WebLab: a data-centric, knowledge-sharing bioinformatic platform

Wang Jun 王珺

Center for Bioinformatics

Peking University

wangjun@mail.cbi.pku.edu.cn

Contents

1 Introduction

2 Data Management

3 Analysis Service

4 Team Work

5 Others

Contents

1 Introduction

2 Data Management

3 Analysis Service

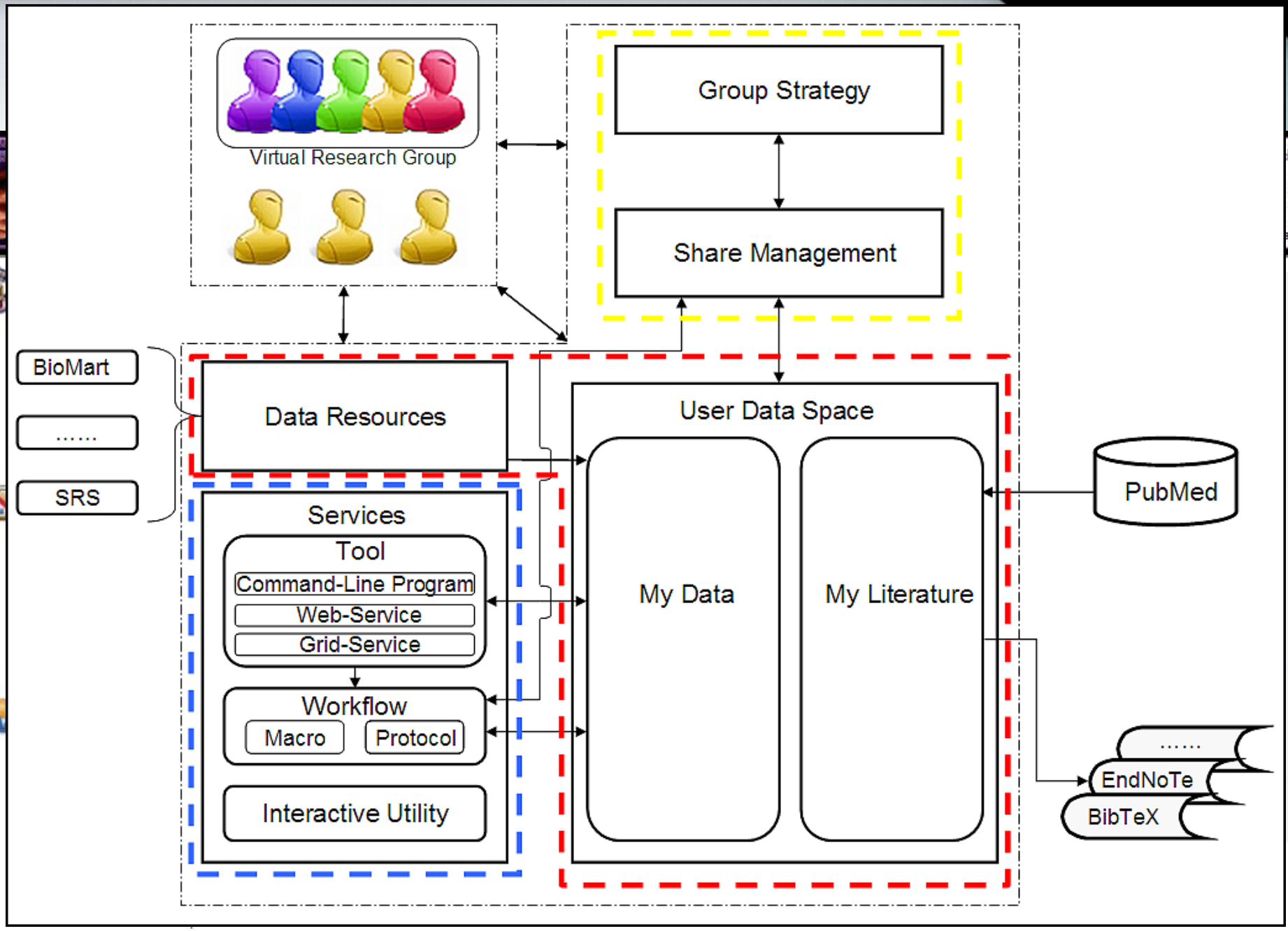
4 Team Work

5 Others

Introduction

Why great demands are proposed for an integrative system?

1. Large biological data are produced
2. Thousands of literatures are pouring
3. Plentiful complex bioinformatic analysis tools
4. Collaboration of researchers from various fields



Copyright© 2004-2009, Center for BioInformatics All Rights Reserved.

