Mem. S.A.It. Vol. 80, 342

Memorie della

Theory in the Virtual Observatory

G. Lemson[1,2] and J. Zuther[2,3]

[1] ARI-ZAH – Moenchhofstrasses 12-14 Heidelberg Germany
[2] Max-Planck-Institut für extraterrestrische Physik, Garching, Germany
e-mail: gerard.lemson@mpe.mpg.de
[3] I. Physikalisches Institut, Universität zu Köln, Köln, Germany

Abstract. In this contribution we introduce the subject of this workshop by giving an overview of efforts that are underway to introduce the results of theoretical astrophysics in the international Virtual Observatory (VO) efforts. To this end we shortly discuss the VO as a project and why theory deserves special attention. This is followed by an overview of activities in this field, both by regarding standardisation of theory in the International Virtual Observatory Alliance (IVOA), and by presenting some examples of theory services. We end by summarising the conclusions of the white paper written in the context of work package 4 of the EuroVO DCA project, the work package that has organised this workshop.

Key words. Cosmology: techniques - Virtual Observatory

1. Introduction

This workshop is organised in the context of work package 4 (WP4) of the EuroVO Data Center Alliance (DCA)[1]. The task of WP4 was to investigate the inclusion of theory data and services in the general Virtual Observatory (VO, VObs) framework. In this contribution we will give an overview of this effort. To this end we first shortly review VObs activities in general. From its inception these have mainly targeted observational astronomy. Its efforts and procedures can in general not be directly applied to the results of theoretical astrophysics. We will discuss the fundamental reasons for this and thus show that extra efforts must be undertaken to remedy this situation. We will then review what efforts have been undertaken to allow the inclusion of theory in the VObs framework. This includes both efforts at standardisation in the context of the International Virtual Observatory Alliance (IVOA[2]) and individual activities in national VO projects, and also in the EuroVO DCA.

Before continuing we will limit the scope of what we mean by *theory* in the VO context. We are mainly interested in computational astrophysics, as this is the part of theory that produces results ("data") that are open for handling in the VO framework. And while we thus "ignore" analytical results, we are interested in computations on all scales, from the largest cosmological simulations, producing results easily exceeding that of the largest observational surveys to date, to small scale calculations of stellar evolutionary tracks or model spectra. The latter are actually of interest as they may be published as online simu-

Send offprint requests to: G. Lemson

[1] http://cds.u-strasbg.fr/twikiDCA/bin/view/EuroVODCA/WebHome

[2] http://www.ivoa.net

lation services, whereas the former require sophisticated archiving techniques.

2. Virtual Observatory

Under the term *Virtual Observatory* are grouped a host of internationally organised efforts that aim to take advantage of the latest Internet technologies to facilitate the access to results of astronomical research. Particular attention is given to the data on which most of that research is based, but also visualisation, analysis and other real-time services are being investigated. The goal is to develop a system that allows astronomers everywhere to gain access to their colleagues results as if these reside on their own desktop computer.

The benefits of such a system are manifold. For many scientists a written publication, describing observations, data reduction and analysis is still the end point of one's research project. The data in general find representation as figures or diagrams, sometimes a table. For this to be the starting point of new research by others it would be of great benefit if those third parties had access to the actual data themselves. This allows one to check the results derived from the data, analyse the data with new techniques, compare them to similar data sets etc. Of special interest is the combination of results about the same objects, but obtained in different observations, with different instruments at different wavelengths, possibly at different times.

The professional astronomical community has actually always been at the forefront of publishing and sharing its results through online means. What is new in the VO effort is the emphasis on *interoperability*. With interoperability is meant the ability of different services on different, geographically distributed systems to communicate with each other, exchange and understand each others data products and protocols.

