



---

## ADMIRE – Platform Delivery and Support progress report

---

<b>Project Title</b>	ADMIRE
<b>Document Title</b>	ADMIRE – Platform Delivery and Support progress report
<b>Deliverable Number</b>	D3.1
<b>Authorship</b>	Rob Baxter, Radek Ostrowski
<b>Document Filename</b>	ADMIRE-D3.1-progress-report.doc
<b>Document Version</b>	1.0
<b>Distribution Classification</b>	Public
<b>Distribution List</b>	<i>ADMIRE Project Team</i>
<b>Approval List</b>	<i>Marcin Choinski, Project Manager, Executive Board</i>

### Document History

<i>Personnel</i>	<i>Date</i>	<i>Comment</i>	<i>Version</i>
RMB	25/07/08	First draft	0.1
RMB	08/08/08	Revision against latest UC doc	0.2
MC	11/08/08	Reviewed	0.3
RMB	25/08/08	Final draft	0.4
RMB	25/08/08	Signed off	1.0

## Contents

<b>Contents</b> .....	<b>1</b>
<b>Executive Summary</b> .....	<b>2</b>
<b>A note on work package numbering</b> .....	<b>2</b>
<b>Related documents</b> .....	<b>2</b>
<b>1 WP 3 Summary: Project Month 6</b> .....	<b>3</b>
<b>1.1 Products delivered to other work packages</b> .....	<b>3</b>
<b>2 Progress against Planned Activities</b> .....	<b>4</b>
<b>2.1 WP 3.1 – scenarios and use cases</b> .....	<b>4</b>
<b>2.1.1 Planned activity</b> .....	<b>4</b>
<b>2.1.2 Actual activity</b> .....	<b>4</b>
<b>2.1.3 Deviations from plan</b> .....	<b>8</b>
<b>2.2 WP 3.2 – set up user support services</b> .....	<b>8</b>
<b>2.2.1 Planned activity</b> .....	<b>8</b>
<b>2.2.2 Actual activity</b> .....	<b>8</b>
<b>2.2.3 Deviations from plan</b> .....	<b>9</b>
<b>2.3 WP 3.3 – set up testbed and initial reference services</b> .....	<b>10</b>
<b>2.3.1 Planned activity</b> .....	<b>10</b>
<b>2.3.2 Actual activity</b> .....	<b>10</b>
<b>2.3.3 Deviations from plan</b> .....	<b>11</b>
<b>2.4 WP 3.4.1 – use case and requirements report</b> .....	<b>11</b>
<b>2.4.1 Planned activity</b> .....	<b>11</b>
<b>2.4.2 Actual activity</b> .....	<b>11</b>
<b>2.4.3 Deviations from plan</b> .....	<b>12</b>
<b>2.5 WP 3.5 – set up Editorial Board and draft book contents</b> .....	<b>13</b>
<b>2.5.1 Planned activity</b> .....	<b>13</b>
<b>2.5.2 Actual activity</b> .....	<b>13</b>
<b>2.5.3 Deviations from plan</b> .....	<b>13</b>
<b>3 Risks and Issues</b> .....	<b>14</b>
<b>3.1 Project issues</b> .....	<b>14</b>
<b>3.2 Significant problem reports</b> .....	<b>14</b>
<b>3.3 New risks</b> .....	<b>14</b>
<b>4 Plans for next Period</b> .....	<b>15</b>
<b>4.1 WP 3.4.2 – revised use cases and requirements</b> .....	<b>15</b>
<b>4.2 WP 3.5.3 – finalise book structure</b> .....	<b>15</b>
<b>4.3 WP 3.6 – ADMIRE Platform Release 1</b> .....	<b>15</b>
<b>Appendix A WP 3 overall timeline</b> .....	<b>16</b>
<b>Appendix B WP 3 deliverable schedule</b> .....	<b>17</b>

## Executive Summary

**This document is ADMIRE deliverable D3.1 and describes progress to date on Work Package 3 within the ADMIRE project.**

**At project month 6 ADMIRE WP3 is on target against original plans, having successfully delivered a collaborative project infrastructure and “level 0” DMI Testbed, an initial use case and requirements report drawing from sources within and outwith the project, and a nightly build-and-test system in preparation for the next stage of the project.**

### 1.1 A note on work package numbering

Sub-work packages within WP3 are numbered according to the list of steps described on p51 of the Description of Work [1] and reproduced (in slightly corrected form) in Appendix A. Thus, Step 4:

4	M4-6	Work with WP6 and other identified users to develop a clarified and structured tabulation of scenarios, use cases and quantified user requirements.
---	------	---

will be numbered in these reports as ‘WP3.4’.