Contents

1 Introduction

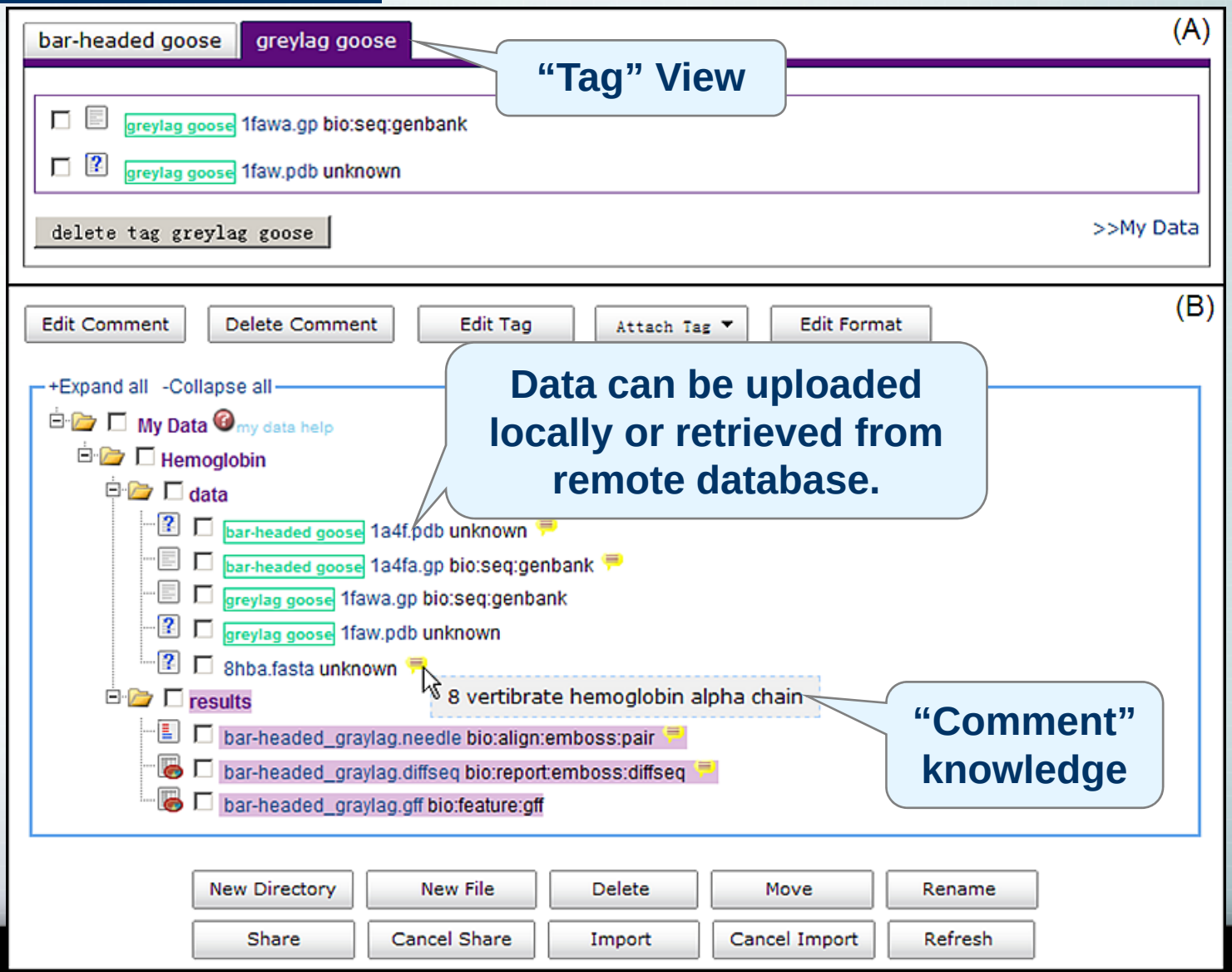
2 Data Management

3 Analysis Service

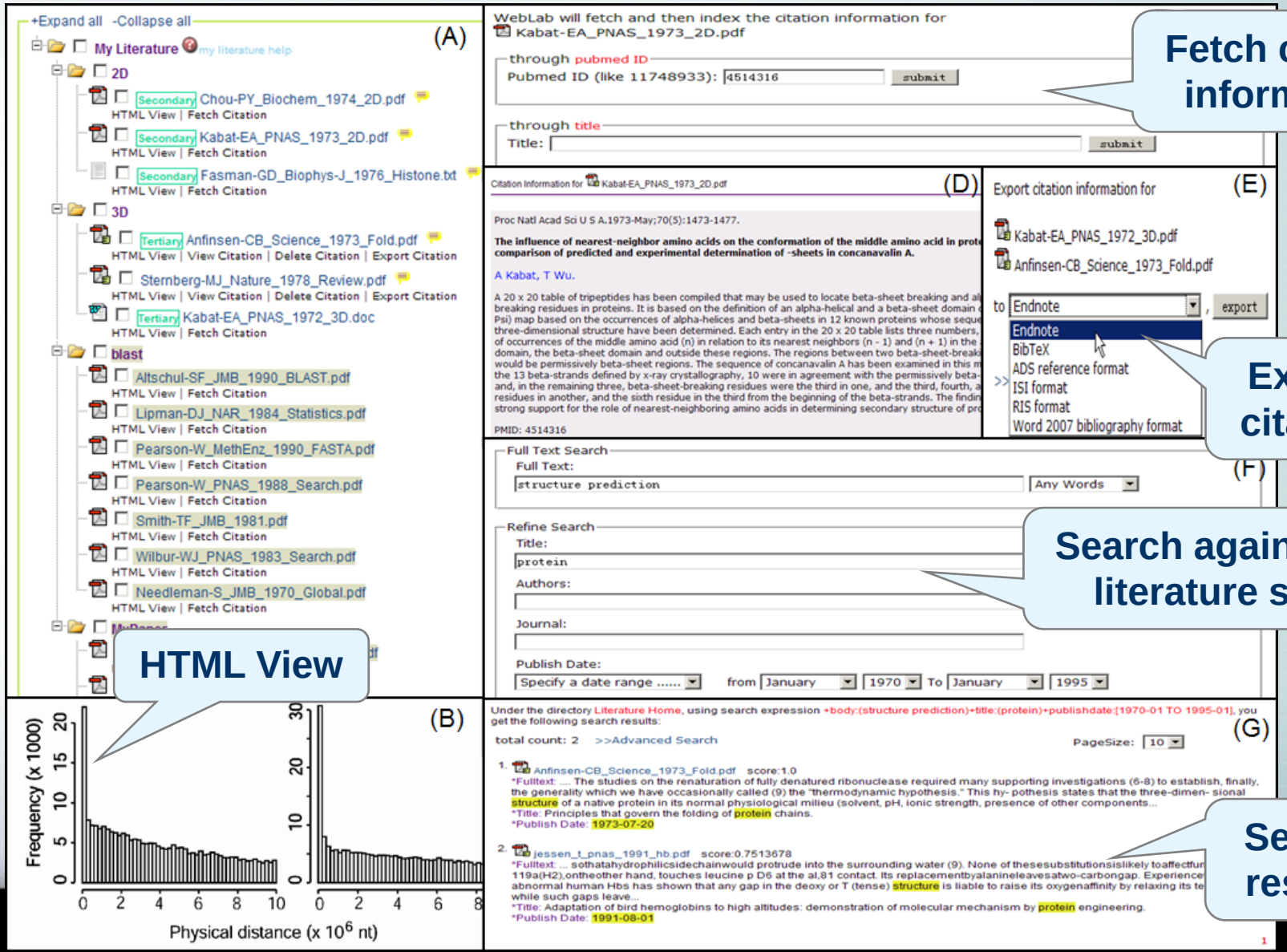
4 Team Work