Whereas it is relatively straightforward to makes one's raw or processed data sets available for download via FTP for example, it is much more work to do so in a usable fashion that allows this interoperability. Fig. 1 illustrates the difficulty a prospective user is faced

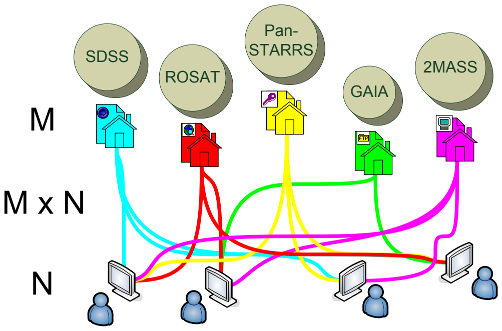

Fig. 1. Babylonian confusion. If every user needs to understand the protocols and data formats of all published resources, the total effort scales quadratically.

with when trying to use resources from different providers. Each publisher designs their own web page, with its own interface, sometimes offering individual search options. These all give access to files or databases with a whole range of possible formats for representing the data, which each must be handled by the end user's own analysis software. And whereas for a single service this is manageable, when it scales ot many services the total effort scales unfavorably.

2.1. Standardisation

Standardisation is the proposed solution to the problem that different archives all present their data in their own specific form. We could view this problem as if each archive presents itself to the world in its own language. In this analogy, for clients to use results from an archive requires the client to understand the archive's languages. This can easily lead to a Babylonian confusion, where each user must understand all languages. In the real world the solution to such a problem has been for participants in the discussion to settle on a common language. Though in reality Esperanto has not caught on it provides the proper analogy. Standards are the Esperanto of the VO, and actually of many information integration efforts.

As illustrated in Fig.2, the introduction of standards could change the amount of effort required for a single user to only learning a com-

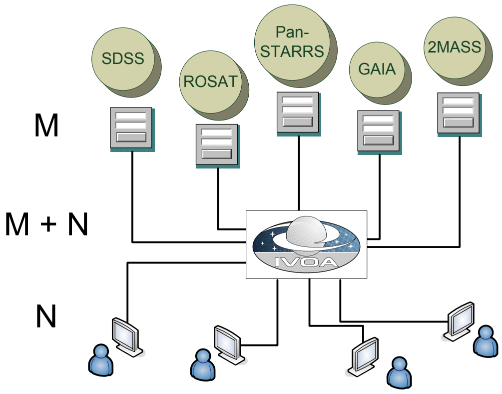

Fig. 2. The inclusion of "Esperanto", i.e. standards, in the situation illustrated in Fig. 1 turns the quadratic scaling into a linear one. Each archive and user needs to "speak" a single language. Of course what needs to be taken into account is the work required to define the standards!

mon standard, implemented by all providers. In contrast to the real world, where English is the defacto Esperanto in many cases, in the VO an Esperanto has to be developed, i.e. the standards have to be defined. Users have to be willing to spend some effort on learning some new techniques. The main other work is on the data provider. Instead of their favorite internal formats they will have to be prepared to offer their resources in the standard defined by the VO. This can in fact be a lot of work, *if* this work is done after the science has finished, at the point of publication. If however the requirements of publishing results to the VO are taken into account at the inception of the project, it need not cost much more effort, can in fact be a one-off effort, usable for multiple projects. But it does require some persuasion from the community to ensure that even this little effort is performed.

2.2. IVOA

The IVOA is the body in charge of the standardisation efforts that are required to make the VO come to life. It is an alliance of the various national VO projects. Its work is divided in a number of working groups, each responsible for developing standards in specific areas. The tasks of these working groups are to write documents with formal definitions The following is a subset of the working groups that will be relevant later on when discussing theory:

- **Registry :** Defines standards for a database containing resource descriptions. Includes standard for resource description ("meta-data") and interface for querying a registry.
- **Data Modelling :** Defines standards for describing features of data and metadata. Current examples are the Space Time Coordinates and the Characterisation data models.
- **Data Access Layer :** Defines protocols for accessing, that is querying for and retrieving of, data products. Most important are protocols for images (Simple Image Access Protocol, SIAP), spectra (Simple Spectral Access Protocol, ssap) and source catalogues (Simple Cone Search, SCS).
- **Semantics:** Defines shared vocabularies of words to be used by other standards when referring to a particular concept.
- **VOTable:** Defines and maintenains a standard for tabular data transfer.

There are other working groups for application, event notification and for the standardisation process as a whole.

