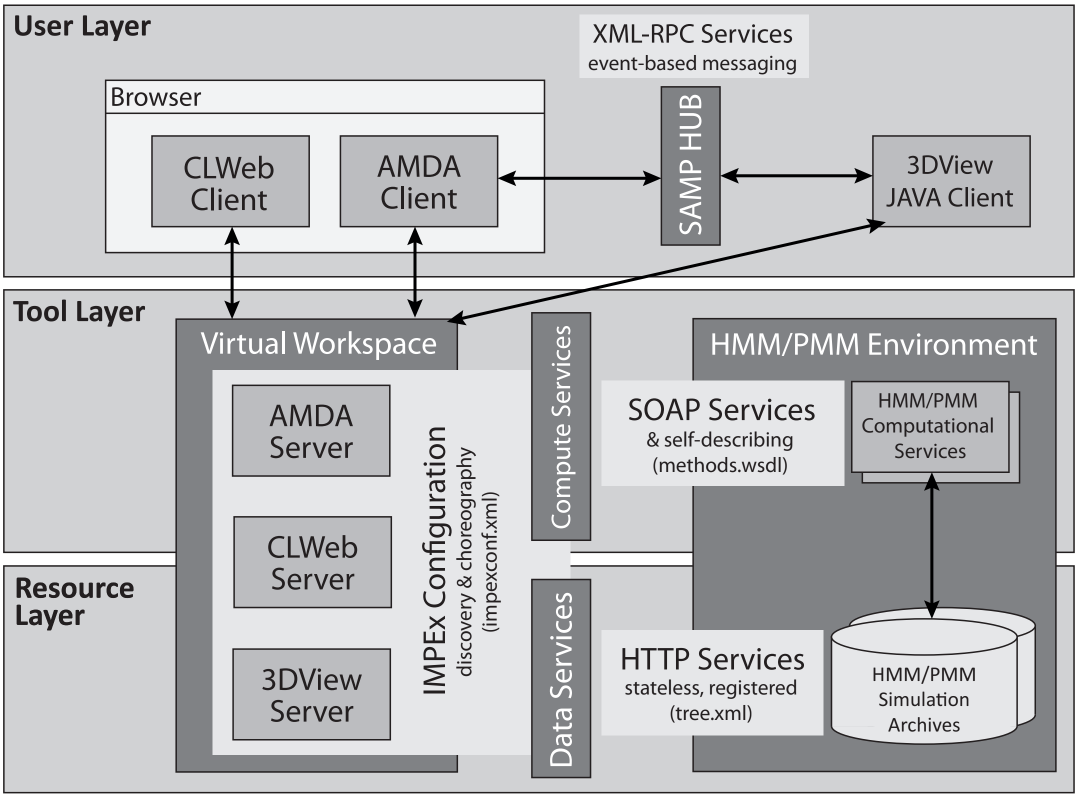


Figure 7.1.: Overall infrastructure of the IMPEx project

Alternatively a "Simple Application Messaging Protocol" (SAMP) hub, developed by IVOA can be activated, which will enable intermediate exchange of e.g. plot and 3D scene configuration data from AMDA to 3DView and vice-versa. The SAMP hub acts as a local messaging middleware launched at the clients machine to communicate between a simultaneously running browser application and a JAVA application (Taylor et al., 2011). The procedure is based on the *publish/subscribe* Web service communication pattern and will be further elaborated in section 7.2.2.

Each computational service and simulation archive provided by the HMM and PMM environment will expose their capabilities by providing descriptions based on XML standards. For the simulation archives, this is done by providing a tree of all stored model runs via HTTP, which is fetched by each participating IMPEx tool regularly, to build up its own interface for browsing and searching trough the metadata. The corresponding `tree.xml` file follows the standards of the IMPEx simulation data model whose XML schema based on SPASE rules and its content for the HMM archive is defined in section 7.3. The capabilities of the related computational services needed for particular science cases are made available via a standard WSDL descriptor for SOAP Web services (`methods.wsdl`). The on-the-fly functionalities include the interpolation along a spacecraft trajectory for provided coordinates where each result is returning a list of URLs via HTTP, which will remain available at the related database for a certain amount of time. The signatures of the needed methods for the HMM environment are described in section 7.2.1.

Finally, the whole IMPEx infrastructure as seen in figure 7.1 is defined with an unique IMPEx configuration file (`impexconf.xml`) which will be used by each system component to discover other participating services and resources. This file must be available at a unique location within the system and co-located at each IMPEx tool after fetching it at the beginning of each user session. The IMPEx configuration provides the basement for collaboration schemes following the *WS-Choreography* standard. There is already information about locations, data sources and service capabilities (`methods.wsdl`) stored in this file (see listing A.1 in appendix A).

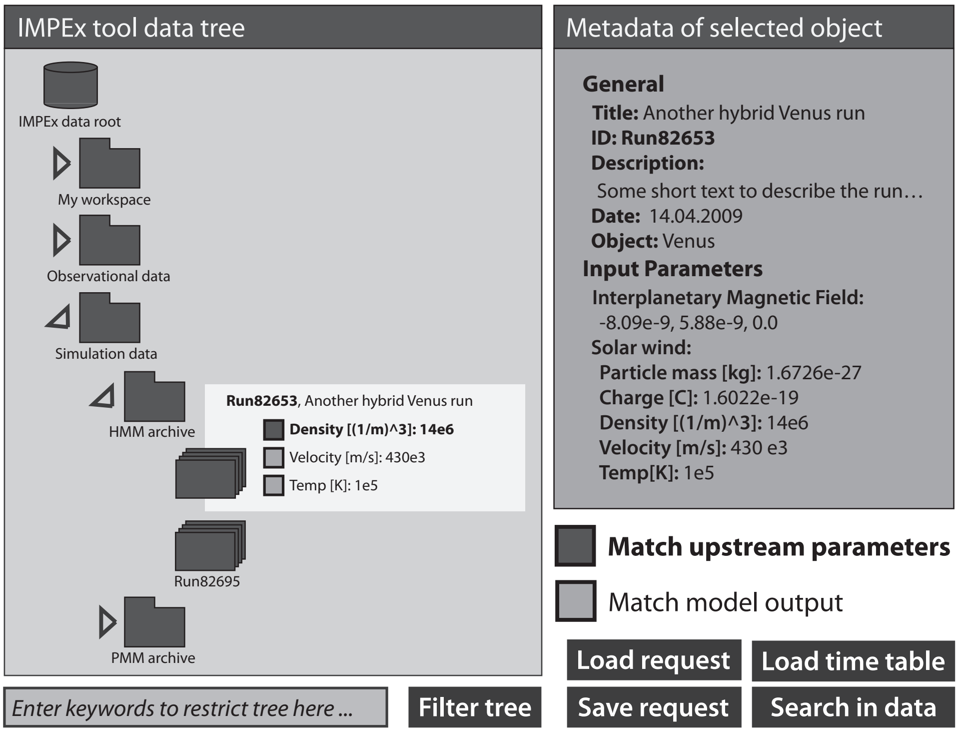
In order to provide autonomous control and execution of scientific workflows, a future extension of the configuration is considered in the architectural design with the establishement of an "Enterprise Service Bus" (ESB). The ESB, according to Erl (2009, p. 704) provides a stronger relationship between collaborating services by introducing a *service broker*, which is an interaction process following the standards of *WS-Coordination*. The underlying messaging middleware enables the definition of message paths and routing for scientific workflows, as well as translation of data models and transformation of delegated data in *WS-Choreography* and *WS-Orchestration* scenarios. The definition of scientific workflows becomes independent of the implementation of each participating tool and resource and the configuration of the overall infrastructure can be managed centralized.