### 1.2 Related documents

- [1] ADMIRE *Annex 1: Description of Work*,  
<http://www.admire-project.eu/trac/browser/projman/plans/ADMIRE-Annex1-DoW.doc>
- [2] ADMIRE website, <http://www.admire-project.eu/>
- [3] ADMIRE internal Trac system, <http://www.admire-project.eu/trac>
- [4] ADMIRE Risks and Issues Log, July 2008,  
<http://www.admire-project.eu/trac/browser/projman/plans/ADMIRE-risksIssues.doc>
- [5] ADMIRE *Use Cases and Requirements Report v0.5*,  
[http://www.admire-project.eu/trac/browser/library/WP3-Use\\_Cases\\_and\\_Testbed/ADMIRE-Use\\_Case\\_%26\\_Requirements\\_Report.doc](http://www.admire-project.eu/trac/browser/library/WP3-Use_Cases_and_Testbed/ADMIRE-Use_Case_%26_Requirements_Report.doc)

## 2 WP 3 Summary: Project Month 6

ADMIRE Work Package 3 will take the products of WP4 (Service Infrastructure) and WP5 (Tools) and package these as conveniently installable downloads, for use by others who wish to use the ADMIRE project and as test platforms for use by WP1, WP2 and WP6. It will provide a website with full documentation and registered, but open, access to the source code of all releases. By combining this with a test-and-build infrastructure, WP3 will provide stable and reliable systems of professional quality that can be safely used by end-users and developers.

WP3 will also provide bug reporting and help-desk advice to users and developers working within and with the ADMIRE project. This community will be built from the partners' existing registered users and new, interested parties. That community will contribute evaluations of ADMIRE technology and encourage other users.

The uptake and understanding of the new methods and the advanced technology will be strongly encouraged through publications, technical notes and conference presentations by *all* of the partners. In addition, ADMIRE will develop a complete draft of a book by the end of the project.

ADMIRE WP3 will thus deliver two key products to the project and user/developer community:

1. A supported and tested distributed DMI platform.
2. A draft of a book on distributed DMI methods and technology.

In months 1-6 WP3's main focus has been on establishing the collaborative infrastructure for the rest of the project. A collaborative working environment has been created, and secure internal testbed set up running existing versions of partners' service platforms. Detailed use cases have been derived from both ADMIRE Pilot Applications and from other sources, including external projects.

### 2.1 Products delivered to other work packages

WP3's main job is to provide the software support and collaboration services to keep the rest of the project on track. It also acts as custodian of the use case and requirements document and the ADMIRE book.

In the first six months WP3 has delivered the following products to the rest of the project:

- The ADMIRE collaborative website, problem report system and code and document repository (Section 3.2);
- The first version of the ADMIRE Testbed including initial data-mining services (Section 3.3);
- The initial the use case and scenarios report (Section 3.1);
- The first draft of the detailed use case and requirements report (Section 3.4).

## 3 Progress against Planned Activities

### 3.1 WP 3.1 – scenarios and use cases

#### 3.1.1 Planned activity

*The initial gathering and description of scenarios and use cases, working with WP6 (DoW [1] p49).*

#### 3.1.2 Actual activity

Four user-led scenarios for ADMIRE were identified in the first month:

- Customer Relationship Management (from the Comarch pilot application);
- Flood and Environmental Modelling (from the IISAS pilot application);
- Genetic Data Mining (in collaboration with the EURExpress-II project);
- Commercial Data Mining (drawn from experience of commercial projects at EPCC).

#### Scenario 1: Customer Relationship Management

The *Comarch CRM for Telco* is a CRM application suite designed to fulfill specific telecommunication companies' requirements and to support all general-purpose CRM system features. The system consists of two logical modules – operational and analytical.

*Operational CRM* supports various front office processes involving interactions between the company and its customers. The operational part of the CRM is not only a business process supporter – it also collects historical information related to customers' behavior, thus itself becomes a valuable source of data.

*Analytical CRM* aims to utilize such information, enabling the measuring, analyzing and finally optimizing of customer relationships. It's designed to support the operational part and the full power of the whole CRM system is available when both parts are integrated and are cooperating.

Typically during the deployment of the analytical CRM system a custom *Data Warehouse* is built and then appropriate data mining models are prepared. All analyses are based on data from the DW. The implementation based on the ADMIRE platform will take quite a different approach.

#### Cross-selling and up-selling

The terms cross-selling and up-selling are often used in the sales and marketing areas, but the actual meaning of these terms can vary significantly. Here we use the terms in the following context:

**Cross-selling** – providing additional products or services to an existing customer. These products fulfill other needs than the products already bought by client, but very often these products are complimentary (i.e. selling VoIP services to the client that already is a broadband internet connection subscriber);

**Up-selling** – providing more expensive products or services to an existing customer. These additional products or services fulfill the same client needs, but are profitable for the company (i.e. selling internet connection with the higher transfer rate – 4Mb/s instead of 2Mb/s).

When a marketing campaign is developed, up-selling and cross-selling sales techniques may be used to prepare the *target group* (the group of people that the marketing campaign is aimed at). Revealing the *characteristics* of the customer that may be interested in buying additional products is also valuable information. These particular uses of these techniques give us two general business use cases: generating a target group, and; gathering user profiles.

## Churn

One of the most oft-quoted marketing slogans is that, from the company point of view, keeping existing customers is much cheaper than acquiring new ones. The process of customers moving to a competitor is called **churning**. In the telecommunication sector switching service provider is very easy for the customer, so for Telco companies, churn prediction becomes a very important feature of the analytical CRM system. This gives us our third business use-case.

