



gridPIR: A central resource for genomic and proteomic information



Baris E. Suzek, Hongzhan Huang, Hsing-Kuo Hua, Peter McGarvey, Cathy H. Wu
Protein Information Resource, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, DC, 20007

Protein Information Resource (PIR)

PIR Activities in caBIG™

- Integrative Cancer Research Workspace
 - Developer
 - Grid-enablement of PIR
 - Adopter
 - SEED Genome Annotation Tool (GeneAnnot)
 - GeneConnect Genomic Identifier Mapping Service
- VCEI Solution for CDR Workspace
 - Development of Organism Identification CDR
- Vocabularies and Common Data Elements
 - Participant

<http://cagrid-portal.nci.nih.gov/>

<http://www.uniprot.org>

<http://pir.georgetown.edu>

<http://pir.georgetown.edu/proteomics>

Protein Science Team (PI)

Executive Team Members

D. Bruce G. Clark, Director, Senior Advisor, Adjunct Professor

D. Zhongping Chen, Associate Director, Team Lead, Research Associate Professor

D. Robert H. Alexander, Senior Protein Scientist, Research Associate Professor

Staff Members

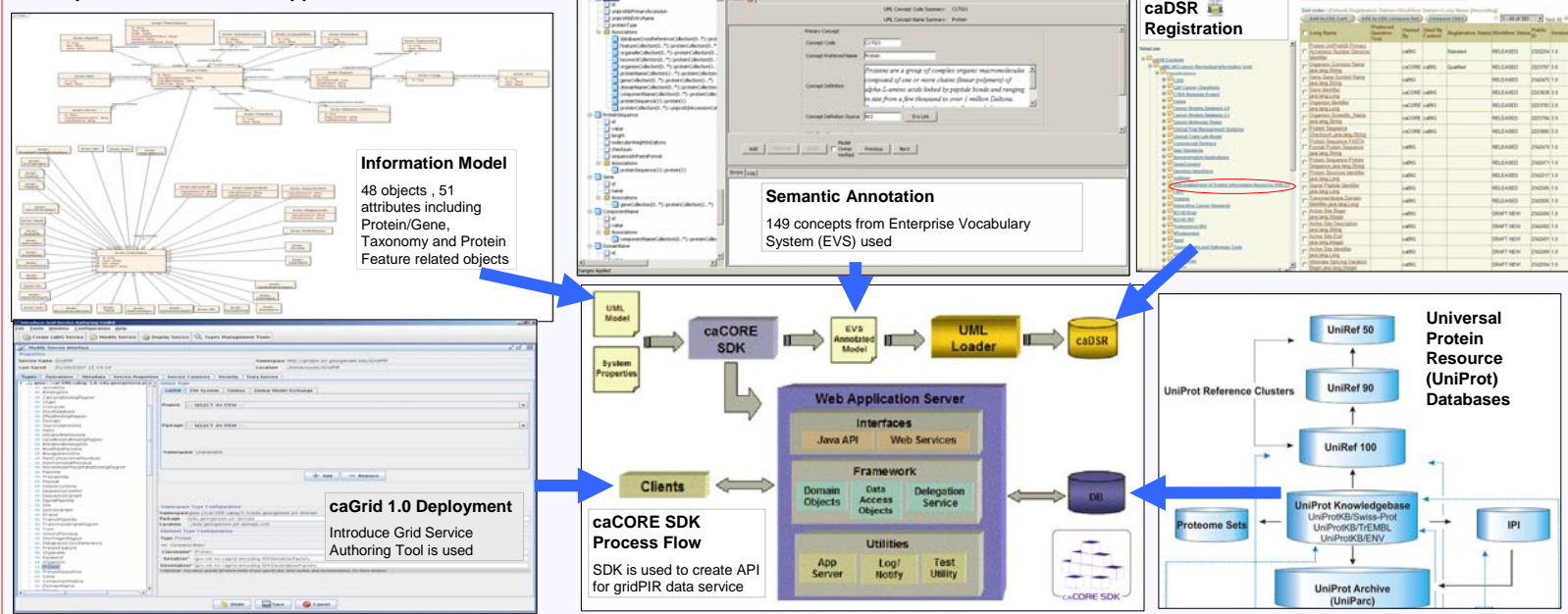
D. Hua Ma, Associate Scientist, Research Associate Professor