7.2. Functional Services

The *functional services* of IMPEx are comprised of the *search-*, *computation-* and *visualization* capabilities defined in the key user requirements. According to the steps of the "Venus Magnetosphere Studies" science case each tool and resource will be extended in its functionality with a defined standardization procedure. The system will be homogenous in the sense of similar IMPEx user interfaces in all participating tools with predefined working steps according to the generic use cases.

7.2.1. Searching in the IMPEx tools

The browse interface shown in figure 7.2 is provided in each of the participating IMPEx tools. The necessary metadata for all archived simulation runs available in the infrastructure is delivered by the according simulation environment. In this case the HMM environment is used and the content of the `tree.xml` corresponding to the URL in the `impexconf.xml` is fetched. A required run can be selected and all input parameters available for multiple selection are displayed. Alternatively the whole data tree can be filtered with keywords, which are generated out of the available `tree.xml` files. Note, that XML search and query functionalities are provided individually by the tools, and all queries made within the data tree stored for later (re)use.



Metadata of selected object

General

Title: Another hybrid Venus run
ID: Run82653
Description:
Some short text to describe the run…
Date: 14.04.2009
Object: Venus

Input Parameters

Interplanetary Magnetic Field:
-8.09e-9, 5.88e-9, 0.0
Solar wind:
Particle mass [kg]: 1.6726e-27
Charge [C]: 1.6022e-19
Density [(1/m)^3]: 14e6
Velocity [m/s]: 430 e3
Temp[K]: 1e5

Run82653, Another hybrid Venus run
■ **Density [(1/m)^3]: 14e6**
☐ Velocity [m/s]: 430e3
☐ Temp [K]: 1e5

■ **Match upstream parameters**
☐ Match model output

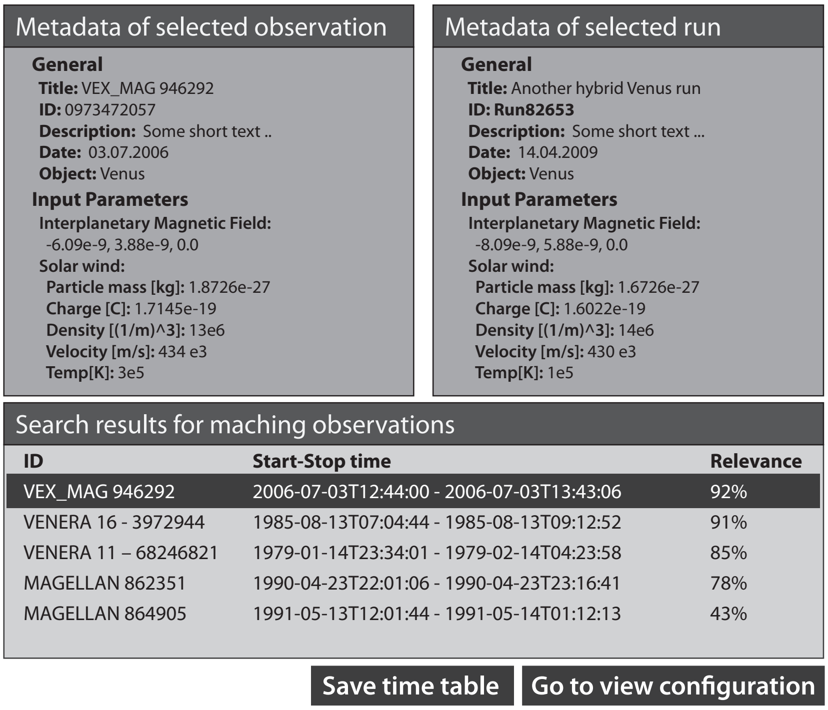
Figure 7.2.: Selection of the model run to be compared

There are two search options after the selection of input parameters from an archived simulation run:

1. **Match upstream parameter** – Solar wind conditions and interplanetary magnetic field data are taken into account for the *best match* search in the observational database. The search is either conducted on all relevant available observational datasets in the IMPEx tool or an individual selection can be loaded via a time-table catalogue.
2. **Match model output** – Only output values of the respective model run are taken into account for the *best match* search in the observational database. In this case a selection of observational data (e.g. magnetic field along a spacecraft trajectory) to be used as input must be loaded via a time-table catalogue.

In case of selecting a match on upstream parameters, the existing internal search interface for observational data is executed with the provided parameters and the respective results are forwared to the next screen in figure 7.3. When selecting a model output match, the according `methods.wsdl` is fetched from defined URL in the IMPEx configuration. A SOAP message is sent by invoking the `getDataPointValue()` method of the HMM computational services. The input of this message is provided by the previously loaded time-table in the search interface which contains 3D coordinates from one or more 2D curves such as a spacecraft trajectory and a unique `ResourceID` of the selected simulation run. The results are returned from the computational services as a time-table and are displayed in the following screen of the search interface as seen in figure 7.3. They consist of the *best matches* of the physical parameters provided together with their measurement coordinates, searched and interpolated within the selected simulation.