5 Others

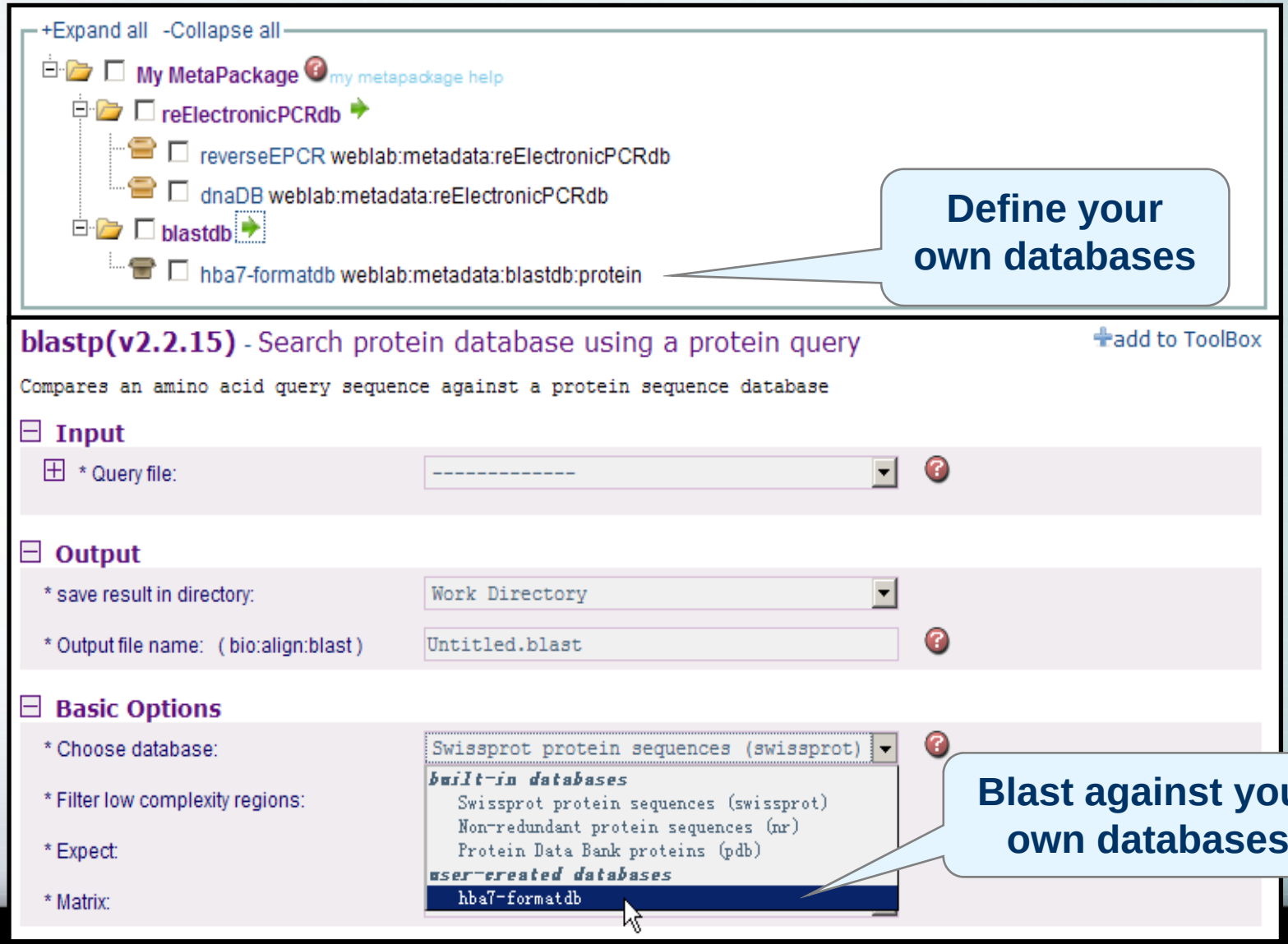
Data Management: “My Data”



Data Management: “My Literature”



Data Management: “My MetaPackage”