## Scenario 2: Flood and Environmental Modelling

This scenario is concerned with the water balance in the river Vah. Its main components are the historical data sets of discharge of waterworks on the river and rainfall data. First we describe the basic workings of the scenario, then several interesting data mining scenarios.

This scenario uses part of the well-known Earth's water cycle<sup>1</sup> to model the relationships between the accumulated rainfall in a river's drainage area, the river's discharge, and another major factor in water movement – evapotranspiration. Evapotranspiration is the process by which groundwater is extracted by plants (via their roots), transferred into leaves, and from there back into the atmosphere by evaporation. As we can see in Figure 1, the rainfall in the river's drainage area is partly absorbed into the ground-water, partly evaporates back into the atmosphere by evapotranspiration, and partly flows into the river, affecting its level and discharge. These major factors affect the water balance of the river's drainage area.

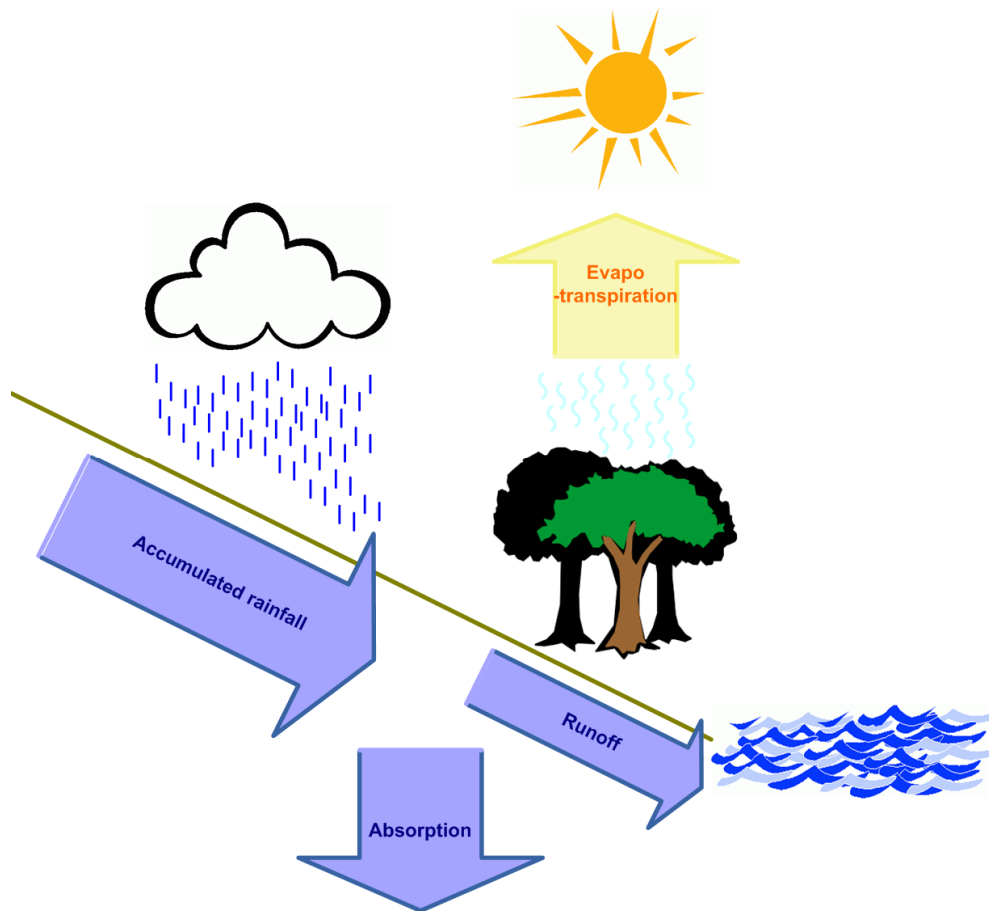


Figure 1. The conceptual scheme of transformation of rainfall into runoff

<sup>1</sup> Water cycle. Article at Wikipedia. [http://en.wikipedia.org/wiki/Water\\_cycle](http://en.wikipedia.org/wiki/Water_cycle)

(a) Water balance: **Accumulated rainfall = Runoff + absorption + evapotranspiration**

Absorption =  $f$  (soil water potential,  $\sum$  ( accumulated rainfall over x past days ), soil type)

Evapotranspiration =  $f$  (Direct sunshine, temperature, wind speed, foliage roughness, root depth, soil water potential)

Principal factors:

- Accumulated rainfall (mm/day) – measured meteorological parameter
- Direct sunshine (hours/day) – measured meteorological parameter
- Foliage roughness – constant phenological parameter
- Foliage area ratio(m<sup>2</sup>/m<sup>2</sup>) – constant phenological parameter
- Root depth (m) – constant phenological parameter
- Wind speed (m/s) – measured meteorological parameter
- Soil water potential (cm<sup>3</sup>/cm<sup>3</sup>) – measured/computed hydrological parameter
- Temperature (°C) – measured meteorological parameter
- Soil type – unknown constant