It is yet to be defined how and where the relevance factor will be calculated in the search results of an IMPEx tool. The simulation environment uses its inbuilt functions to interpolate the relevant parameters at the given points. It is basically capable of providing a mechanism to compare the observational and simulation data and calculating a variance of the corresponding magnetic field vectors. Note, that the mechanism must be able to provide an algorithm for normalization of those vectors and calculation of the euclidian distance[11] between corresponding vectors, resulting in an quantitative relevance index.



General
Title: VEX_MAG 946292
ID: 0973472057
Description: Some short text ..
Date: 03.07.2006
Object: Venus
Input Parameters
Interplanetary Magnetic Field:
-6.09e-9, 3.88e-9, 0.0
Solar wind:
Particle mass [kg]: 1.8726e-27
Charge [C]: 1.7145e-19
Density [(1/m)^3]: 13e6
Velocity [m/s]: 434 e3
Temp[K]: 3e5

General
Title: Another hybrid Venus run
ID: Run82653
Description: Some short text ...
Date: 14.04.2009
Object: Venus
Input Parameters
Interplanetary Magnetic Field:
-8.09e-9, 5.88e-9, 0.0
Solar wind:
Particle mass [kg]: 1.6726e-27
Charge [C]: 1.6022e-19
Density [(1/m)^3]: 14e6
Velocity [m/s]: 430 e3
Temp[K]: 1e5

ID	Start-Stop time	Relevance
VEX_MAG 946292	2006-07-03T12:44:00 - 2006-07-03T13:43:06	92%
VENERA 16 - 3972944	1985-08-13T07:04:44 - 1985-08-13T09:12:52	91%
VENERA 11 – 68246821	1979-01-14T23:34:01 - 1979-02-14T04:23:58	85%
MAGELLAN 862351	1990-04-23T22:01:06 - 1990-04-23T23:16:41	78%
MAGELLAN 864905	1991-05-13T12:01:44 - 1991-05-14T01:12:13	43%

Figure 7.3.: Listing of search results and selection of best fit

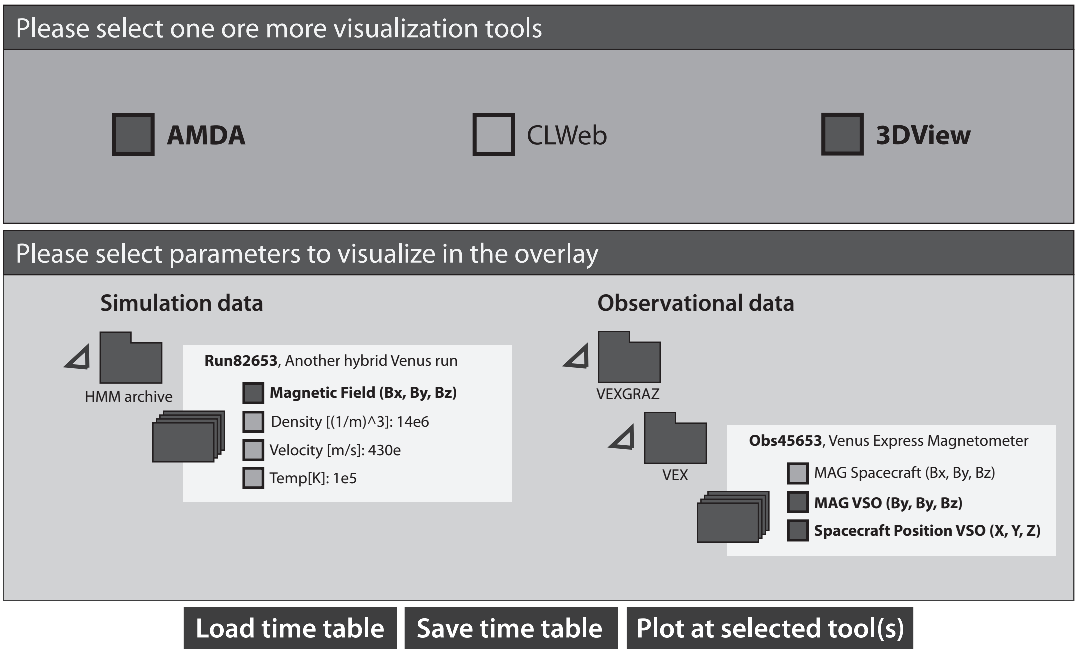
---

[11] Euclidian distance in n-dimensional space: `http://mathworld.wolfram.com/Distance.html`

At this point the whole search result including the comparison configuration can be saved in a time-table catalogue for later processing and visualization. Additionally the result of this step can directly be delegated to the view configuration.

7.2.2. Visualization in the IMPEx tools

At the view configuration screen a summary of the selected parameters in the simulation and observational data branch are listed for a final check. Again the loaded time-table catalogues can be saved or previously generated time-tables can be loaded again. The visualization tools available in IMPEx are listed for selection. Either the IMPEx tool used for search can be activated or the complete view configuration can be delegated to an other, external tool (see figure 7.4).



Please select one ore more visualization tools

☒ **AMDA** ☐ CLWeb ☒ **3DView**

Please select parameters to visualize in the overlay

Simulation data

HMM archive

Run82653, Another hybrid Venus run

☒ **Magnetic Field (Bx, By, Bz)**
☐ Density [(1/m)^3]: 14e6
☐ Velocity [m/s]: 430e
☐ Temp[K]: 1e5

Observational data

VEXGRAZ

VEX

Obs45653, Venus Express Magnetometer

☐ MAG Spacecraft (Bx, By, Bz)
☒ **MAG VSO (By, By, Bz)**
☒ **Spacecraft Position VSO (X, Y, Z)**