Contents

1 Introduction

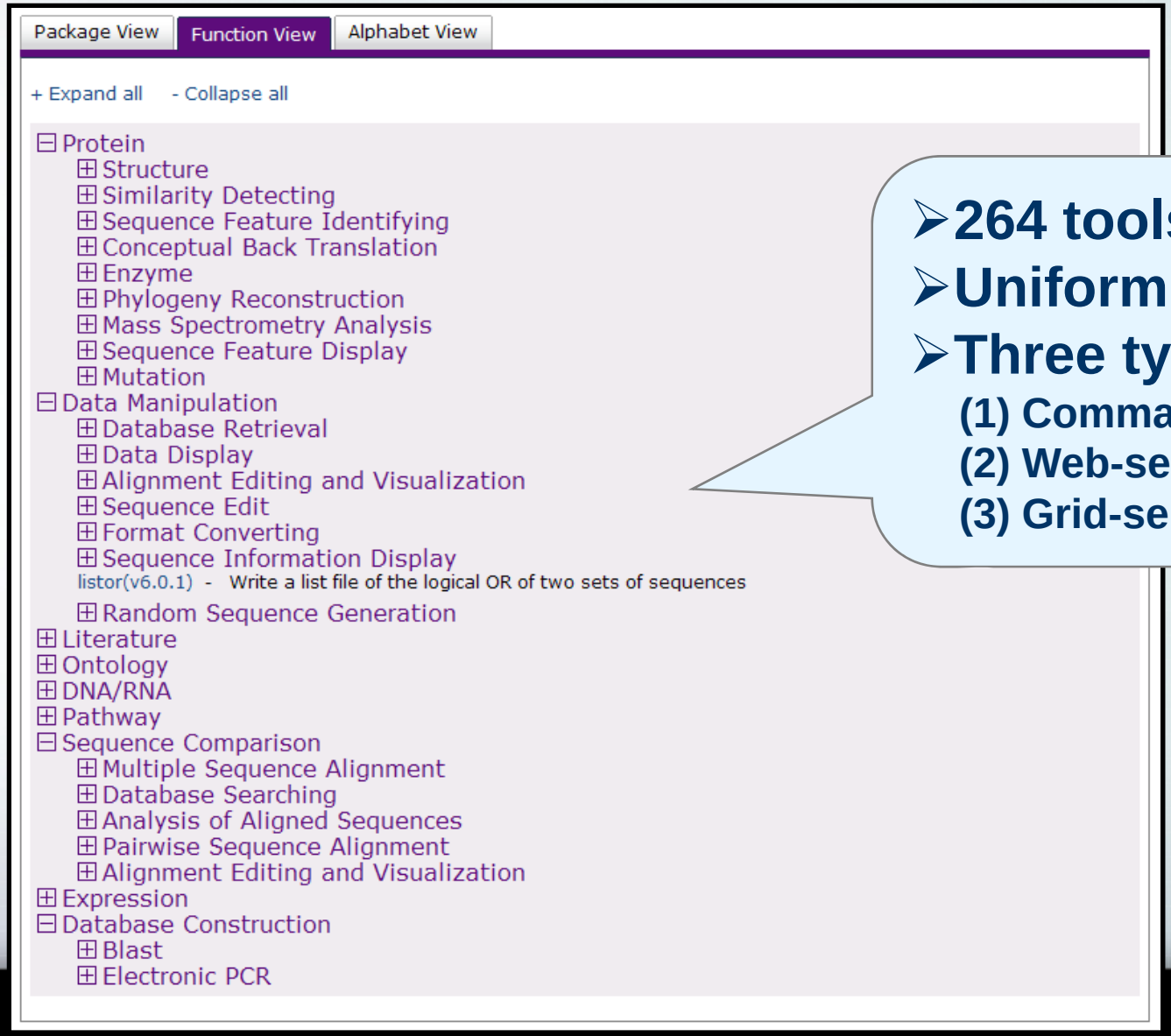
2 Data Management

3 Analysis Service

4 Team Work

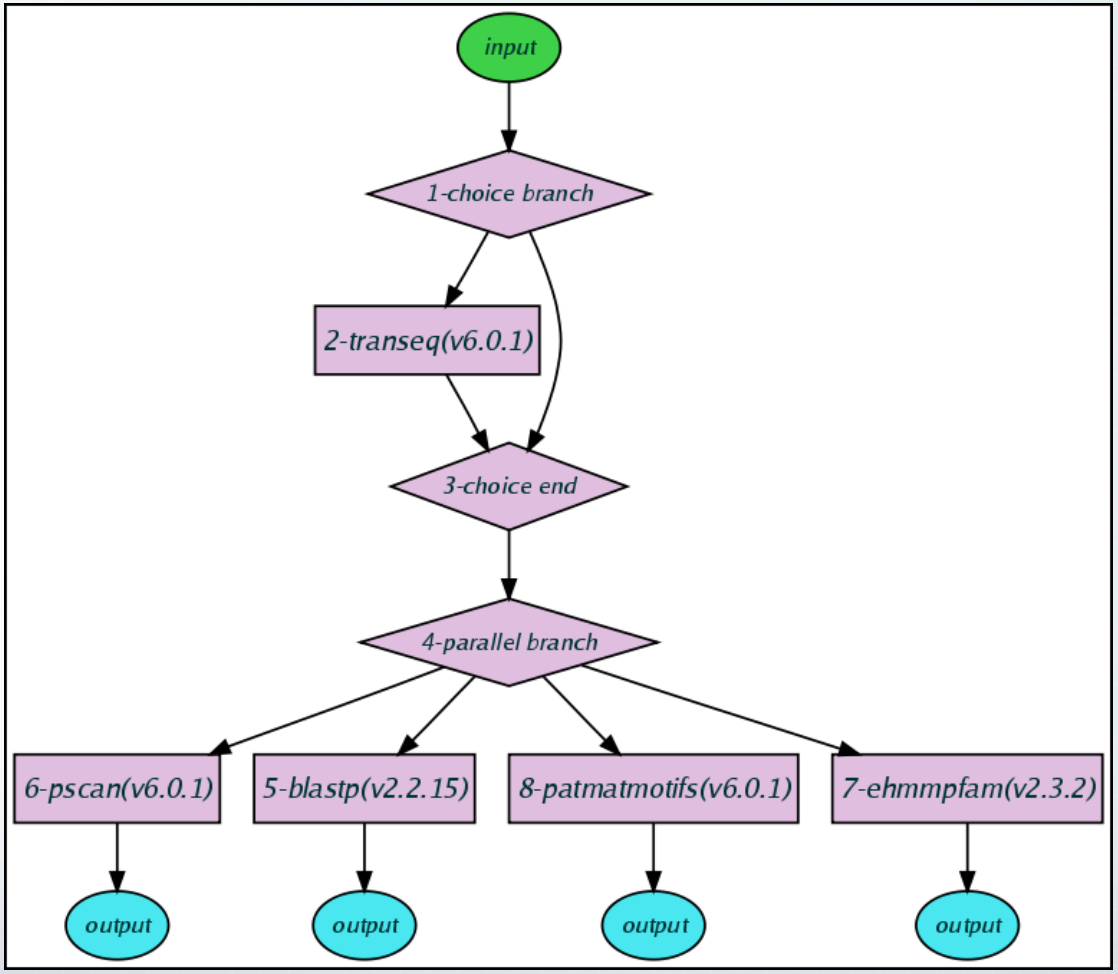
5 Others

Analysis Service: “Program”