There are also so called *interest groups*. These represent the interests of a particular subset of the community to the VO. Requirements can be collected and communicated to one or more particular working groups with the request to either create standards, or, more frequently, to ensure that in a standard-under-construction the particular requirements are taken into account. The theory community is represented by the Theory Interest Group (TIG), which will be described below.

3. Theory in the VO

3.1. Motivation

Before discussing the *how* of publishing theory,we should discuss why we think it is interesting to attempt this at all? Many of the reasons are the same as they are for observational results, i.e. for scientific results in general. Publishing your results online...

- gives your colleagues access to the data described in your publication, thus providing an extra motivation to extend your work, which leads to new science and increases the publication's impact;
- enables scientists not directly involved in your project to verify your results, which in general helps to find mistakes, thus improving your results;
- provides others with a benchmark to which to compare their own results, both for reasons of reproducibility and for checking their research;
- is increasingly mandated by funding agencies;
- can serve as showcase for future proposals;
- may facilitate the refereeing process, by giving referees access to the data (possibly not yet public);

Publishing your results in the VO:

- makes your results available in a standardised manner, facilitating their discovery and greatly increasing their re-usability;
- enforces a good practice to follow VO-like standardisation: it forces you to think carefully about your own results, which improves re-usability even for you;
- is the proper thing to do if you also want to use the VO for your own research, "what goes around, comes around";
- may not give obvious benefits to you, but you may agree that it is good for people to think about how others should publish their scientific results in a homogeneous manner, so that you and others have an easier job interpreting and using these.

In any case, having your results seen and reused by others will increase the impact of your research.

On top of all of this, publishing the results of *theoretical* research to the VO...

- allows others to compare observations to models, facilitating their interpretation, or enabling more sophisticated predictions (survey planning, improved exposure time calculators, ...);
- allows one to see physical processes in action;
- allows fellow theorists to compare and test their analytical models on state of the art simulations;

3.2. Issues when standardising theory

The main challenge for the VO to be able to support theoretical data is that in general they are very different from the observational data products that have so far been mainly handled. Though it is often possible to create synthetic images, spectra and source catalogues, these are generally the results of post-processing of more fundamental simulation results which should also be handled. And there is one big difference between publishing these virtual observations and the real ones: they do not observe the same universe. Hence in contrast to data access protocols like SIA, SSA and SCS there are no common coordinate systems to use in protocols to find familiar objects on a common sky; there are in fact no common objects to be identified.

So whereas in principle publishing mock observations should be as straightforward as making real observational data VO compliant, in practice specific standards must be developed, or existing standards must be improved to ensure a proper handling by the VO. In any case, reducing the contribution of theory to the production of mock images or spectra would limit the possible impact of publishing theory in the VO. Indeed, the virtue of the theoretical results is to provide an in-depth understanding of the physical and chemical processes at play in Nature, most of the relationships between the various components of a simulation, and undoubtedly the most interesting ones, being unobservable. Good examples are the structure of stellar interiors, the dark matter and dark energy component included in most N-body simulations, or the ionization structure in photoionised regions.

Hence we believe that it is still true what was claimed already in (Lemson & Colberg 2004), namely that theory requires special attention within the VO framework. The focus on observational astronomical data sets and services of the VO in general and the IVOA in particular has had the consequence that many of the current IVOA data models and data access standards are irrelevant for theoretical data products. The most important parameters for the query protocols for discovering interesting data sets contain positions on the sky, the IVOA query language contains definitions for regions on the sky, data models characterise the spatial, temporal and/or wavelength extent of observational data sets, registry resources can indicate their footprint on the sky. In the few cases where theory resources have been taken into account in the construction of standards (for example the inclusion of theory spectra into the Simple Spectral Access Protocol), in general this has not happened in great detail yet. It is good to investigate this issue in some more detail, as it will indicate also the problems that must be solved when supporting theory with standardisation.