Figure 7.4.: Selection of parameters and visualization tools

In case of choosing an external tool for visualization, a broker service implementing the "Simple Application Messaging Protocol" (SAMP) will be started. According to (Taylor et al., 2011, p. 7) a variety of messaging concepts can be implemented, such as the *publish/subcribe* pattern and event-based or point-to-point messaging. In this scenario a *publish/subscribe* pattern between provider and consumer is used as asynchronous messaging middleware. The so-called SAMP hub can be realised as an integrated adapter service or an independently started JAVA application and is implementing the XML-RPC protocol (see also Topf, 2012, pp. 17-19). IMPEx will use a similar implementation as that for the connection of AMDA with the JAVA application Aladin (see André et al., 2011). The IMPEx tool will provide a pop-up window after execution of the view configuration, where an instance of another IMPEx tool can be started, for example 3DView. After the start of the second tool, it can be registered at the previously activated SAMP hub. The second tool directly implements a discovery mechanism, which immediately recognizes the started broker service (at the first tool) and automatically subscribes to it. The initiating service is then publishing the view configuration at the hub which is then broadcasted to the participants. Each registered tool must be able to translate the encoded SAMP messages into their own view configuration mechanisms. The same capabilities must be provided, when a visualization was already conducted at one of the IMPEx tools and the displayed time-tables need to be displayed at another IMPEx tool within a shared user session. The complete message semantics and the vocabulary of the SAMP protocol are found in Taylor et al. (2011, p. 55-60).

7.3. Simulation data model

The IMPEx simulation data model is needed for providing a vocabulary and semantics for the search in `tree.xml` metadata files and for standardization of the outputs from the HMM computational services. The XML standard definitions elaborated, are based on the experiences made with the SPASE data model in basic Web service implementations for AMDA (Topf, 2012, p. 43). Since the SPASE data model does not provide all necessary elements to describe simulation data, an extension namespace for IMPEx should be provided by the simulation data model. There are three types of data to be described and standardized by this extension for accomplishing the needed search capabilities in IMPEx:

- a simulation run, including its input parameters,
- the catalog of outputs of a simulation run,
- outputs datasets and their included parameters.

The extension and its rules are provided by an additional XSD schema, which will be used in the namespace definition of the `tree.xml` file. The header of each file is comprised of the standard descriptions defined in the SPASE data model (see Topf, 2012, pp. 54). The whole `tree.xml` is considered as an *infrastructure resource* according to SPASE (The SPASE Consortium, 2011).

For simulation runs the file contains a number of `<SimulationRun>` elements, which is treated similar to the `<Observatory>` element in the observational data descriptions as *origination resource* (The SPASE Consortium, 2011). It contains access information to the output archives of the run by providing URIs to the corresponding SPASE descriptions in the `ResourceID` element and details about the simulated planetary object. Besides of that, a `<SimulationDomain>` element is introduced by the SPASE extension for IMPEx. It contains all necessary input parameters and details of the physical processes and particle populations which were used in a specific run. Of certain interest for the science case is the `<InputParameter>` element, which provides the necessary information for the matching of upstream parameters. All possible elements and their constraints are further described in the respective IMPEx specification in Hess, Gangloff, Modolo, Génot, and Jarvinen (2012).

The simulation run outputs metadata linked in the `<SimulationRun>` element are also stored in the `tree.xml` file and are represented with standard `<Repository>` elements from the SPASE data model for each available data catalogue. These post-processed results may contain 3D boxes, 2D cuts or simulated parameters along spacecraft trajectories as time-series. They also represent *infrastructure resources*, according to The SPASE Consortium (2011). Each `<Repository>` element provides general information about the data product, such as the access URL and a `<NumericalData>` element, containing metadata of each simulated parameter. The structure and vocabulary for the `<NumericalData>` element are similar to the dataset descriptors for the VEX-MAG observational data (Topf, 2012, pp. 56). The full `tree.xml` file of the HMM environment is provided in Hess et al. (2012, pp. 36-50).

7.4. Authentication & Authorization

The IMPEx authentication and authorization concept is based on the requirement to have a possibility to connect existing tool user accounts within a dedicated IMPEx user session defined. This assumes a simplified single sign on solution with a federated user management for the whole provided infrastructure. The goal is to make the individual workspaces share their user related data with each other, in oder to simplify the exchange of time-tables, for example. It is also of particular interest, when a user who returns to one of the IMPEx tools to continue to work on previously generated searches or to view configurations with one of the other IMPEx tools.

There is no mandatory need of a centralized IMPEx user database in the concept, although the implementation of an identity provider such as OPENID[12] is considered in future extensions for the IMPEx infrastructure. The provided concept encourages the use of existing IMPEx tool user databases, so new IMPEx users may register at any participating authentication system and start working with IMPEx functionalities.

When working within the IMPEx environment a unique `IMPEx ID` is assigned to each tool user account involved in the execution of a scientific workflow. This `IMPEx ID` represents a relation between separate tool user accounts, which identifies them in the IMPEx environment as one unique IMPEx user. In the described scenario in figure 7.5 a user starts working with the IMPEx infrastructure with an existing user account at one of the participating IMPEx tools, AMDA, CLWeb or 3DView.

---

[12] OPENID Single Sign On Solution: `http://openid.net/`

The following step-by-step explaination of the authentication procedure with two participating IMPEx tools describes the sequence of system actions needed when establishing a shared user session according to figure 7.5:

Step 1. The user requests a login at one of the participating tools and gets authenticated with the inbuilt authentication system. A valid user session at Tool 1 is created in **step 1.1**.

Step 1.3 The user requests a restricted functionality at Tool 2, being part of the IMPEx infrastructure. An `IMPEx ID` is assigned to the Tool 1 user acccount in **step 1.2** at the first usage of any IMPEx functionality available in Tool 1.