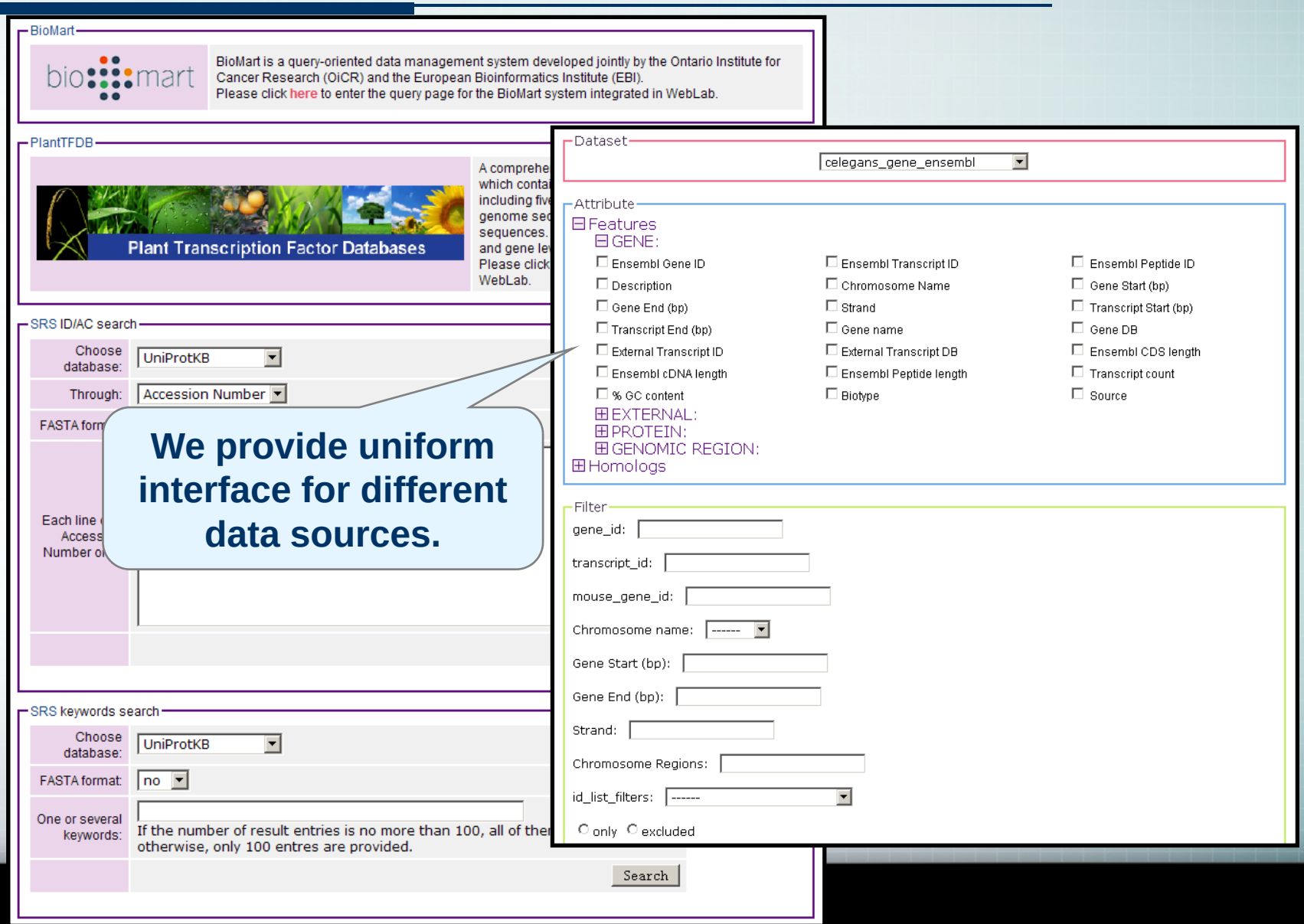
- 264 tools in total
- Uniform interface
- Three types of tools:
  - (1) Command-line programs
  - (2) Web-services
  - (3) Grid-services

Analysis Service: “Workflow”



- **Workflow**
  Complex analysis tasks can be run automatically
- **Two types:**
  Macro
  Protocol
- **Customized**

Analysis Service: “Resource”



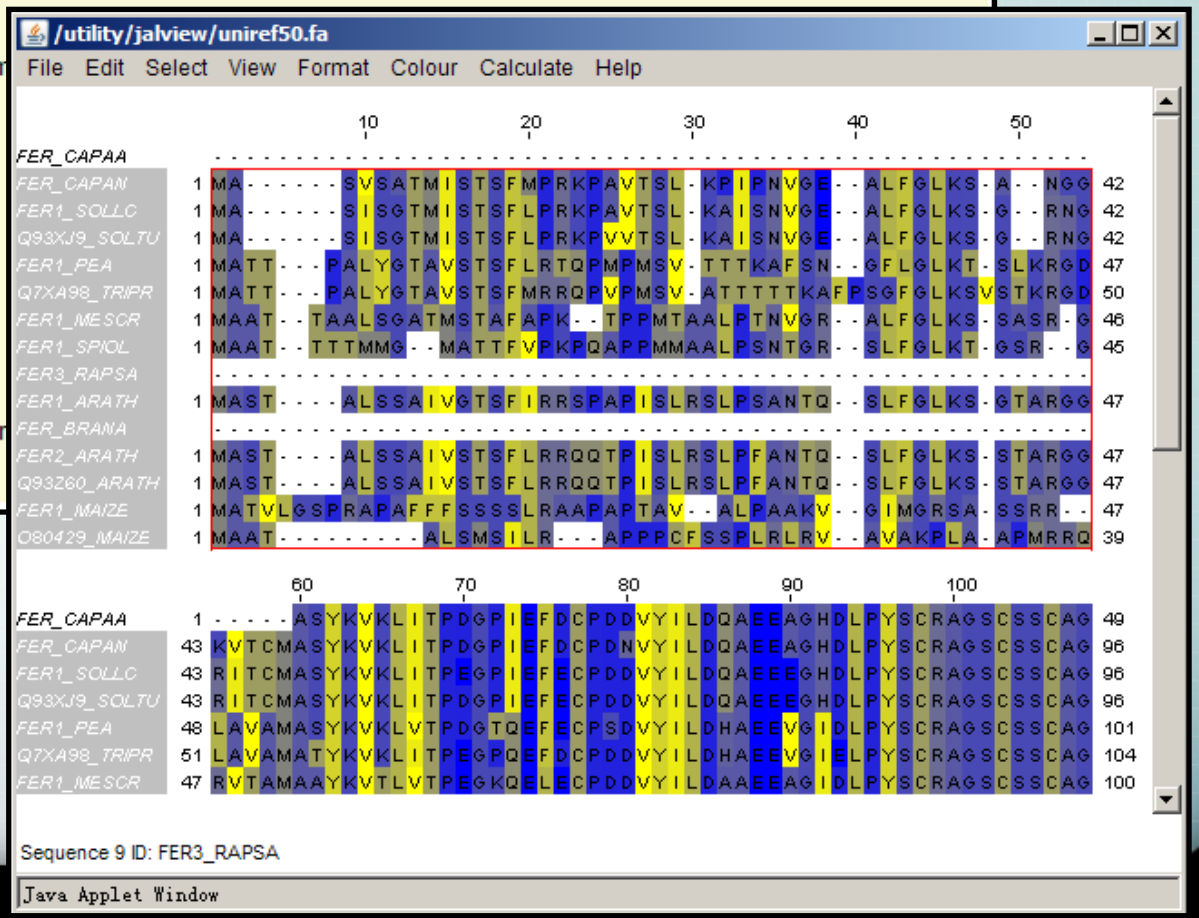
We provide uniform interface for different data sources.

