Subject mediation approach for scientific problem solving in Virtual Observatories

Leonid Kalinichenko

Institute of Informatics Problems, Russian Academy of Sciences,
Vavilov street 44, 2, Moscow, 119333, Russian Federation
email: leonidk@synth.ipi.ac.ru

There exist two principally different approaches to the organization of problem solving in VO: (*i*) information resources driven approach (choice and integrated definition of resources are made independently of the problem specification); and (*ii*) scientific problem driven approach (a specification of a problem domain is created, the relevant to the problem resources are identified and semantically mapped into the domain). Intrinsic difficulties of the first approach: semantic gap between resources and the problem, instability of global schema w.r.t. a set of resources, inability of automatic identification of resources for the problem. To implement the second approach a mediation technology is required. On the consolidation phase of the mediator the efforts of the scientific community are focused on the problem definition by specifying the mediator. During the operational phase relevant information resources are identified and expressed in terms of the mediator. Advantages of the mediator approach include truly semantic integration of heterogeneous resources due to their semantic mapping into the mediator; multiple subjects can be semantically integrated applying recursive structure of the mediators.

Basic methods supporting the mediation technology include: (*i*) identification of relevant resources by metadata, ontologies, information structure, behaviour while registering in mediator resource classes in terms of the mediator classes; (*ii*) rewriting queries (rules) expressed in terms of a mediator into rules over relevant data and services; and (*iii*) running workflows with rules as tasks over distributed heterogeneous resources.

In May 2005 the RVO information infrastructure (RVOII) project report has been published (joint effort of SAO RAS, INASAN, IPI RAS). Basic principles of RVOII are defined as: representing infrastructure as a network of interoperating web (Grid) services, moving processing to data, encouraging code reuse and composition in SOA, emphasizing subject mediators to support various subject domains in astronomy. A Community centre in Moscow for support of scientific astronomical problem solving over distributed repositories of astronomical information has been created. Currently the Centre includes two AstroGrid installations, one of them at the Joint Supercomputer Center of RAS to make possible call supercomputer tasks from workflows. First experience shows that usage of AstroGrid as the RVOII core is suitable according to the RVOII principles.

The mediator approach shows a way for considering IVOA artifacts as well conceptualized and consolidated facilities (not as separate entities – data models, DAL protocols, database schemas, query languages, problem formulation models, applying different modeling facilities). Thus, the early IVOA document "A unified domain model for astronomy" has been attempted actually as a mediator definition for the whole astronomy. In IVOA such sort of models is not supported by suitable methodological idea. IVOA specific data types and classes of problems (e.g., Spectra, Quantity, Passband, Simulation) look conceptually better as specifications of parts of respective mediators. IVOA standards for uniform access to heterogeneous VO data (e.g., image, table, spectrum, etc.) look more consistent as the mediator layer.