Abstract

of attestation master's degree work

subject:

"Semantic Grid for molecular science"

by Tetiana Osmanova

The purpose of work

The purpose of the given work is to research and analyze existing models of semantic Grid in molecular science.

Urgency of spent researches

Scientists in many disciplines use molecular modeling to obtain information about the properties of molecules. The studies of molecular structures are one of the main goals of modern science. Their characteristics and properties are widely used in physics, chemistry, medicine, biology, genetics, paleontology, engineering and other sciences.

Over 30 million chemical compounds are known, many of importance in healthcare, biosciences and creating new materials. It is of fundamental importance to know their properties, including implications for safety. Expanding knowledge base of biomedical domains is reflected in growing volumes and complexity of clinical data generated and utilized in contemporary health services. Most of the required characteristics can be calculated using Sredinger equation, but the number of calculations at this extremely high. The calculation of equation is not the main difficulty, is more complex analysis of the results: data and codes - a very heterogeneous without ordered structure. Therefore, recent research - research initiatives is widely used Grid - calculation for simplicity and speed calculations. But still remains the problem of storage, because every science - research laboratory, the results of their calculations stores locally that impedes progress in finding the necessary information several times.

At the chemical community, the main source of data stores in individual research papers in traditional chemical laboratories, complemented by the arrival today of highly automated synthesized and analytical technologies (particularly in the industrial sector). The data generated in laboratories and distributed throughout the global community, each laboratory facilitates the discovery that consolidated the global literature. You must create a system that would do so to gather the world's inventions in molecular studies, thus to facilitate access to and search. This tool and can develop a semantic Grid - infrastructure.

Tasks solved in work

The work presents theoretical information on the existing Grid systems of the molecular science. The different types of ontologies describe metadata chemical structures in these models, effective ways to identify resources in the Semantic Grid, namely the process of finding resources that meet the needs of users.

The analysis of existing semantic Grid for molecular studies, including the purpose of each and list the tasks they perform with the current requirements of Semantic Grid. These requirements include the interoperability, scalability, decentralization and dynamism.

The achieved results

Results of investigations are to require the Semantic Grid in Computational Chemistry, consistent with the molecular studies in Ukraine and conformity to international standards and existing models, such as: Molecular GRID (World Wide Molecular Matrix (WWMM)), and a draft project CombeChem Collaboratory for Multi-scale Chemical Science (CMCS).

Scientific novelty

The innovation of the executed work consists in analyzing and setting for the basic requirements for application of semantic grid for molecular studies. Now the world is very widely used and developed technology of Semantic Grid in many areas of science, but unfortunately does not have the rapid application of molecular science. Therefore, this paper can serve as a starting point in researching the development of semantic grid for molecular research in Ukraine.

The practical value

The practical value of the work is to obtain systematic, theoretical knowledge base to build infrastructure for the semantic grid for molecular studies. The results may be used as recommendations and proposals to build Ukrainian Grid infrastructure in molecular sciences, as well as guidelines for the use of existing ones.

Conclusions and recommendations

This paper describes the main requirements were imposed in constructing the infrastructure of Semantic Grid. Status of molecular science in the world and shows the problems that arise during molecular studies thus show the relevance and appropriateness of using this infrastructure in the molecular sciences. The paper presented a detailed description of models representing semantic grid in the world, namely: Molecular GRID (World Wide Molecular Matrix (WWMM)), and a draft project CombeChem , Collaboratory for Multi-scale Chemical Science (CMCS).

Based on detailed analysis of these systems revealed certain shortcomings, namely, all systems are dedicated to that search and decide on storage, in neither of them realized sourcing agent. In terms of interoperability, data models covering a very narrow application area, such as a description of molecules, but the project CombeChem through the use of very detailed ontology is able to expand this area.

This work cannot be considered whole section for the development and application of semantic grid for molecular studies because each individual group of scientists engaged in a certain range of research and analyzed the model is the only example of Semantic Grid technologies for a range of tasks and problems arising during molecular studies.

Also in the analysis of research in Ukraine and found that a large number of scientific institutions using molecular researches. Therefore, proposed ways of using Semantic Grid in the national grid - infrastructure. There are two ways to do this: create your own model or use an existing, for each variant prompting some recommendations.

Work on 112 sheets contains 22 illustrations. By preparation of work the literature from 27 different sources was used.

The list of keywords:

Semantic Grid, Semantic Web, Ontology, molecular dynamics, metadata, CML, XML-schema, Protein Data Bank, World Wide Molecular Matrix (WWMM)), CombeChem, Collaboratory for Multi-scale Chemical Science (CMCS), Dublin Core, InChl, molecular structure, folding proteins.