Analysis Service: “Utility”

Utility includes some mini popular bioinformatics analysis tools written in javascript or applet. If you know some other applications which are pretty useful, please give us some advice at your free time. utility help

+ Expand all - Collapse all

- ⊞ Bioverse
- ⊟ Dotlet
  - Dotlet(applet) Dotlet is a program for com
- ⊞ GlycanBuilder
- ⊞ Vimida
- ⊟ Sequence Manipulation Suite
  - ⊞ Sequence Analysis
  - ⊞ Sequence Figures
  - ⊞ Format Conversion
  - ⊞ Miscellaneous
  - ⊞ Random Sequences
- ⊟ jalview
  - jalview(applet) Jalview is a multiple alignr
- ⊞ WebMol



Contents

1 Introduction

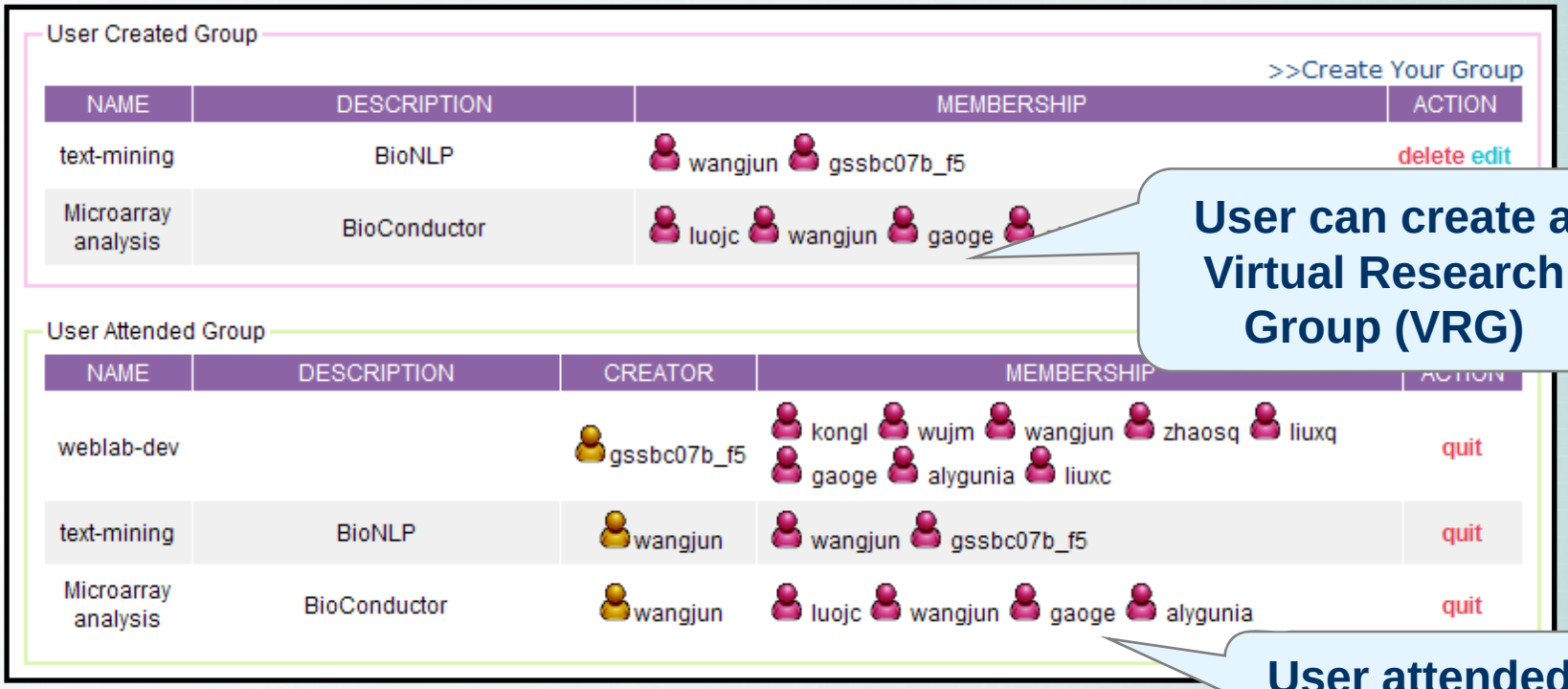
2 Data Management

3 Analysis Service

4 Team Work

5 Others

Team Work: “Group Strategy”