One of its causes is that the ideas for the VO originated in the observational community, and people from that community were the first developers and obviously their requirements lead to the first standards. But one can argue that the observational community is a priori more suitable for the VO framework in many ways:

1. Simple observables. Observations described using small number of parameters (space/time/wavelength/flux/polarisation). Hence standardisation is relatively simple, at least wrt. the data part. It is in the definition of appropriate metadata descriptions that most effort is spent.
2. Common sky. When looking at some part of the sky, one can expect to see the same objects as other observations of that same part of the sky.
3. Small set of observatories (of the order of a few hundreds). These are reused by many scientists; so many different scientists have very similar data products to begin with. Instead of re-observing, might first look at archive. Not much work for these to agree on common standard, since FITS has been a common format for astronomical data for many years.
4. Archiving obviously useful. Even 100 years old observations may be of use to present day astronomers. Consequently data centres existed before the VO concept was explicitly announced, and they contained expertise of use to VO development.
5. Large, relatively homogeneous community. Most astronomers are observers, all of whom are interested in the resources available. Moreover, though clearly there are large differences of specialisation in different wavelength regimes, to first order science-ready observations can be compared to each other by anyone, if only by overlaying one image on top of another. Consequently use cases abound for interoperability.

Though one may argue about the detailed validity of these points, we believe they are definitely true when compared to the analogous situation in theory:

1. Complex "observables". Anything that can be imagined can be modelled and simulated. Consequently much more complex to standardise.
2. No common worlds. In general, computer simulations start from random initial conditions and in the general case it is impossible to identify common objects in different simulations. Hence any possible interoperability between different services must necessarily follow a different pattern.
3. Every computer is a laboratory. Every computer can be used to produce data from some simulation run on it. And everyone who knows how to program computers can create a simulation package for relatively little costs. Hence the potential for heterogeneity are very much larger than in the observational world. In some cases ad hoc standards are developing, but this is mainly true for the more complex types of simulations. For example cosmological simulations, for which there are only a few software packages that are freely available and reused.
4. Moore's law. Old simulations can be redone after a few years using cheaper, faster resources. So special reasons must be found why one would spend a lot of atten-

tion (and resources) to archiving one's results.

5. Very diverse, self-reliant community. The freedom one has in modelling naturally leads to great diversity in models. Furthermore many theorists/simulators will rather write their own codes than use someone else's results. Alternative use cases have to be found that are of sufficiently general appeal to warrant creating standards for.

3.3. Theory in the IVOA: work of the Theory Interest Group

To oversee the application of the VO framework to theory, and to propose theory specific standards efforts to the IVOA the Theory Interest Group (TIG[3]) was founded. It was initiated by (Lemson & Colberg 2004) which was the first attempt to discuss the introduction of theoretical results and data into ongoing VO and in particular IVOA activities. The white paper had two main target audiences. The first was the community of theorists who either wish to take the initial steps necessary to publish their results online, or who want to make their existing online presence "VO compatible", in both cases ensuring consistency with well defined standards. For this audience, the authors attempted to describe existing VO efforts, with particular emphasis on the standardization efforts embodied by the IVOA. The second audience comprises developers working in the various VOs, especially those involved in the IVOA. For this group, the authors described how theoretical archives and related services imply interesting new requirements on these efforts.

The TIG has as its charter to:

- Provide a forum for discussing theory specific issues in a VO context.
- Contribute to other IVOA working groups to ensure that theory specific requirements are included.
- Incorporate standard approaches defined in these groups when designing and implementing services on theoretical archives.
- Define standard services relevant for theoretical archives.
- Promote development of services for comparing theoretical results to observations and vice versa.
- Define relevant milestones and assign specific tasks to interested parties.

Since its formation the TIG has been the most active interest group in the IVOA with many well motivated participants. Starting with a project to gather use cases specific for theory, it has recently concentrated on the preparation of standards for accessing large scale simulations.

One of these is the Simulation Database (SimDB). SimDB is an online service offering query capabilities to a database containing meta data describing results of simulations and their post-processing as well as about the codes used in these algorithms. The simulations are supposed to be those that produce a representation of 3+1D space, (possibly reduced spatial dimension through assumptions of symmetry). A SimDB also contains information about web services giving access to the simulation results themselves. The more detailed specification of such services is the goal of the Simulation Data Access Protocol (SimDAP) specification. For developments on both these standards please check the TIG pages on the IVOA website.