D. C. Wang, Senior Protein Scientist, Research Associate Professor

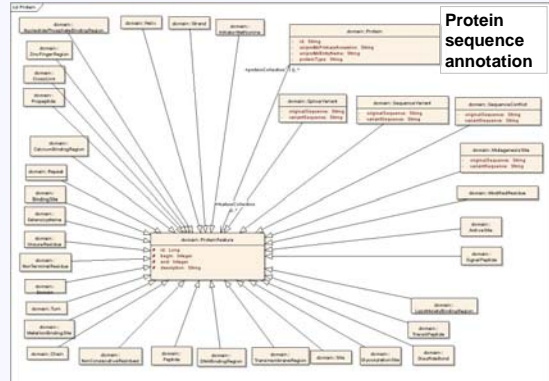
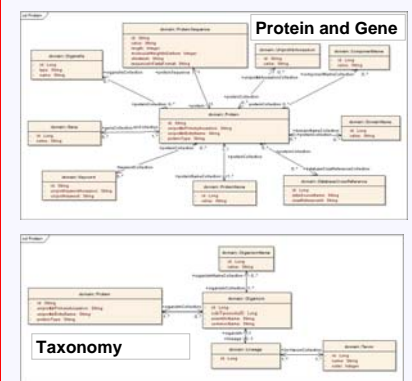
D. Craig A. Kelly, Senior Protein Scientist, Research Associate Professor

Harold P. Pierce, PhD Student

Development - Model Driven Approach



Information Model



Discovering gridPIR on caGrid

Step 1: Find the EVS concept for "Protein" using EVS API

```

EVSDiscoveryClient client = new EVSDiscoveryClient("http://cagrid-
service.nci.nih.gov:8080/wsf/services/cagrid/EVSDiscovery");
EVSDiscoveryClient service = client.getService("EVSDiscovery");
EVSDiscoveryRequest request = new EVSDiscoveryRequest("Protein");
EVSDiscoveryResponse response = client.getService("EVSDiscovery").
getEVSDiscoveryRequest(request);
EVSDiscoveryConceptList concepts = response.getEVSDiscoveryConceptList();
EVSDiscoveryConceptList conceptList = client.getService("EVSDiscovery").
getEVSDiscoveryConceptList(concepts);
  
```

Step 2: Discover data services using the "Protein" concept and discovery API

```

DiscoveryClient client = new DiscoveryClient("http://cagrid-
service.nci.nih.gov:8080/wsf/services/cagrid/Discovery");
DiscoveryClient service = client.getService("Discovery");
DiscoveryRequest request = new DiscoveryRequest("Protein");
DiscoveryResponse response = client.getService("Discovery").
getDiscoveryRequest(request);
DiscoveryResultList results = response.getDiscoveryResultList();
for (int i = 0; i < results.getDiscoveryResultList().length; i++) {
    DiscoveryResult result = results.getDiscoveryResultList().get(i);
    String serviceURL = result.getServiceURL();
  }
  
```

<http://141.161.54.149:8080/wsf/services/cagrid/GridPIR>

<http://128.252.178.222:9092/wsf/services/cagrid/GeneConnect>

<http://cagrid-portal.nci.nih.gov:8080/wsf/services/cagrid/CABIOSvc>

Querying gridPIR data service (1)

Retrieve the proteins for gene "BRCA2" (Breast Cancer Gene 2)

```

<CQLQuery uri="http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/CQLQuery">
  <TargetName> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Target>
    <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  </Target>
</CQLQuery>
  
```

Send the CQL query to gridPIR data service

Retrieve matching Protein objects from service

BRCA2_HUMAN
BRCA2_RAT
BRCA2_MOUSE

From multiple organisms

Querying gridPIR data service (2)

Retrieve human proteins for gene "BRCA2" (Breast Cancer Gene 2)

```

<CQLQuery uri="http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/CQLQuery">
  <TargetName> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Target>
    <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  </Target>
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Target>
    <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  </Target>
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Organism
  <Target>
    <Association> http://cagrid.nci.nih.gov:8080/wsf/services/cagrid/Protein
  </Target>
</CQLQuery>
  
```

Send the CQL query to gridPIR data service

Retrieve matching Protein objects from service

BRCA2_HUMAN

Only human proteins

Interoperability with other data services