User Created Group

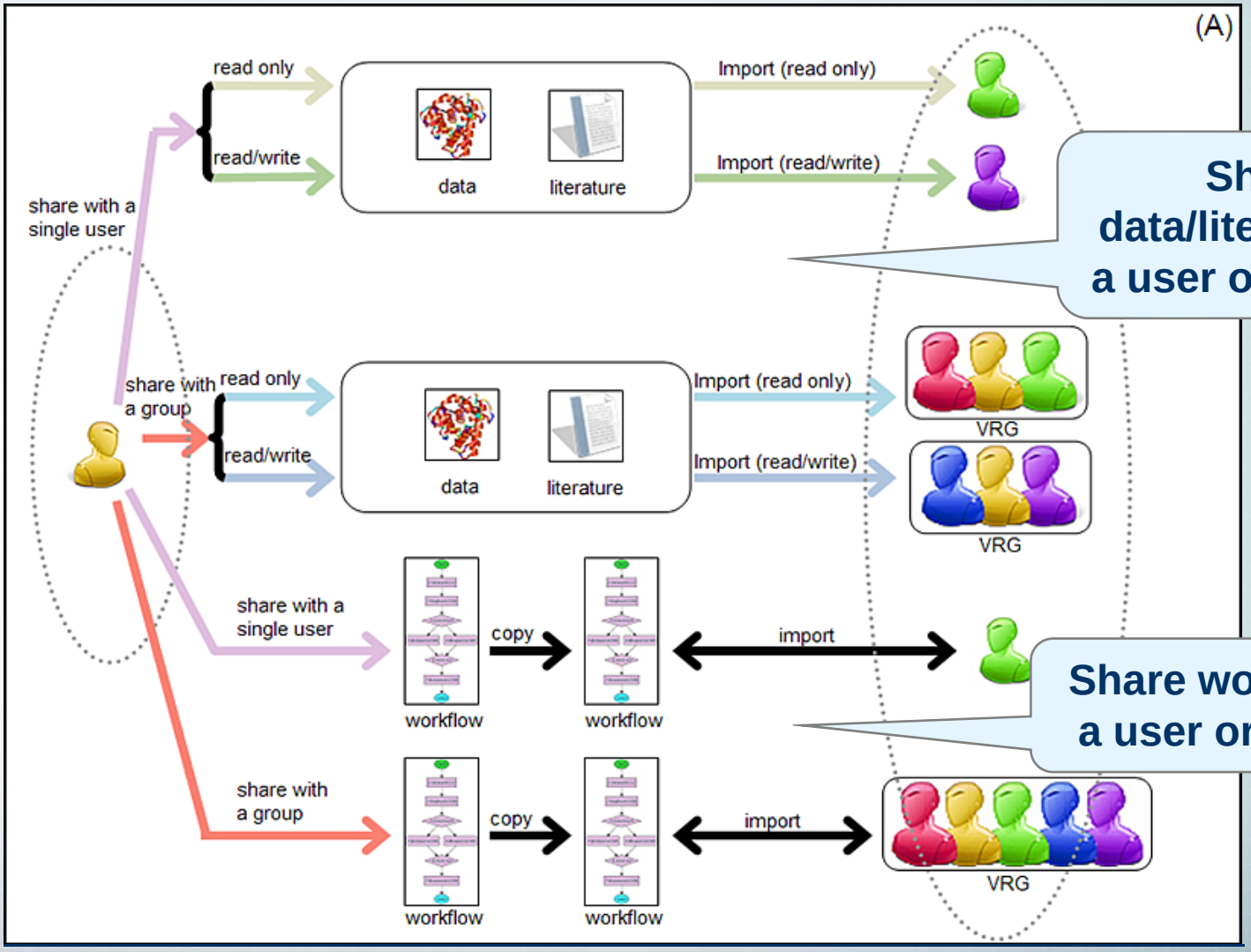
>>Create Your Group

NAME	DESCRIPTION	MEMBERSHIP	ACTION
text-mining	BioNLP	wangjun gssbc07b_f5	delete edit
Microarray analysis	BioConductor	luojc wangjun gaoge	

User Attended Group

NAME	DESCRIPTION	CREATOR	MEMBERSHIP	ACTION
weblab-dev		gssbc07b_f5	kongl wujm wangjun zhaosq liuxq gaoge alygunia liuxc	quit
text-mining	BioNLP	wangjun	wangjun gssbc07b_f5	quit
Microarray analysis	BioConductor	wangjun	luojc wangjun gaoge alygunia	quit

Team Work: “Sharing Mechanism”