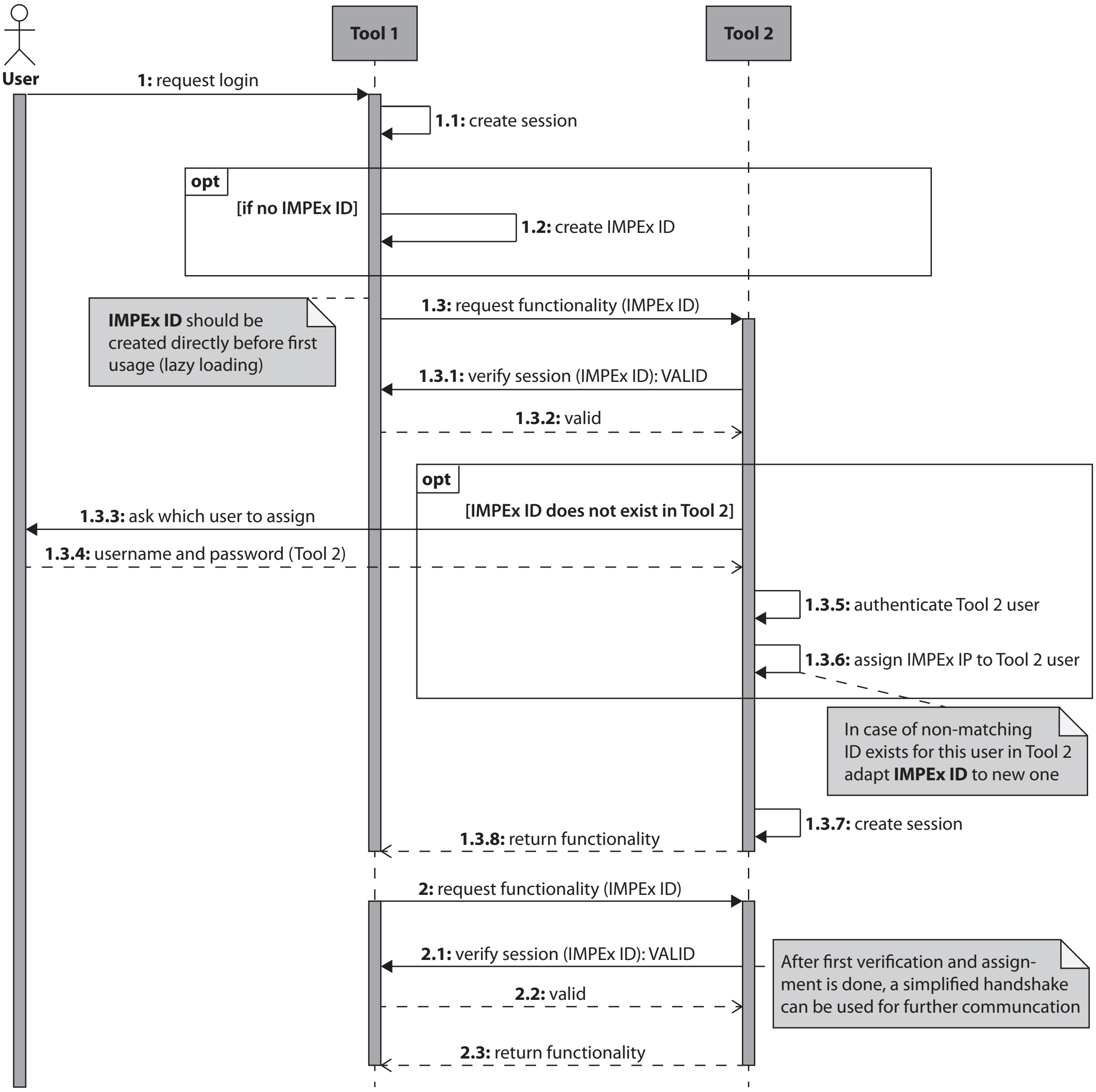


Figure 7.5.: Sequence diagram of the consolidated IMPEx authentication concept

Step 1.3.1 The active Tool 1 user session is verified with the `IMPEx ID` against the Tool 2 authentication system.

Step 1.3.2 If there is a Tool 2 user account existing with the same `IMPEx ID` the session is validated and the user can continue with the use of the accquired IMPEx functionality.

Step 1.3.3 If there is no Tool 2 user account with the same `IMPEx ID`, the user is asked which Tool 2 user to assign to the active IMPEx session.

Step 1.3.4 The user provides username and password for the Tool 2 user account and gets authenticated in step **step 1.3.5** as seen in figure 7.5. Alternatively, the user may now be able to obtain a Tool 2 user account.

Step 1.3.6 The active `IMPEx ID` is assigned to the Tool 2 user account. In case of a non-matching `IMPEx ID`, the `IMPEx ID` of the Tool 2 user account is updated to the `IMPEx ID` of the currently active session. A valid user session at Tool 2 is created in **step 1.3.7** as seen in figure 7.5.

Step 1.3.8 The user is finally successfully authenticated at both tools and the Tool 2 returns the requested IMPEx functionality to Tool 1.

After the first verification and assignment of the common `IMPEx ID` a simplified handshake protocol can be used for further communication within the active IMPEx user session as seen in **steps 2.0-2.3** in figure 7.5. An IMPEx functionality is requested, the session and `IMPEX ID` are verified and the IMPEx functionality is returned. It is yet to be studied, if appropriate SOAP extensions, being part of the *WS-Security* definitions introduced in Topf (2012, p. 17), can be integrated in the service-oriented infrastructure of IMPEx to provide a tailored protocol for this concept.

8. Results and Discussion

The experiences made within the EuroPlaNet-IDIS project and the increasing service-orientation in interactive research environments represent the basement for the infrastrucutural design of the IMPEx project. The establishment of Web services, each dedicated to a specific purpose and distributed among the global research network, are emerging in the design of so-called "Virtual Observatories". The defintions from IVOA and SPASE provide the needed protocols, vocabularies and semantics to let particular Web services expose their capabilities in a standardized way. The VO paradigm also aims at a sustainable and reusable environment, where each integrated tool or resource stays completely autonomous in its functionality. Compared to the achievements of the EuroPlaNet-IDIS architecture, IMPEx integrates computational services and simulation data to the already existing service-oriented data analysis tools connected to observational data. The standardized registry-based access to observational data in AMDA represents the initial condition of the IMPEx infrastructure and this interconnection between tools and resources defines the path for integration of simulated data. The overall architectural design of IMPEx therefore reuses the existing services, defines new services with extended standards and creates a mechanism to make all participating tools and resources composable within scientific workflows. In order to make services reusable and composable, they must first be made discoverable via *registries* in the same way as in the EuroPlaNet-IDIS system.

The composability of services can be assured by coordination through a centralized or decentralized controller mechanism and classification of services according to their purpose in the overall process. In respective VO standards, each tool and resource is defined by its interaction characteristics with the user. The flow of data through the whole aggregated process must be identified, a message-path must be elaborated and each singular interaction between the participating services must be described. Depending on the way how the assigned task is achieved in the composed complex service activity, a certain degree of governance is needed to solve the respective problem. It is either considered to have context information shared among all participating services or the workflow logic is held by one controller service as defined in the presented industry standards. The *WS-Coordination* concept shares the workflow logic created by registered services trough activiation services to users and ensures atomicity of the process by monitoring each single result of the integrated service activities. In contrast, the *WS-Orchestration* concept defines service roles, relationships and all possible message exchange scenarios, which are co-located at the collaborating tools.

