## Preamble for GROC consideration

The Data Area met at noon Monday 6th October at GGF9. We propose the workshop described below to take place on the Sunday preceding GGF10 in Frankfurt. The details below were agreed for the purpose of presentation to GROC, but the chairs and programme committee may refine it. The programme committee will represent the groups in the data area. There will be additional members to bring in external topics which should influence our thinking and systems.

We believe that this consistent with GROC requirements on workshops:
1. It serves to communicate with others in GGF the work, achievements and plans of the Data Area, and
2. It serves to communicate externally the achievements and challenges in this area.

## Preamble for JoGC Editorial Board

We propose to run the following workshop and to select and shepherd papers to develop a collection appropriate for a special issue. The discussions will develop editorial material. The *Data Area Structure and Function Analysis* will provide a basis for a good survey and framework. The selected papers revised and shepherded after the workshop will provide good quality papers to complete the issue. I therefore send this note to ask if we can develop the idea of the workshop leading to a special edition of the journal.

# The Future of Grid Data Environments:

# A Global Grid Forum (GGF) Data Area Workshop

Sunday 7th March 2004, Frankfurt, Germany, GGF10
Chairs: Malcolm Atkinson, Ann Chervenak, Andre Merzky & Susan Malaika

Please note: This draft should be revised by the chairs and the programme committee.

## Background and Objective

The area of data in grid environments has become a very active at GGF, with new groups forming at every GGF meeting. In view of the increased activity level, a *Data Area Structure and Function Analysis* report for the GGF Data Area is being prepared. A first draft will be available for review by the GGF chairs in October 2003 and revised for public review by 21st Jan 2004. It will be presented as part of this workshop at GGF10 in Frankfurt, Germany on 7th March 2004. The Data Area group chairs, BoF leaders and Group chairs developing data-related standards in other GGF Areas will be invited to comment. Research and experience papers will solicited by a call and be selected by a programme committee for presentation at the workshop. A selection of those papers revised as a result of the workshop will be published as a special issue in the Journal of Grid Computing.

This workshop is intended to shape the future standardization work for the data in grid environments at GGF. During the morning we will communicate the *consolidated*

achievements and plans of the Data Area.  In the afternoon we will review these in the light of ongoing research.  We expect to receive set of high quality papers identifying important results and exposing challenges.

The workshop will be open to anyone registered for GGF10.

## *Format*

The workshop will run for one day in two parts.
Part 1 – **Data Area Structure and Function Analysis**
- Presentation of the *Data Area Structure and Function Analysis*
- Observations and responses from the GGF Groups
- Discussion to identify consensus and required revision

Part 2 – **Research and Experience Papers**
- Presentation of papers
- Discussion session to review research and experience impact on the Data Area's road map.

## *Submission of abstracts*

We invite research and experiences papers to be presented in part 2.  We particularly encourage papers that address topics such as:
- Data Virtualization, Access, Integration and Transformation
- Data Discovery, Quality, Provenance and Metadata
- Data Movement, Distribution, Caching and Replication
- Data Management and Quality of Service
- Data Recovery and Coordination
- Data Experiences in Grid Environments

Please send your abstracts to KKK by 25$^{th}$ November 2003.   Acceptances will be notified by 15$^{th}$ December 2003.  Full papers are required by 8$^{th}$ February 2004.

## *Considerations*

We expect presenters to remember that the goal of the Data Area is to develop compatible standards well-designed for grid application requirements.  It would be helpful if presenters, particularly in the morning session, considered the following:
1. Are the requirements you address clear?
2. Is the vocabulary for their area already well defined and consistent across the Data Area or are new concepts and definitions required?
3. Is there existing work in the Data Area on which this builds or is this an identification of new area of R&D which should be considered?
4. Is the interaction of their topic with other systems: scheduling, resource management, work flows, etc., well understood?
5. Are the scale and dependability parameters of their application or technology well characterised – where possible specifically quantified?

## *Important Dates*

25 Nov 2003:  Abstract submission
15 Dec 2003:  Notification of paper acceptance
21 Jan 2004:  Public Review of *Data Area Structure and Function Analysis* starts

8 Feb 2004:   Public Review of *Data Area Structure and Function Analysis* ends
8 Feb 2004:   Full Papers Required
7 Mar 2004:   Workshop

## *Program Committee*

Chairs of WGs and RGs in the Data Area or their representatives.
>So far
>>Vijay Dialani  DAIS WG, University of Southampton
>>Paul Watson   DAIS WG, University of Newcastle

Others invited to provide stimulation:
>So far
>>Beth Plale      Indiana University