Share data/literature to a user or a group

Share workflow to a user or a group

Contents

1 Introduction

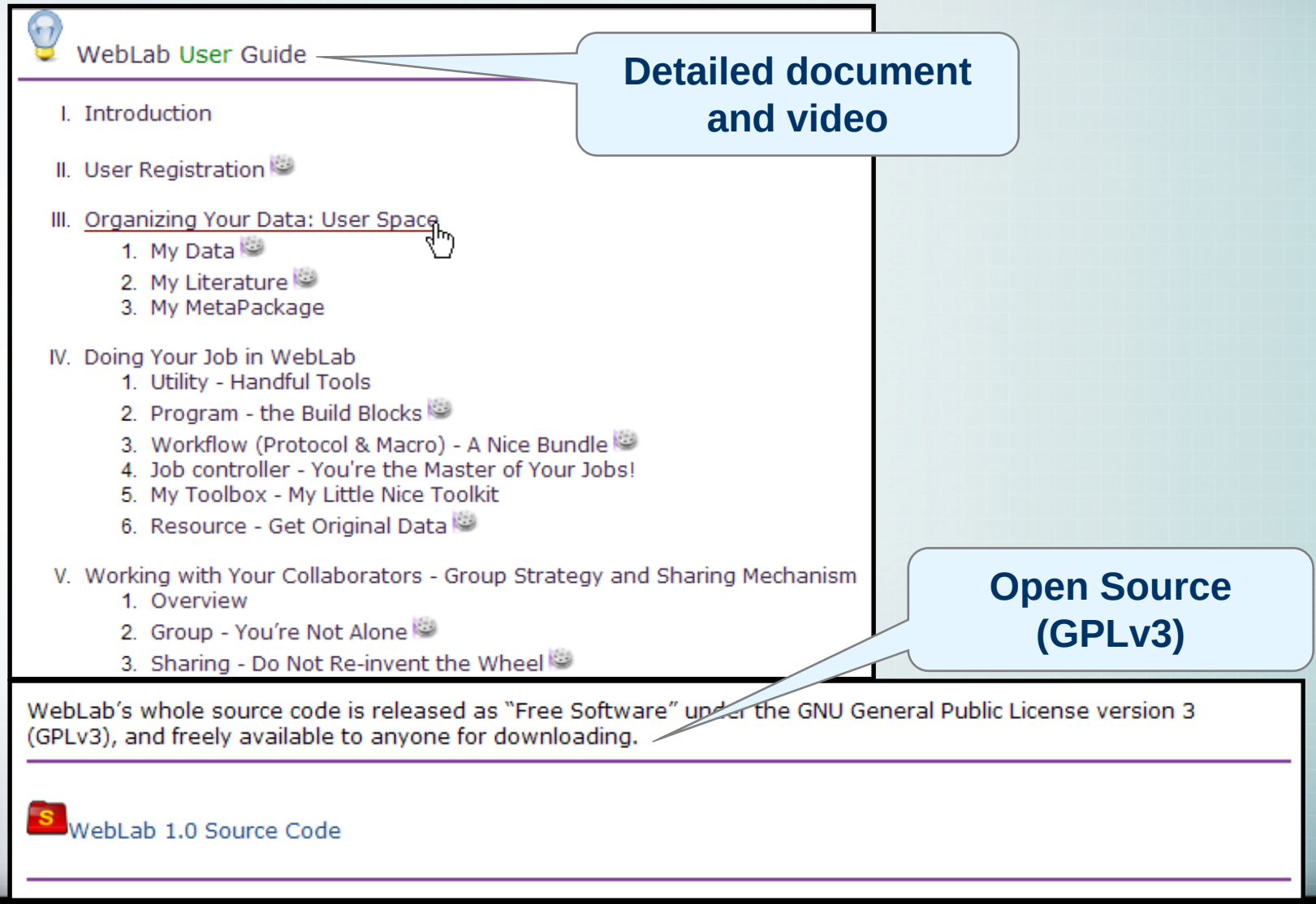
2 Data Management

3 Analysis Service

4 Team Work

5 Others

Others: “Document & License”



Others: “EMBOSS Link”

http://emboss.sourceforge.net

Web interfaces

Most web interfaces to EMBOSS provide a form for setting options and running applications. Difficulties can arise because of 'dependencies' in the application input. These include where the value of one parameter depends on the value or property of another, for example sequence length, or where the available options depend on the input, for example whether a sequence is DNA or Protein. Some web interfaces handle dependencies better than others.

To use a web interface, all you need are a web broswer and internet access. Some of these interfaces run 'anonymously' so any user can run the applications. In most cases however you will be expected to register with the site before you can use their services.

The interfaces we have seen so far are:

WebLab

WebLab collect provide easy web access to the most popular bioinformatics programs. It is task-oriented instead of tool-oriented. It provides pre-defined protocols to guide biologists to work step by step through the entire analysis procedure.

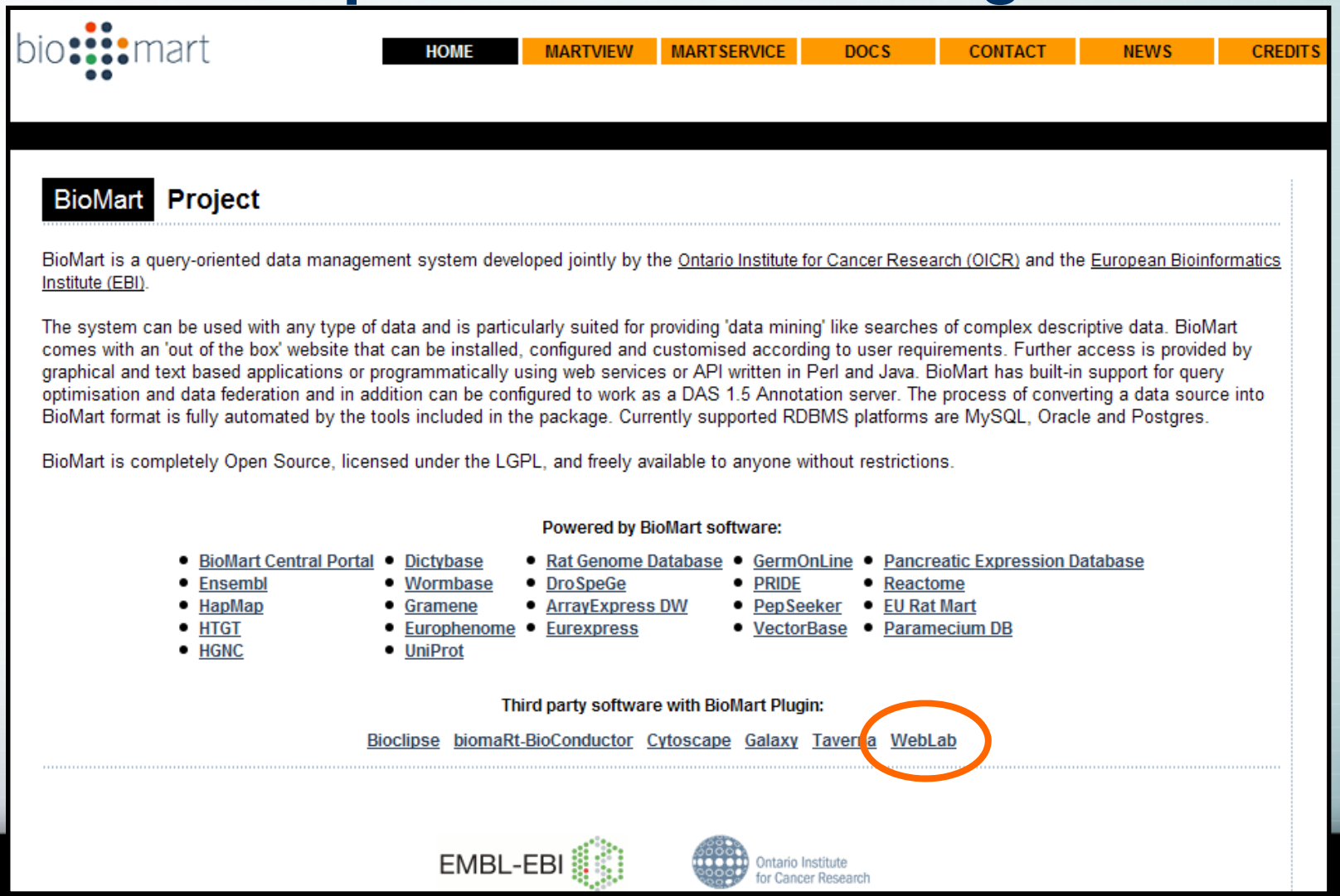
Pise

PISE (french pronunciation like the english word 'peas') was developed by Catherine Letondal at the Institute Pasteur. Using an XML definition for each application, PISE generates a Web interface and other common interfaces automatically.

For EMBOSS, the ACD files were converted with scripts to generate the XML definitions. Where ACD files caused problems, we simply fixed the ACD files.

EMBOSS Explorer.

An extremely popular web interface to EMBOSS from Luke McCarthy at Plant Biotechnology Institute, Canada

Others: “BioMart Link”

http://www.biomart.org