(b) **Runoff = accumulated rainfall – absorption – evapotranspiration**

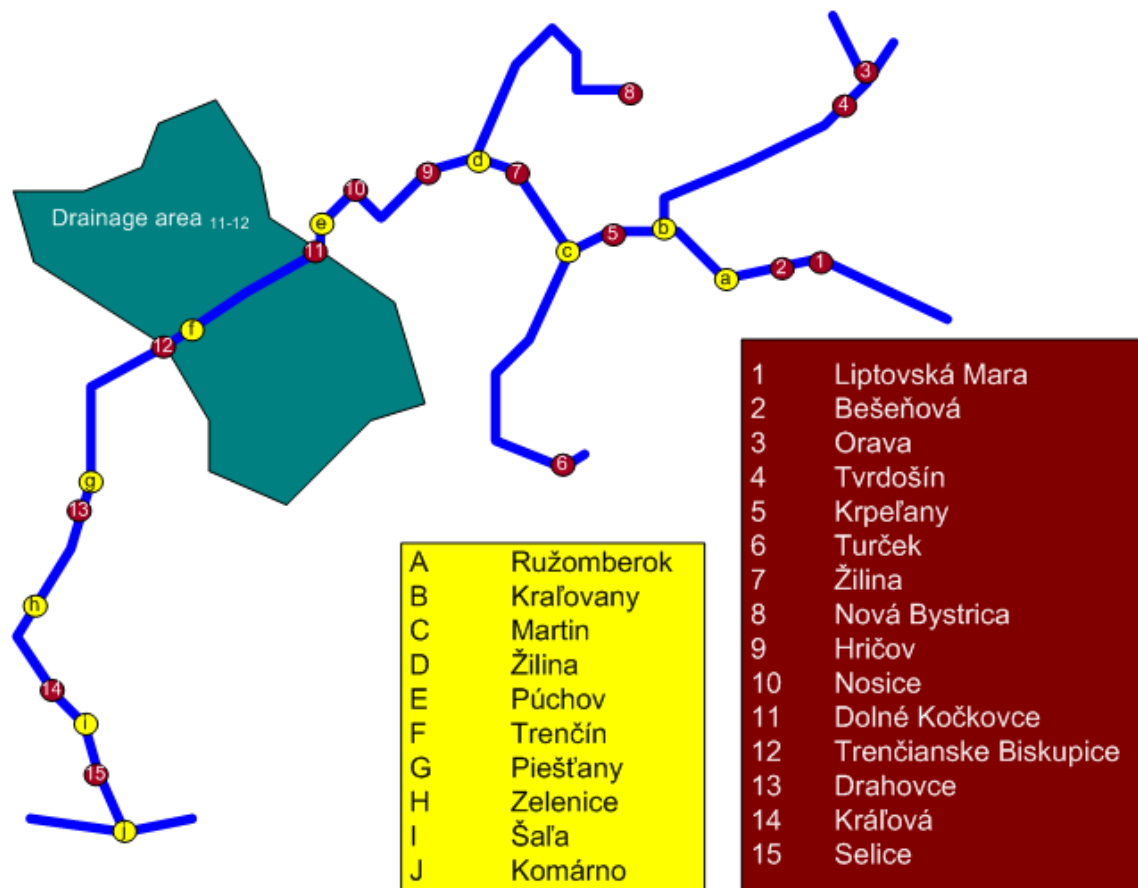


Figure 2. The river Váh cascade. See also

<http://maps.google.com/maps/ms?ie=UTF&msa=0&msid=103581285281031070524.000453a1c96740c899351>

### The River Váh Cascade

There are 15 major waterworks on the river Váh and its main tributaries – see the map in Figure 2. The rainfall-runoff scheme (Figure 1) is applicable on each part of the river between two waterworks. Additionally, the outflow of each waterworks is part of the inflow of the next one in the cascade. So we can say that

(c)	$\mathbf{Outflow_{n+1} = Outflow_n + Runoff_{n-n+1}}$
-----	---

### Detailed usage scenarios

Altogether we have identified five data analysis scenarios to be addressed in greater depth:

- Hydrological modelling of the Vah cascade;
- Radiolocation meteorology for improved weather prediction;
- Relation discovery between meteorological phenomena and measured values from meteorological stations;
- Combining meteorological data with spatio-temporal data from other domains;
- Agro-meteorological scenarios.

### Scenario 3: EURExpress-II Genetic Data Mining

Eurexpress-II (<http://www.eurexpress.org>) is a EU-FP6 project that ends December 2008. Its aim is to build a transcriptome-wide atlas database for the developing mouse embryo established by RNA *in situ* hybridisation. The reality is that the project is restricted to one stage of development. For that one stage they use automated processes for *in situ* hybridisation on all genes. The outcome of these are many images, which are then annotated by human curators. The annotations consists of tagging each file with terms from the ontology for mouse anatomy development. If an image is tagged with a term it means that that anatomical component is “expressing” the gene that was hybridised against.