In the IMPEx project, all services and tools are highly distributed and used for other collaborative processes, so the decentralization of the ownership of scientific workflows is needed to ensure compatability to other service connections and to make particular IMPEx functionalities available to external entities. There are also certain aspects of possible inter-tool communication within IMPEx, where a centralized broker service for message exchange will simplify steps in a scientific workflow. In order to identify all the according requirements for the IMPEx infrastructure, the participating tools and resources were analysed and put in a workflow context with the "Venus Magnetosphere Studies" science case. Science cases are seen as blueprint for the design of service oriented scientific workflows. The interaction of the system with the user and particular data exchange actions are documented. This design process is needed to define according interfaces between the particular IMPEx capabilities within a scientific workflow. A standardized access to simulation data by using XML metadata is seen as a pre-requisite for all new IMPEx functionalities which will be needed by the required *search capabilities* in all participating data analysis and visualization tools. The *computation capabilities* of all simulation enviroments will need to make their functionalities remotely accessible by WSDL interfaces. Finally the *communication* and *authentication capabilities* will provide a set message exchange protocols each dedicated to a specific use case scenario and a federated authentication system among all participating data analysis and visualization tools.

The overall infrastructure of IMPEx is defined through a configuration file shared among all participating entities. It describes the different roles and capabilities of each tool and resource in particular by providing access to the *registries* of archived model runs via HTTP protocol and the computational functionalities of the simulation environments via SOAP interfaces. The synchronisation between plotting capabilities of AMDA and visualisation functionalities of 3DView will be accomplished by a client side SAMP hub, implementing event-based messaging with the XML-RPC protocol. A shared "virtual workspace" is interconnecting the individual workspaces and user authentication systems existing in the IMPEx tools in order to be able to access user generated searches and view configurations in any IMPEx tool at any time. This architectural aspect is accomplished by implementing a procedure to link existing tool user accounts with an `IMPEx ID` to form a unique IMPEx identity and to establish a shared user session among aggregated capabilities from different IMPEx tools.

Finally, the "Enterprise Service Bus" concept was compared with the actual implementation of the IMPEx configuration during the architectural design phase and identified as a possible future extension of the overall workflow management in IMPEx. This service oriented infrastructure pattern may provide a stronger relationship between collaborating services by publishing detailed information about the interaction processes via an independent *service broker*. It will enable a centralized management of scientific workflows making the processes completely independent of the composed services by providing common data transformation and message translation capabilities for all participating services.

9. Conclusions and Outlook

The overall design of the IMPEx infrastructure builds on top of the well established standardization of scientific data access and exploitation. A general trend of service-orientation in distributed research environments is considered as a crucial argument in the design of interfaces and protocols between IMPEx tools and resources. A certain attention is directed to the fact, that each particular service within the system must stay autonomous and reusable for other tasks. The composition of service activities within the IMPEx infrastructure, designed for the execution of scientific workflows is extending the service-orientation aspects, by introducing complex *service processes*, which need to be coordinated in a suitable way. The elaborated concepts are clearly going in the right direction with regard to automatization and aggregation of particular tasks, however due to the highly distributed framework, some functionalities may be not suitable for complete integration into particular *service processes*. In that respect, an important question has to be raised in the design of this service-oriented infrastructure: "How much distribution does make sense in a complex interactive research environment such as IMPEx?"

In particular, the searching, plotting and visualization capabilities need implementation of complex procedures to exchange information among each other. The VO paradigm propagates the establishment of highly distributed environments in order to keep separate tools and functionalities in the broader scientific community. In contrast, the IMPEx project possesses a specialized scientific background where each of the participating tools is sharing base capabilities which could be integrated in a singular system. Client side solutions for inter-tool communcation such as the SAMP protocol are suitable for quick solutions, however they are considered as not perfect by means of data integrity (Taylor et al., 2011).

The extension of IMPEx with an ESB may provide additional management capabilties for scientific workflows which are published centralized via service registries and executed by an integrated controller mechanism. The configurations of each participating service are managed in one location and the base capabilities of IMPEx like search functionalities for remote databases may be directly accessible via the ESB. A federated authentication system is additionally fostered in this extension, where a unique identity provider can be provided on top of the ESB implementation. Sucessive projects of the IMPEx initiative may increase these efforts by focusing on overall process management and definition of concrete technological goals in order to prevent possible goal conflicts among the participating institutions.

A. The IMPEx configuration

The following listing represents the actual draft of the IMPEx configuration file, which is fetched at the start of each user session to update all participating tools on the location and capabilities of all other services and resources available in IMPEx.

Listing A.1: The IMPEx configuration XML file

```
<?xml version="1.0"?>
<impexconfiguration xmlns="http://www.impex.org/2012/↩
   configuration.xsd">
  <database type="simulation">
  <!-- type="simulation|observation" -->
    <name>HMM environment</name>
    <description>
      Hybrid simulation database for space plasma physics
    </description>
    <dns>hwa.fmi.fi</dns>
    <methods>/methods.xml</methods>
    <tree>/tree.xml</tree>
    <protocol>http</protocol>
  </database>
  <database type="simulation">
    <name>PMM environment</name>
    <description>Dynamic paraboloid model</description>
    <dns>smdc.sinp.msu.ru</dns>
    <methods>/methods.xml</methods>
    <tree>/tree.xml</tree>
    <protocol>http</protocol>
    <info>
      http://smdc.sinp.msu.ru/index.py?nav=model-para
    </info>
  </database>
  <tool>
    <name>AMDA</name>
    <description>
      A Web tool for space plasma physics data analysis
    </description>
    <dns>cdpp-amda.cesr.fr</dns>
```

```
    <methods>/methods.xml</methods>
    <protocol>http</protocol>
    <info>
      http://cdpp-amda.cesr.fr/DDHTML/HELP/about.html
    </info>
  </tool>
  <database type="observation">
    <name>AMDA DB</name>
    <dns>cdpp-amda.cesr.fr</dns>
    <tree>/tree.xml</tree>
    <protocol>http</protocol>
  </database>
  <tool>
    <name>CLWeb</name>
    <description>
      Multimission space plasma data plotting tool
    </description>
    <dns>clweb.cesr.fr</dns>
    <methods>/methods.xml</methods>
    <protocol>http</protocol>
    <info>http://clweb.cesr.fr/clweb_poster.pdf</info>
  </tool>
  <tool>
    <name>3DView</name>
    <description>
      3DView multimission data visualization tool launcher
    </description>
    <dns>http://3dview.cesr.fr/</dns>
    <methods>/methods.xml</methods>
    <protocol>http</protocol>
    <info>http://3dview.cesr.fr/</info>
  </tool>
</impexconfiguration>
```

