eIUS: Astronomy Final Use Case

In these examples, underlined elements are based either on the original interview data or on feedback sent by the informants in response to draft versions of the use cases. Key *activity types* are highlighted in *italics* and **relevant ICTs** in **bold**.

Narrative

1. Peter, an astronomer based at the University of Edinburgh, is *querying* the **ADS digital library** to find new literature on one of his main research areas, the formation and evolution of galaxies in clusters of galaxies. He just read an **email** from his colleague Betty concerning their planned journal publication, which will report work from a long-term research project in this context. Peter finds some useful publications fitting their data collection and analysis approach and replies to Betty informing her of his progress.
   [use quote: "I'd always (..) [start with] using ADS as (..) [my] route into the literature."]
2. This particular project, a collaboration between researchers at a number of universities in the UK and beyond, started some years ago. It is aimed at *identifying* galaxy clusters at a young, early stage to get more indications of the processes affecting the evolution of the galaxies within them, using various kinds of observational data. In the project's early stages Jeff, a student, was instrumental in *developing algorithms to search* an x-ray data archive, subsequently *building up a* ***catalogue*** **of several hundred clusters of galaxies**. Observations using x-ray data are very useful for discovering clusters of galaxies and the database has been a very rich resource for Peter and his colleagues over recent years.
3. Nowadays various astronomy projects systematically map large areas of the sky in particular bands of the spectrum, and their observation data is loaded up into a database and made accessible online. At one of the f2f project meetings Peter got a very useful hint to a new set of data from a Leicester colleague, who specialises in such x-ray data. The data had recently become available after an 18 month embargo period in which it was restricted to exclusive use by its original owners.
4. To further identify the individual galaxies in the clusters Peter *cross-correlates the x-ray data* with data from **optical and near-infrared archives** (like the **WSA** and **SDSS** archives). The Edinburgh data centre specialises in these data, so Peter can discuss questions with colleagues locally. Based on that he *defines* ***SQL queries*** to *find and filter* adequate data from these databases.
5. During the project also some new observations had to be made regarding clusters without extant entries in the optical archives in order to gather more comprehensive data for the analysis process. Peter over some time worked on this with Betty who has an office in the same building which helped their collaboration in meeting up f2f regularly. As with all archive data they add metadata in *recording relevant information in images*. This can be done manually or in a semi-automatic way *using* ***image analyser code and tools*** *to scan images*

to identify sources which then are *described with basic standard attributes*, e.g. specifying whether they are stars or galaxies, how bright they are, etc.

6. A particular challenge for astronomy is to make archives interoperable and establish standards to make analyses on various kinds of data (optical, x-ray, radio) easier. Peter and a lot of his colleagues make use of the **Virtual Observatory** being developed by the **International Virtual Observatory Alliance (IVOA)** and its UK member **AstroGrid** in trying to adapt to and use standards for describing data, using metadata and access protocols.
   [use quote: The "move into having standard means of access (.) [to] the data (..) means [that, as a researcher,] you can expend all your effort actually analysing the data not just (.) getting it [and integrating it]".]
7. The attributes (metadata) describing sources detected in images in archives are the object of most analysis (not the images themselves). Peter opens the **IDL workbench software** which he has used in the past to *write his own code*. This time, for the evolution of galaxies project, he modifies the code to *analyse attributes* of clusters and individual galaxies in order to find statistical correlations between both.
8. In a next step of analysis Peter *accesses theoretical models* developed by a colleague at another astronomy department which predict the properties of galaxies. This way he can match his observational data against the predictions in these models and, thereby, *test the models*.
9. Peter *gives feedback* to this colleague from his own research practice on how to refine one of these models at one of their rare project **Access Grid** sessions.
10. In the same session the group also discusses and mostly agrees to recent developments in the discipline's dissemination practices to provide not only a final publication but also to link this to part of the actual source data. The **CDS data centre in Strasbourg** is currently the leader for curating data this way in linking articles within the journal 'Astronomy and Astrophysics' to the respective source data. **Nature** offers extra descriptions of the underlying data online.
11. Eventually, after some more work by Peter on data analysis and further exchange with Betty the article is *written up* and *submitted* to the journal, one of the important publications in the discipline.
    [use quote: The paper "illustrates that these projects often (.) run over many years and have different stages in them: of analysing data which are in archives, then actually making new observations, analysing those data and then using those data and to think (.) what are the (.) [further] observations which are needed and then making new observations and analysing those data and that's sort of how it all moves on."]