Two top-level use cases have been identified for the EURExpress-II data:

**Gene pattern extraction.** Create a query that selects useful data from the database. Useful could mean, selecting only patterns that have sufficient annotations and expression patterns that adhere to a minimum expression level. Model the data to fit the machine learning or other data mining algorithm. This may include transformations of the data. Also, it may needs partitioning of the data to facilitate an n-fold validation process. Run the machine learning algorithm. Prioritise the output results and make a selection of significant results. Link results back to the data to show provenance. This task challenges the data mining process, more specifically, the process of building up data models and getting things ready for the actual data mining step. The owners of the data will be interested in getting patterns from their data. To validate output results from the data mining process and to make these more interpretable, the results should be linked with third-party data. Moreover, this third-party data, such as the Gene Ontology, can be used to compute the significance of the results.

**Automating curation of images.** This is essentially a typical classification task. The input is the set of image files, the output for each image should be correctly identification of the anatomical components that exhibit gene expression patterns in the image. A more advanced task would also identify the strength and type of pattern. The challenges posed by this task are primarily dealing with large data volumes, as the images are large and there are many of them, and the complexity of the classification task itself. At the MRC Human Genetics Unit, much expertise exists on image processing, which may lower the complexity of the task by first pre-processing the images. These pre-processing tasks may pose interesting challenges to the design of the data fetching and integration processes.

### Scenario 4: Commercial Data Mining

This scenario describes a fictional data mining project that is based on similar projects carried out by EPCC for commercial companies. The data and results of such projects are typically confidential

hence the fictional nature of this scenario. The contents of this scenario are however based on work that was actually carried out and the types of patterns that were detected in data sets – in this it provides a strong corollary to Scenario 1, sharing many of the same questions and thus approaches to data mining.

We focus primarily on what has to be done to the data prior to running the data mining algorithm on the final set of data. This is what consumes most of the project time. This scenario thus demonstrates the highly interactive nature of a data mining project.

Although commercial data cannot be made available to the ADMIRE project we believe it would be possible to construct a dummy data set that has the properties described here and contains some correlations that we would expect to discover using data mining techniques.

Analysis of this scenario has identified two top-level business usecases:

**Data preparation and model generation for knowledge discovery in commercial data:** This use case consists of preliminary data analysis and model building phases for knowledge discovery in commercial data. It captures all the specific tasks and activities basing on EPCC knowledge and experience in commercial data mining projects. The use case should present more generic and overall approach to data mining than other use cases described in this document as it is a summary of experience in commercial projects. It is not expected to cover all the project specific details. The use case is captured in a way that it also contains information about held analysis and its results to present detailed view of data mining process that *are strongly driven by data analysis results*.

**Data mining model execution for knowledge discovery in commercial data:** This use case presents the data mining phase of knowledge discovery in commercial data. The model built in the previous use case is executed here and the results are collected.

### 3.1.3 Deviations from plan

No deviations from plan.

## 3.2 WP 3.2 – set up user support services

### 3.2.1 Planned activity

*Set up user support service, including working relationships with special projects, documentation and help desk, based on partners' existing services (DoW [1] p49).*

### 3.2.2 Actual activity

#### Support services, documentation and helpdesk

The main focus of WP3.2 in PM1-6 has been in establishing a comprehensive collaborative working environment for the project based around the Trac system<sup>2</sup>. Trac is “*a web-based software project management and bug/issue tracking system emphasizing ease of use and low ceremony. It provides an integrated Wiki, an interface to version control systems, and a number convenient ways to stay on top of events and changes within a project*”.

Already used by Fujitsu Labs for their internal USMT<sup>3</sup> development, and chosen independently by EPCC, the ADMIRE Trac-based website provides all project partners with lightweight support mechanisms and helps build the project community essential to ADMIRE's progress [3]. Opening the Trac system to other registered users will be straightforward in due course; Trac has very a fine-

---

<sup>2</sup> <http://trac.edgwall.org/>

<sup>3</sup> Universal Service Management Technologies (formerly Infrastructure Services Bus, ISB), Fujitsu Labs' core technology contribution to ADMIRE.

grained permissions system – users may be allowed to access the problem report system, for instance, but not the wiki pages.

The ADMIRE Trac system is installed as a restricted-access part of the public ADMIRE website [2] which was also set up under WP3.2 this stage.

### Relations with special projects

ADMIRE has also built strong working relationships with three key projects:

**OGSA-DAI:** funded as one of three OMII-UK sites which together aim to provide software and support to enable a sustained future for the UK e-Science community and its international collaborators, OGSA-DAI develops middleware to assist with access and integration of data from separate sources via the grid. ADMIRE has already been able to exchange staff with the core OGSA-DAI architecture team to ensure a consistency of vision and direction.

**EURExpress-II:** ADMIRE is already collaborating closely with the UK Medical Research Council's Human Genetics Unit, which is an important partner in the FP6 EURExpress-II project<sup>4</sup> as it collects and organises all the intermediate and final data in the project. These data are obtained from expensive wet-lab *in-situ* hybridisation experiments to discover transcriptomes in the developing mouse embryo. Furthermore, the data are then annotated using a highly skilled team of developmental biologists. We shall pursue two important use cases with them to further enrich these unique data. This involves filling in missing annotations in the database as well as facilitate searches for higher order structures (see *Scenario 3* in Section 3.1.2 above).