Glossary

3DView	3DView Multimission, a 3D visualization envionment for spacecraft trajectories and planetary ephemerides, p. 15.
Aladin	Aladin Sky Atlas, an interactive astronomy database, developed by CDS, Strasbourg, p. 29.
CLWeb	A web-based scientific analysis tool developed by CESR, Toulouse, p. 16.
DaME	Data and Models Environment, a work package of the EU FP7-SPACE project IMPEx, p. 15.
ELS	Abbreviation of Electron Spectrometer, p. 17.
EMDAF	European Modelling and Data Analysis Facility, a workpackage within the EU FP7 project EuroPlanet, p. 2.
ESB	Enterprise Service Bus, a design pattern for Web service infrastructures, p. 26.
FMI	Finnish Meteorological Institute, a multi-disciplinary research agency located at Helsinki., p. iii.
HMM	Hybrid and MHD Models, a work package of the EU FP7-SPACE project IMPEx, p. iii.
HTTP	Hypertext Transfer Protocol, the cornerstone for data communication in the World Wide Web, p. 25.
IMA	Abbreviation of Ion Mass Analyzer, p. 17.
IRAP	Institut de Recherche en Astrophysique et Planétologie, a french research institution for astrophysics and planetology located in Toulouse, p. iii.
JAVA	An object-oriented programming language, originally developed by Sun Microsystems, p. 20.

MAG	Abbreviation of Magnetometer, p. 17.
FP7	7th Framework Programm for Research and Technological Development, a funding programme created by the European Union, p. iii.
IMPEx	Integrated Medium for Planetary Exploration, an EU FP7-SPACE collaborative project, p. iii.
EuroPlaNet	European Planetary Network, a EU FP7 research infrastructure project, p. iii.
AMDA	Automated Multi-Dataset Analysis Tool, a web-based scientific analysis tool developed by CDPP, Toulouse, p. iii.
IVOA	International Virtual Observatory Alliance, a community dedicated to the development of a common VO, p. 1.
IDIS	Integrated and Distributed Information System, a workpackage within the EU FP7 project EuroPlanet, p. 5.
WSDL	Web Service Description Language, an XML based describtion vocabulary for Web services, p. 8.
IPDA	International Planetary Data Alliance, a community dedicated to the development of a common planetary data access protocol, p. 12.
SAMP	Simple Application Messaging Protocol, an IVOA standard based on the XML-RPC protocol, p. 25.
SOAP	Simple Object Access Protocol, an XML based messaging protocol for Web services, p. 26.
XML-RPC	A Remote Procedure Call protocol based on XML, p. 29.
PMM	Paraboloid Magnetospheric Models, a work package of the EU FP7-SPACE project IMPEx, p. 16.
RTD	Abbreviation of Research & Technical Development, p. iii.
SPASE	Space Physics Archive Search and Extract, a data model for describing and accessing heliophysics data, p. 24.
URI	Uniform Resource Identifier, a unique identifier for an abstract or physical resource, p. 30.

URL	Uniform Resource Locator, a special version of an URI used in the World Wide Web to address Web resources, p. 26.
VEX	Venus Express, an ESA planetary mission to Venus, p. 31.
VO	Virtual Observatory, often referred to an distributed online data analysis environment for space sciences, p. 12.
W3C	World Wide Web Consortium, defining standards and specifications for the Internet, p. 2.
XML	eXtensible Markup Language, a collection of encoding rules to make documents machine-readable, p. 1.
XSD	XML Schema Definition Language, a set of rules an XML document must conform in order to be considered valid, p. 30.

List of Figures

Listings

References

Alexeev, I., et al. (2010). Mercury's magnetospheric magnetic field after the first two MESSENGER flybys. *Special Issue on Icarus*, *209*(1), 23–39. Retrieved from `http://adsabs.harvard.edu/abs/2010Icar..209...23A`

André, N., Cecconi, B., Budnik, E., Jacquey, C., Génot, V., Fedorov, A., … Topf, F. (2009, November). Connecting the CDPP/AMDA service to planetary plasma data: Venus, Earth, Mars, Saturn (Jupiter and comets). In *Proceedings SF2A 2009* (p. 231). Besançon, France. Retrieved from `http://hal.archives-ouvertes.fr/hal-00442838/PDF/sf2a_09_andre.pdf`

André, N., et al. (2011, November). HST Auroral Campaign Observations of Jupiter and Saturn enabled by the CDPP/AMDA and IVOA/Aladin tools. In *Proceedings of the PV 2011 Conference, 2011, Toulouse, France.* Retrieved from `http://typhon.obspm.fr/idis/docs/Andre_PV2011.pdf`

Arviset, C., Gaudet, S., & the IVOA Technical Coordination Group. (2010, November). *IVOA Architecture Version 1.0* (Note). IVOA. Retrieved from `http://www.ivoa.net/Documents/Notes/IVOAArchitecture/`

Blanc, M., et al. (2008, December). *EuroPlaNet Research Infrastructure ANNEX I - Description of Work.* Seventh Framework Programm Capacities Research Infrastructure - Proposal Grant Aggreement No. 228319. (Project Summary: `http://cordis.europa.eu/fetch?CALLER=FP7_PROJ_EN&ACTION=D&DOC=1&CAT=PROJ&QUERY=0130ad0c101f:b682:561aca31&RCN=92505`)

Booth, D., Haas, H., McCabe, F., , Newcomer, E., Champion, M., … Orchard, D. (2004, February). *Web Services Architecture* (W3C Working Group Note). W3C. Retrieved from `http://www.w3.org/TR/ws-arch/`

Cecconi, B., Gangloff, M., Pallier, E., Bourrel, N., André, N., Érard, S., … Topf, F. (2011, February). *Interoperable Data Access Data Model: Initial Prototype for Plasmas* (Specification). EuroPlaNet. Retrieved from `http://typhon.obspm.fr/idis/docs/Data_Model_v118a.pdf`

Djorgovski, S. G., & Williams, R. (2005, December). Virtual Observatory: From Concept to Implementation. In N. Kassim, M. Perez, W. Junor, & P. Henning (Eds.), *From Clark Lake to the Long Wavelength Array: Bill Erickson's Radio Science* (Vol. 345, p. 517). Retrieved from `http://arxiv.org/pdf/astro-ph/0504006v1.pdf`

Erl, T. (2004). *Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services* (First ed.). Prentice Hall PTR.