3.4. Theory in the EuroVO DCA

EuroVO DCA had a work package especially devoted to theory. Its two deliverables were the organisation of this workshop and a white paper describing a framework for inclusion of theory in the VO. The conclusions of the latter will be discussed in the concluding section of this contribution. Several of the partners in EuroVO DCA also had their own theory related activities, some of which have been described elsewhere in these proceedings (see contributions by Allen, Caniglia *et al*, Lemson & Zuther, Manzato *et al*, Rodrigo and Wozniak).

An important activity of the EuroVO DCA was the census of data centers. This census contained an explicit questionnaire for theory

[3] http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaTheory

archives and services. Here we shortly summarise its results, which give an overview of the state of theory work in Europe. There were 24 responses concerning simulation archives and 10 responses concerning services. The census indicates that there is quite a range of types of simulations published online already. These can be classified as large-scale simulations and micro simulations. The former are primarily related to cosmological simulations of dark matter halos and galaxy mergers. The latter ones model and simulate the microphysics of mainly stellar spectra and stellar populations. Other simulations are related to molecular line libraries, astrophysical jets, stellar accretion disks, quantum mechanical calculations of Stark broadening, and web links to other theory web sites.

The most interesting difference with respect to observational archives/services is the type of data provided. Simulations in general do provide more physical data than observational data. Their structures are different from those of classical images or spectra. Whereas, e.g. spectral line shapes are data types relevant to both observations and simulations, there are many other, non-classical data types like hierarchical merger trees from galaxy evolution, isochrones for stellar evolutionary models, quantum-mechanical parameters, particle properties (coordinates, velocities, masses, etc.), html links to other theory web sites, and physical parameters of jet simulations. Theoretical services are mainly related to the observational interface to the simulations.

The interest in adopting VO standards for metadata and data access is high. The level of VO compliance is varying, but almost all archives offer at least VOTable access and simple access protocols (SSA, SIA). Two archives have implemented the current versions of the theoretical database model and data access (SimDB and SimDAP). And about 14 implementations exist of S3, a proposal for a simple protocol for describing simple HTTP services. As an example of such efforts we will now describe a few theory services that have been implemented by our VO project, GAVO.

4. Theory in the German Astrophysical Virtual Observatory (GAVO)

Standardisation efforts form the core of the international VO activities. The fact that within the IVOA no theory specific standards have been developed so far, has not stopped people from investigating how to publish theoretical results on line, and implementing corresponding services. Some of that work was done within the context of national VO projects. Those in particular have in general tried to implement services which aim to follow the VO philosophy and have been in general focused on the research fields covered by the national community of theorists, or some projects in challenging astrophysical computation. Here we describe a few of those that have been implemented by the German Astrophysical Virtual Observatory (GAVO[4]). Many of the contributions in these proceedings point to similar such activities.

GAVO has from its inception had special attention to theory. This resulted in a number of services, some prototypes, aimed at illustrating concepts, others fully supporting science cases.

4.1. Millennium DB as example for TAP/ADQL

The *Millennium Database*[5] is a service created by GAVO and the Max-Planck Institut für Astrophysik (MPA). The service exposes a relational database containing post-processing products of the *Millennium Simulation* (Springel et al. 2005). Users can submit SQL queries to the database and receive the results back. In this it is similar in behaviour to what the purpose is of the Table Access Protocol (TAP, under development in the IVOA Data Access Layer working group[6]) together with the Astronomical Data Query Language (ADQL) of the query language working group[7]. In particular most users

4 http://www.g-vo.org

5 http://www.g-vo.org/Millennium

6 http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaDAL

7 http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaVOQL

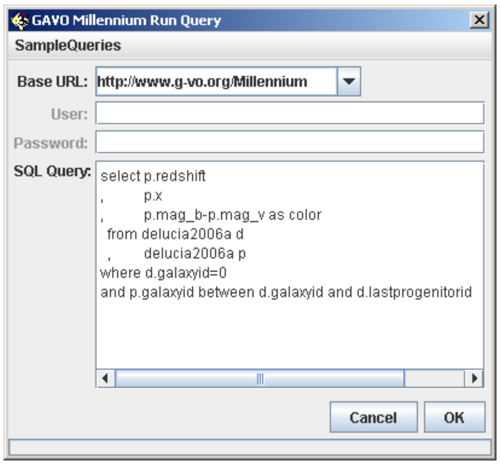

Fig. 3. TOPCAT's SQL query entry window for accessing the Millennium database.