**BEinGRID:** The mission of BEinGRID, the EU largest Integrated Project, is to establish effective routes to foster the adoption of Grid technologies across the EU and to stimulate research into innovative business models<sup>5</sup>. ADMIRE has built strong links with the Data Management Technical Solutions Cluster and has already identified that one new OGSA-DAI software component required by the Comarch CRM application within ADMIRE is also required by one of the BEinGRID Business Experiments<sup>6</sup>. Early development of this component within ADMIRE WP5 has been prioritised as a result.

**NERIES:** The Network of Research Infrastructures for Earthquake Seismology (NERIES<sup>7</sup>) is a FP6 research infrastructure project with 25 participants and duration of 4 years. NERIES, the largest earth science EC project, is coordinated by ORFEUS<sup>8</sup> and has currently started its third year. A major goal within this project is to provide a unified single access point for a multitude of earthquake seismological data. This is being realized by implementing an advanced distributed archive structure in combination with a Service Oriented Architecture, an activity which may well provide additional use-cases for ADMIRE. ADMIRE has formed collaborative links with both ORFEUS and the European-Mediterranean Seismology Centre<sup>9</sup> to investigate data analysis and data mining scenarios for seismological data.

### 3.2.3 Deviations from plan

No deviations from plan.

---

<sup>4</sup> <http://www.eurexpress.org/>

<sup>5</sup> <http://www.beingrid.eu/>

<sup>6</sup> See <http://www.gridipedia.eu/611.html> for a description of the “OGSA-DAI JDBC Driver” component.

<sup>7</sup> <http://www.neries-eu.org/>

<sup>8</sup> The Observatories and Research Facilities for European Seismology group, <http://www.orfeus-eu.org/>

<sup>9</sup> <http://www.emsc-csem.org/>

### 3.3 WP 3.3 – set up testbed and initial reference services

#### 3.3.1 Planned activity

*Set up initial reference services based on partners' existing technologies configured to enable ADMIRE research by M3 and support these to M6 (DoW [1] p49).*

#### 3.3.2 Actual activity

##### ADMIRE public key infrastructure (PKI)

Given the distributed nature of the ADMIRE project and Testbed we required a security infrastructure to provide proper, lightweight authentication mechanisms. WP3.3 thus invested time in setting up an X509 certificate-based authentication infrastructure. All project members can now acquire ADMIRE certificates to enable them to participate fully in the project. The ADMIRE PKI also recognises certificates signed by two other Certificate Authorities – the UK e-Science CA and the SlovakGrid CA.

##### ADMIRE Testbed hardware

A good deal of activity in WP3.3 has been in deploying a number of servers to form the basis of the ADMIRE Testbed. The Testbed will comprise a number of dedicated servers drawn from across the project consortium across which we intend to deploy two service environments:

- a production environment, based on stable releases of the ADMIRE Platform software, and;
- a development environment, based on incremental updates to the previous Platform plus experimental services deployed as candidates for the next ADMIRE release.

At this stage in the project only the development environment is up and running.

Currently four servers have been set up at EPCC – `admire3` to `admire6.epcc.ed.ac.uk` – running Scientific Linux 5. An additional five servers at IISAS (Linux-based) and two at FLE (potentially running MacOS X) are awaiting the completion of USMT v1.0 before being incorporated into the Testbed configuration.

Our aim with the Testbed is to provide a mix of capabilities and operating systems to exercise the ADMIRE Platform tools and services across a range of hosting environments; introducing Microsoft Windows-based servers will be a target for the next project stage.

##### Deployed services

We have successfully deployed the following baseline services on the EPCC testbed nodes:

[admire3.epcc.ed.ac.uk](http://admire3.epcc.ed.ac.uk)

- USMT v0 (Jetty container):
  - baseline version of USMT service bus “nucleus”;
- OGSA-DAI v3.0 (Axis/Tomcat container)
  - baseline version of OGSA-DAI data access services, fronting the test database *ProductDB* – “polluted” and “clean” (synthetic) customer and contract data;
- MySQL RDBMS – an instance of MySQL hosting the *ProductDB* synthetic “customer database”:
  - available tables: *customer*, *contract*, *customer\_clean*, *contract\_clean*, *littleblackbook*, *extblackbook* and *weather*;
- eXist XML DBMS – an instance of the eXist XML database server;
- Distributed decision tree component from WP5
  - an ADMIRE version of an example data mining tool built by UVIE and UEDIN.

- OGSA-DAI workflow runner & visualiser  
[admire4.epcc.ed.ac.uk](http://admire4.epcc.ed.ac.uk)
- USMT v0 (Jetty container)
  - baseline version of USMT service bus “nucleus”;
- OGSA-DAI v3.0 (Axis/Tomcat container)
  - baseline version of OGSA-DAI data access services, fronting the test database *StatsDB* – (synthetic) postcode mappings and census information data;
- MySQL RDBMS – an instance of MySQL hosting the *StatsDB* synthetic “stats database”:
  - available tables: *stat*, *postcode* and *measurements*;

Together the test databases hosted by OGSA-DAI on admire3 and 4 provide a synthetic but representative example of a typical commercial data mining configuration.