Erl, T. (2005). *Service-Oriented Architecture: Concepts, Technology and Design* (First ed.). Prentice Hall PTR.

Erl, T. (2007). *SOA Principles of Service Design* (First ed.). Prentice Hall PTR.

Erl, T. (2009). *SOA Design Patterns* (First ed.). Prentice Hall PTR.

Erl, T. (2012a, April). *SOA Glossary* (Definition for Service-Oriented Computing Terms). SOA Systems Inc. Retrieved from `http://www.soaglossary.com/`

Erl, T. (2012b, April). *SOA Specs* (Primary SOA Standards and Specifications). Arcitura Education Inc. Retrieved from `http://www.whatissoa.com/soaspecs/`

ESA. (2010, November). *Venus Express Mission Overview.* Retrieved June 4, 2012, from `http://www.esa.int/esaMI/Venus_Express/`

Hass, H., & Brown, A. (2004, November). *Web Services Glossary* (W3C Working Group Note). W3C. Retrieved from `http://www.w3.org/TR/ws-gloss/`

Hess, S., Gangloff, M., Modolo, R., Génot, V., & Jarvinen, R. (2012, May). *IMPEx Simulation Data Model: Format of the XML Files describing the simulation inputs and outputs* (Specification Draft). IMPEx.

Hull, E., Jackson, K., & Dick, J. (2011). *Requirements Engineering* (Third ed.). Springer, London.

Kallio, E., & Janhumen, P. (2003, September). Solar Wind and magnetospheric ion impact on Mercury's surface. *Geophys. Res. Letters*, *30*(17), 1877. Retrieved from `http://www.agu.org/pubs/crossref/2003/2003GL017842.shtml`

Kallio, E., et al. (2008). The Venusian induced magnetosphere: A case study of plasma and magnetic field measurements on the Venus Express mission. *Planetary and Space Science*, *56*(6), 796–801. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0032063307003790`

Kavantzas, N., & Burdett, D. (2004, March). *WS Choreography Model Overview* (W3C Working Draft). W3C. Retrieved from `http://www.w3.org/TR/ws-chor-model/`

Khodachenko, M., et al. (2009, January). *IMPEx Collaborative Project ANNEX I - Description of Work.* Seventh Framework Programm Capacities Collaborative Project - Proposal Grant Agreement No. 262863. (Project Summary: `http://cordis.europa.eu/fetch?CALLER=FP7_PROJ_EN&ACTION=D&DOC=1&CAT=PROJ&QUERY=01309ac26125:224b:4536ecf7&RCN=99058`)

Newcomer, E., & Robinson, I. (2007, July). *Web Services Coordination (WS-Coordination) Version 1.1* (OASIS Standard). OASIS. Retrieved from `http://docs.oasis-open.org/ws-tx/wstx-wscoor-1.1-spec-errata-os/wstx-wscoor-1.1-spec-errata-os.html`

Rosen, M., Lublinsky, B., Smith, K. T., & Balcer, M. J. (2008). *Applied SOA: Service-Oriented Architecture and Design Strategies* (First ed.). Wiley Publishing, Inc.

Taylor, M., et al. (2011, September). *SAMP - Simple Application Messaging Protocol Version 1.3* (Proposed Recommendation). IVOA. Retrieved from `http://www.ivoa.net/Documents/SAMP/`

The European Commission. (2007, May). *FP7 Funding schemes.* Retrieved June 3, 2012, from `http://cordis.europa.eu/fp7/what_en.html`

The European Commission. (2009, July). *Work programme 2010 - Cooperation Theme 9 - SPACE.* Retrieved April 10, 2012, from `ftp://ftp.cordis.europa.eu/pub/fp7/docs/wp/cooperation/space/j_wp_201001_en.pdf`

The IMPEx consortium. (2012, May). *The Integrated Medium for Planetary Exploration - Web portal.* Retrieved June 3, 2012, from `http://impex-fp7.oeaw.ac.at/`

The SPASE Consortium. (2011, October). *A Space and Solar Physics Data Model* (Specification). SPASE. Retrieved from `http://www.spase-group.org/docs/dictionary/spase-2_2_2.pdf`

Topf, F. (2012). *Service Oriented Architectures in Planetary Sciences.* Unpublished Bachelor Thesis, FH Campus 02, Graz, Austria.

Topf, F., Scherf, M., Rucker, H. O., Jacquey, C., Khodachenko, M., Génot, V., ... Kallio, E. (2012, March). *Venus Magnetosphere studies under the aspect of modern information technologies in the field of plasma physics.* Talk at the session 17: Laserplasmen und innovatiove Anwendungen at the DPG-Frühjahrstagung 2012. Retrieved from `http://europlanet-plasmanode.oeaw.ac.at/fileadmin/documents/publications/vex-mag-dpg12.pdf`

Weerawarana, S., Curbera, F., Leymann, F., Storey, T., & Ferguson, D. F. (2005). *Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging, and More* (First ed.). Prentice Hall PTR.

Williams, R., Bob, H., Linde, T., McDowell, J., Moore, R., Ochsenbein, F., ... Tody, D. (2004, June). *Virtual Observatory Architecture Overview Version 1.0* (Note). IVOA. Retrieved from `http://www.ivoa.net/Documents/latest/IVOArch.html`

Zhang, T. L., et al. (2008, September). Induced magnetosphere and its outer boundary at Venus. *J. Geophys. Res.*, *113*(E00B20), 9. Retrieved from `http://www-ssc.igpp.ucla.edu/personnel/russell/papers/InducedMagnetosphere.pdf`

Index