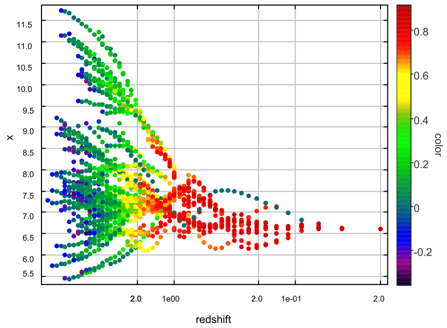

Fig. 4. Result of the query in Fig. 3, plotted in TOPCAT. The X axis show redshift, Y axis the X-coordinate of the progenitor galaxies of the descendant galaxy at the final redshift at the extreme right. Color indicates the B-V color of the simulated galaxies.

do not use the web browser interface, but create an HTTP request as a URL string and send this through command line tools such as wget. This feature was explicitly added to the service so that users could write scripts that query the database multiple times and handle the results.

An interesting feature of many theoretical data sets is the occurrence of data structures that one will in general not find in observational data sets. They offer therefore interesting applications of standards such as ADQL and TAP. The data products published by the Millennium Database for example contain hierarchical structures and strong cross-linking. The galaxies and dark matter halos are organised in merger trees and special techniques had to be invented to make those efficiently queriable from the database.

Another interesting feature is that the popular TOPCAT[8] visualisation tool has a plugin feature which allows users to query the Millennum database explicitly (see Fig. 3). In Fig. 4 we show the result of a query run from TOPCAT and displayed in 3D.

Since the service was announced via the Los Alamos preprint server in August 2006 (see Lemson et al. 2006), the service has seen extensive usage. About 6.5 million queries have been submitted, in total close to 70 billion rows have been returned. More importantly, close to 200 papers[9] have appeared dealing with Millennium simulation data, a significant fraction (30%) has used this online service for obtaining the data.

4.2. Virtual telescopes

A recurring concept in the theory VO is that of the "virtual telescope". A virtual telescope is a piece of software that simulates the observation of a real physical object, by "observing" a simulation of the object, including instrumental effects where possible. An important, almost defining feature is that the results are similar to results obtained by real observations. These may be mock images, spectra, source catalogues. The idea is that it is easier to apply instrumental effects and noise and such to idealised, synthetic observations of a simulated object about which everything is known, than it is to subtract such effects from "dirty", real observations.

As an example GAVO has published a simplified simulator created by the Planck group at MPA[10]. The code produced by the group is easily wrapped with a web interface that al-

[8] http://www.star.bris.ac.uk/~mbt/topcat/

[9] http://www.mpa-garching.mpg.de/millennium/

[10] http://planck.mpa-garching.mpg.de/

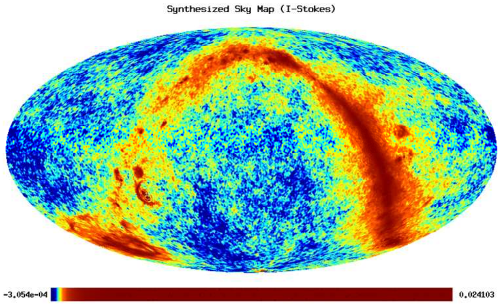

Fig. 5. The CMB sky simulated with the default settings on GAVO's Planck simulator website at http://www.g-vo.org/planck.

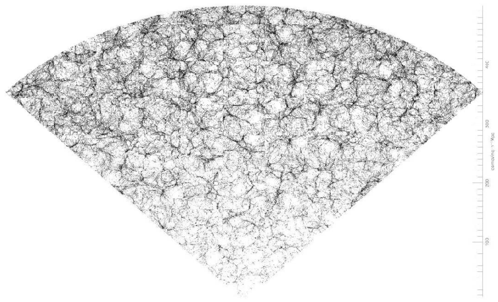

Fig. 6. "Light-cone" cut out through the dark matter halo distribution in the milli-Millennium simulation, performed by GAVOs MoMaF2 virtual telescope web service at http://www.g-vo.org/mpasims/MoMaf2.

lows a user to set various parameters that configure the mock-Planck satellite. It furthermore allows one to add particular physical foregrounds which will also exist in the real signal. The final result is a FITS file containing the predicted signal as a map on the sky in a format that is equivalent to the way real maps will be produced. An example of such a result is the full sky intensity map in Fig. 5.