### Nightly build and test process

In preparation for the ADMIRE Platform Release 1 due in PM12 we have deployed the first stages of a nightly build-and-test system on the ADMIRE Testbed. We have adopted Apache Ivy<sup>10</sup> as a dependency manager at the heart of the build system.

The ADMIRE system test suite will run nightly and reports test failures on a series of dashboards. So far we have ported the existing OGSA-DAI system test suite to the ADMIRE Testbed, a series of 607 tests of the OGSA-DAI installation. As of the time of writing, all tests pass. We have begun to review the Metronome<sup>11</sup> distributed test framework as a possible candidate for a future ADMIRE test system.

New system tests have been developed and added to the suite to cover the new *distributed decision tree algorithm* developed in WP5.

### 3.3.3 Deviations from plan

No deviations from plan.

## 3.4 WP 3.4.1 – use case and requirements report

### 3.4.1 Planned activity

*Work with WP6 and other identified users to develop a clarified and structured tabulation of scenarios, use cases and quantified user requirements (DoW [1] p49).*

### 3.4.2 Actual activity

The four scenarios identified in WP 3.1.1 (Section 3.1) form the basis for the *ADMIRE Use Case and Requirements* document [5] which has been drafted in WP3.4.1. This document will be updated throughout the project.

The approach used in this and subsequent requirements work package is as follows:

- identify key user scenario;
- perform top-level analysis on user scenarios to identify key top-level use cases;
- describe each use case in a common form;
- derive functional requirements from each use case description;

<sup>10</sup> <http://ant.apache.org/ivy/>

<sup>11</sup> <http://nmi.cs.wisc.edu/node/65>

- collect functional requirements together and identify both common and unique requirements to feed in to later stages of ADMIRE development.

Altogether 21 detailed use cases have been identified at this stage from the four key usage scenarios. These are:

- UC1.1:** Generating target groups with up-selling and cross-selling sales techniques;
- UC1.2:** Gathering user characteristics with up-selling and cross-selling sales techniques;
- UC1.3:** Churn prediction.
- UC1.4:** Decision Tree model creation (training)
- UC1.5:** Decision tree models browsing
- UC1.6:** Scoring data with decision tree model
- UC1.7:** Preparing filter for customers data set
- UC2.1:** Change in Foliage Properties
- UC2.2:** Prediction of Future Flood Threat Levels
- UC2.3:** Radiolocation reflection prediction
- UC2.4:** Model prediction correction based on radiolocation data
- UC2.5:** Relations discovery in spatio-temporal meteorological data
- UC2.6:** Weather influence on crop production
- UC2.7:** Integration of vulnerability data with the results of meteorological modelling
- UC2.8:** Monitoring soil quality
- UC2.9:** Monitoring extreme values of soil content
- UC2.10:** Prediction of soil content
- UC3.1:** Gene pattern extraction.
- UC3.2:** Automating curation of images.
- UC4.1:** Data preparation and model generation for knowledge discovery in commercial data
- UC4.2:** Data mining model execution for knowledge discovery in commercial data

These use cases have been documented in a common form derived from the IBM Rational Unified Process<sup>12</sup> and using the following template:

- Use Case Name
  - Description
  - Actors
  - Flow of events
    - Main flow
    - Alternate flows
  - Pre-conditions
  - Post-conditions
  - Special requirements
  - Extension points

### 3.4.3 Deviations from plan

No deviations from plan.

<sup>12</sup> The format chosen to describe use cases is the same as that used in the IBM Rational Unified Process (described in [www.dama-nj.org/presentations/TQ%20Writing%20Good%20Use%20Cases.pdf](http://www.dama-nj.org/presentations/TQ%20Writing%20Good%20Use%20Cases.pdf)).

---

## **3.5 WP 3.5 – set up Editorial Board and draft book contents**

### **3.5.1 Planned activity**

*Establish Editorial Board and develop an agreed structure and contents for the book (DoW [1] p49).*

### **3.5.2 Actual activity**

The chair of the Editorial Board has been chosen – Project Architect Prof. Malcolm Atkinson. In consultation with the Executive Board it has been decided to hold off filling the remaining places on the Board, and on drafting the first outline of the book's contents until after the first phase of ADMIRE deliverables (PM6).

### **3.5.3 Deviations from plan**

As described above. The bulk of this task has been deferred until after the first Milestone. This will allow the chair of the Editorial Board to fill the remaining places and begin work on the book's contents with a better idea of the most suitable candidates from around the project consortium.

---

## 4 Risks and Issues

### 4.1 Project issues

No issues recorded against WP3 in this stage.

### 4.2 Significant problem reports

Two significant PRs were raised against WP3 this stage. Both were fixed and closed:

#### PKI for ADMIRE (Trac ticket #43)

**Opened:** 13/03/08

**Status:** Closed as FIXED 12/06/08

**Problem:** Working in a distributed computing project ADMIRE needs some level of security. The minimum requirement for security is an X509 public key infrastructure (PKI).