Relevant ICTs

ICT	Comments
The SAO/NASA Astrophysics Data System (ADS)[1]	ADS is a Digital Library portal for researchers in Astronomy and Physics, operated by the Smithsonian Astrophysical Observatory (SAO)

	under a NASA grant. It contains scans of older journal articles and a database of online articles. The research process usually starts with querying this online repository.
The WFCAM Science Archive (WSA)[2]	The WFCAM (Wide-Field CAMera) Science Archive is curated by the Wide-Field Astronomy Unit in Edinburgh and houses data from the UK Infra-Red Telescope in Hawaii (UKIRT).
The Sloan Digital Sky Survey (SDSS) / SkyServer[3]	The SDSS is one of the largest surveys of the night sky and provides the SkyServer online database with images, spectra, photometric data and spectroscopic data as well as various tools to query and filter this content.
SQL (Structured Query Language)	Computer language to query databases in a structured way
Code and tools to scan images	Various standardised or non-standardised software packages and languages (e.g. Java) to create the description metadata attributes for images in archives
The International Virtual Observatory Alliance (IVOA)[4] and AstroGrid in the UK[5]	The international Virtual Observatory initiative is coordinated by the IVOA and aims at achieving interoperability and developing standards for the community by embedding the existing standards for describing data, using metadata (registries) and access protocols. The UK member AstroGrid also hosts a community portal and offers a free downloadable Desktop Suite of astronomy applications.
IDL[6] data analysis and visualisation software package	IDL is a licensed software package used for the analysis and visualisation of images (or more precise their metadata attributes) in Astronomy. It includes the IDL workbench, a development environment.
Access Grid (AG)	The AG is an advanced videoconferencing system making use of large-format displays and multimedia support. It interfaces to Grid middleware.
The Astronomy Data Centre in Strasbourg (CDS)[7]	The CDS enables links between articles published in the Astronomy and Astrophysics

[1] http://adswww.harvard.edu/
[2] http://surveys.roe.ac.uk/wsa/
[3] http://cas.sdss.org/astrodr7/en/
[4] http://www.ivoa.net/
[5] http://www.astrogrid.org/
[6] http://www.ittvis.com/ProductServices/IDL.aspx
[7] http://cdsweb.u-strasbg.fr/

	journal and the underlying source data.

Commentary

This use case exemplifies the research lifecycle in Astronomy in the area of formation and evolution of galaxies and clusters of galaxies based on a long-term project. The examples have all been provided by the interviewee and only some details have been added to provide the flow of the narrative. Furthermore only the student has been mentioned explicitly as a person, other persons are based on the provided data on collaborations in the project and in the community in general.

Not mentioned in the narrative: The interviewee also discusses the level of service provided by the Wide-Field Astronomy Unit in Edinburgh for their repository to enable users to do plots of histograms or scatter plots (i.e. graphical representations of data sets) which are then just displayed in a browser and the dataset as a whole does not have to be downloaded (the data stays on the institutes server). These features are considered to be still quite basic, it is not possible to address more complex astronomy questions, which at the moment require the download of all the data for further analyses.

It has to be pointed out that resolving of research questions in astronomy usually is a long lasting process independent of the sophistication of the e-infrastructures. Proper tools can help to make single steps along the way less cumbersome.

Comments by Informant

The interviewee provided extensive feedback on the narrative and the underlying experience report which has been incorporated into those documents.

Other Editorial Considerations

Element	**Usage**
Links to direct quotes?	Yes
Year?	No
Month?	No
Time of day?	No
Location given?	Yes
Real institutions named?	Yes
Real journals named?	Yes
Real conferences named?	No