Another example of the principle that it is easier to add problems to "clean" data than to remove them from "dirty" data is used when calculating "observed redshifts" for a simulated catalogue extracted from a cosmological simulation. It is easy to add effects from peculiar velocities to the simulated sample than to correct for these in an observed redshift survey. Examples of this can already be found in the pre-calculated mock catalogues published in the Millennium database. An online service has existed for a while based on the MoMaF code (Blaizot et al. 2005). It is again a trivial matter to wrap such code with a web interface, see Fig. 6 for an example of this virtual telescope. An updated version working on the Horizon (see Wozniak in these proceedings) is in development (Blaizot, private communications).

5. Conclusions

The activities of EuroVO DCA WP4 have led to a white paper discussing a framework for the inclusion of theory data and services in the VO. The two main conclusions of that white paper were

- Simulation specific data access protocols such as SimDB and SimDAP should be developed as a matter of urgency
- In order to encourage take up of Theory-VO services we must rapidly develop proto-type Theory-VO services allowing access to scientifically useful theory simulations and models.

We concluded that it is well possible to introduce theory in the VO framework, and the effort is on-going. But issues remain to be resolved. The architecture of the VO needs no adjustment to accommodate theory. A resource registry for discovering interesting resources is as relevant for theoretical data products and services as it is for observational ones. Data models are required (and can be constructed) for describing theoretical resources. The approach to data access protocols based on a discovery and a retrieval phase is suitable for theory as well. The IVOA query language is relevant for filtering the sometimes very large data sets. Services can in principle be deployed, protected and chained together in workflows according to the standards of the Grid and Web services working group.

It is therefore mainly in the details of this process that work remains to be done. Particularly, the standardisation for data description, discovery and data access protocols require more work above and beyond what is currently available in the IVOA. The reasons for this are simply that the existing standards so far have mainly dealt with observational data products and theory data products can be very different and very diverse. Work along these lines is progressing though. The SimDB and SimDAP standards follow the IVOA approach for what we call cosmological simulations. Efforts have started in the IVOA to investigate their applicability to other types of simulations such as isochrones, stellar evolution etc. It seems likely that these can be served with only minor changes to the models and data access service specifications, but alternative approaches are being explored as well such as a simple self-described service protocol.

The question that remains is whether developments in this area will be sufficient for the astrophysical community to start participating. This may in the end be the greatest challenge, convincing theorist to make the extra effort of conforming to the standards and publishing their results in the VO framework. Willingness for participation expressed by the respondents to the Census of European astronomical data centres is a very good starting point. Here is important to develop successful prototypes and to assess relatively simple manners to take those extra steps. By financing visits and workshops and in that way enabling and starting collaborations, projects such as the EuroVO are of great importance in this development.

Acknowledgements. Thanks to the members of the EuroVO DCA Theory Experts Group for the collaboration in writing the white paper that was the basis for much of the current document. Thanks to Jai Won Kim for careful reading of the document. GL works for the German Astrophysical Virtual Observatory (GAVO), which is supported by a grant from the German Federal Ministry of Education and Research (BMBF) under contract 05 AC6VHA. JZ acknowledges support from the European project EuroVO DCA under the Research e-Infrastructures area.

References

Blaizot, J. et al. 2005, MNRAS, 360, 159
Lemson, G. & Colberg, J. C. 2004, IVOA white paper, "Theory in the VO" http://ivoa.net/pub/papers/TheoryInTheVO.pdf
Lemson, G. and the Virgo Consortium, 2006, astro-ph/0608019
Lemson, G. et al. 2008, EuroVO DCA WP4 white paper, "Framework for the inclusion of theoy data and services in the VObs".
Springel V. et al. 2005, Nature, 435, 629