**Solution:** The ADMIRE PKI now supports X509 certificates from three certificate authorities: the ADMIRE CA, the UK e-Science CA and the SlovakGrid CA.

#### Possible multicast problem at EPCC (#51)

**Opened:** 04/06/08

**Status:** Closed as FIXED 04/07/08

**Problem:** USMT v0 use IP multicast for clients to be able to find the nucleus registry – initial tests of USMT failed on the EPCC nodes of the ADMIRE Testbed.

**Solution:** Two new machines (`admire5` and `admire6`) were deployed on a multicast-enabled network segment at EPCC to get around the problem. However note that this highlighted a risk of USMT relying on multicast which is typically disabled by default (Risk TR14 in [4]).

### 4.3 New risks

No risks recorded against WP3. One new risk (TR14) recorded against WP4 from activities in WP3.

---

## 5 Plans for next Period

ADMIRE WP3's plans for months 7-12 are unchanged from the original Description of Work (See [1] and Appendix A).

### 5.1 WP 3.4.2 – revised use cases and requirements

*Continue support and intelligence gathering; issuing a report with a refined catalogue of scenarios, use cases and quantified user requirements (DoW [1] p49).*

The Use Case and Requirements Report [5] will be expanded and extended and used in the first evaluations of the ADMIRE Architecture, Models and Languages.

### 5.2 WP 3.5.3 – finalise book structure

*Establish Editorial Board and develop an agreed structure and contents for the book (DoW [1] p49).*

The Editorial Board will be finalised and the first agreed structure and contents for the book will be completed.

### 5.3 WP 3.6 – ADMIRE Platform Release 1

*Set up and support revised and extended reference service based on outputs from WP4 and WP5. First ADMIRE Platform release cycle (DoW [1] p49).*

Months 6-12 will see the first test-and-release cycle for the ADMIRE DMI Platform. Most of the work of WP3 will be geared towards testing components as they are delivered from WP4 and WP5. Early in this stage we plan to decide which components will form the baseline Release. This stage will see the advent of the first version of the ADMIRE Testbed production environment – a stable DMI tools platform running ADMIRE Release 1.

## Appendix A WP 3 overall timeline

WP	Time	Description	Status
3.1	M1-2	Initial gathering and description of scenarios and user cases, working with WP6.	Done
3.2	M1-3	Set up user support service, including working relationships with special projects, documentation and help desk, based on partners' existing services.	Done
3.3	M1-6	Set up initial reference services based on partners' existing technologies configured to enable ADMIRE research by M3 and support these to M6.	Done
3.4	M4-6	Work with WP6 and other identified users to develop a clarified and structured tabulation of scenarios, use cases and quantified user requirements.	Done
	M6-12	Continue support and intelligence gathering; issuing a report with a refined catalogue of scenarios, use cases and quantified user requirements.	Next
3.5	M1-9	Establish Editorial Board and develop an agreed structure and contents for the book.	Ongoing
3.6	M6-12	Set up and support revised and extended reference service based on outputs from WP4 and WP5.	Next
	M6-12	First ADMIRE Platform release cycle.	Next
3.7	M13-18	As for step 5 using services deployed in step 6 and setting up new user relationships to fill gaps in requirements identified by users or other WPs.	
3.8	M13-18	As for step 6 using new releases from WP4 and WP5.	
	M13-18	Second ADMIRE Platform release cycle.	
3.9	M18-21	Produce draft Chapter of Scenarios illustrating uses of large-scale and multi-function DMI infrastructures as envisaged by ADMIRE.	
3.10	M18-24	As for step 7 using latest releases from WP4 and WP5 and a possibly refreshed set of collaborating projects.	
3.11	M18-24	As for step 8 using latest releases from WP4 and WP5.	
	M18-24	Third ADMIRE Platform release cycle.	
3.12	M25-27	Produce draft Chapter of quantified user requirements.	
3.13	M25-30	As for step 10 using latest releases from WP4 and WP5.	
3.14	M25-30	As for step 11 using latest releases from WP4 and WP5.	
	M25-30	Fourth ADMIRE Platform release cycle	
3.15	M31-33	Produce draft Chapter on methods of evaluating DMI systems and the results obtained during the ADMIRE project.	
3.16	M31-36	As for step 10 using latest releases from WP4 and WP5.	
3.17	M31-36	As for step 11 using latest releases from WP4 and WP5.	
	M31-36	Fifth ADMIRE Platform release cycle.	

## Appendix B WP 3 deliverable schedule

Deliv.	Time	Description	Status
D3.1	M6	Platform Delivery and Support report against construction of base testbed and installation of pre-existing services.	<i>This document</i>
D3.2	M12	ADMIRE Platform software Release 1	
D3.3	M18	ADMIRE Platform software Release 2	
D3.4	M24	ADMIRE Platform software Release 3	
D3.5	M24	Book Delivery progress report.	
D3.6	M30	ADMIRE Platform software Release 4	
D3.7	M30	Book Delivery progress report.	
D3.8	M36	ADMIRE Platform software Release 5	
D3.9	M36	ADMIRE Book Draft.	