

caBIG dotNET and the  
xl-caBIG Smart Client

bringing caBIG cancer data to the .NET Developer and  
Microsoft Office User Communities

*<http://xl-cabig-client.sourceforge.net/>*

*Robert Macura*

*Tom Macura*

# eScience Workshop October 2005

---

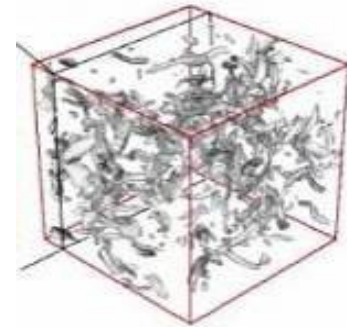


Microsoft Faculty Summit 2006

# Science Paradigms by Jim Gray

- Thousand years ago:  
science was **empirical**  
describing natural phenomena
- Last few hundred years:  
**theoretical** branch  
using models, generalizations
- Last few decades:  
a **computational** branch  
simulating complex phenomena
- Today:  
**data exploration** (eScience)  
unify theory, experiment, and simulation  
using data management and statistics
  - Data captured by instruments  
Or generated by simulator
  - Processed by software
  - Scientist analyzes database / files

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K \frac{c^2}{a^2}$$



# Sharing Cancer Research Data

---

- **Cancer Research is an empirical, hypothesis driven process:**
  - based on past knowledge and a novel idea a hypothesis is formed
  - an experiment is constructed in an attempt to (dis)prove the hypothesis
  - experiment is performed and resulting data is collected
  - results are analyzed, and conclusions from the analysis are presented at meetings and/or published in peer-reviewed Journals

# Sharing Cancer Research Data

---

- **Value in sharing raw research data -- not just in their interpretations. Research data can be reused**
  - To provide evidence towards other hypothesis
- **Research data, especially clinical data, is very expensive to collect**
- **Make data accessible to others**
  - Discoverable
  - Syntactically Interoperable (others have to be able to read the data)
  - Semantically Interoperable (others have to understand it).



**caBIG** *cancer Biomedical  
Informatics Grid*

NATIONAL<sup>®</sup>  
CANCER  
INSTITUTE



- **caBIG** (*cancer architecture Biomedical Informatics Grid*) is a 60m (3yrs) National Cancer Institute informatics project to connect data, research tools, scientists, and organizations.
- **caBIG** 's goal is to speed the delivery of innovative approaches for the prevention and treatment of cancer.
- By combining shared vocabulary, data elements, and data models **caBIG** is expected to become a common-standard informatics platform used for all cancer research supported by NIH grants -- the *WWW of Cancer Research*

# caGRID Overview

---

- caGRID is *caBIG*'s enabling infrastructure forming the “grid” in caBIG
- Features of caGRID
  - Provide syntactic and semantic interoperability, based on shared information models
  - Security framework providing authentication and permission control
- On the caGRID there are data-services (providing data), analytic-services (providing algorithm implementations), and core *caBIG* services (e.g. indexing service listing available data- and analytic-services)

# http://cabig.nci.nih.gov

The screenshot shows a Windows Internet Explorer browser window displaying the caBIG website. The address bar shows the URL <https://cabig.nci.nih.gov/>. The page header includes the National Cancer Institute logo and the text "National Cancer Institute" and "U.S. National Institutes of Health | www.cancer.gov". The main content area features a "Welcome to the caBIG™ Web site" message and a news item titled "NCI Receives the Computerworld 21st Century Achievement Award in Science for caBIG™". A sidebar on the left contains a navigation menu with items like "Home", "About caBIG", "Events", "Press Room", "Speakers Bureau", "Library", "caBIG Toolkit", "Inventory of Tools", "Software Downloads", "Compatibility Guidelines", "Concurrent Version System (CVS)", "caGrid 0.5", "caBIG Data Elements (CDE Browser)", "caBIG Data Standards", "Training Portal", "Participants Workspace", "Login", "Documents Under Review", and "Architecture". A "caBIG CALENDAR" for July 2006 is visible on the right, with the 17th, 18th, 19th, 20th, and 21st highlighted. Below the calendar is a "caBIG NEWS" section with entries for 07-14-2006 and 07-06-2006. The browser's status bar at the bottom shows "Internet" and "100%" zoom.

caBIG - Welcome to the caBIG™ Web site - Windows Internet Explorer

https://cabig.nci.nih.gov/

File Edit View Favorites Tools Help

caBIG - Welcome to the caBIG™ Web site

NATIONAL CANCER INSTITUTE National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

caBIG™ cancer Biomedical Informatics Grid™

caBIG™ cancer Biomedical Informatics Grid™

Home

About caBIG

Events

Press Room

Speakers Bureau

Library

caBIG Toolkit

Inventory of Tools

Software Downloads

Compatibility Guidelines

Concurrent Version System (CVS)

caGrid 0.5

caBIG Data Elements (CDE Browser)

caBIG Data Standards

Training Portal

Participants Workspace

Login

Documents Under Review

Architecture

home

## Welcome to the caBIG™ Web site

[NCI Receives the Computerworld 21st Century Achievement Award in Science for caBIG™](#)

**About caBIG™** - The cancer Biomedical Informatics Grid, or caBIG™, is a voluntary network or grid connecting individuals and institutions to enable the sharing of data and tools, creating a World Wide Web of cancer research. The goal is to speed the delivery of innovative approaches for the prevention and treatment of cancer. The infrastructure and tools created by caBIG™ also have broad utility outside the cancer community. caBIG™ is being developed under the leadership of the [National Cancer Institute's Center for Bioinformatics](#).

**caBIG™ Participants** - Over 800 people from more than 80 organizations are working collaboratively on over 70 projects in a three-year pilot project. Workspace specific information, materials and

caBIG CALENDAR

July 2006

Su	Mo	Tu	We	Th	Fr	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Select date for more information.

caBIG NEWS

07-14-2006  
[caBIG™ -Related Journal Article Published in "BMC Medical Informatics and Decision Making"](#)

07-14-2006  
[What's BIG This Week - 07/14/06](#)

07-06-2006  
[caBIG™ Town Hall Meeting - July 31st](#)

06-07-2006

Internet 100%

Microsoft Faculty Summit 2006





# for Cancer Research

---

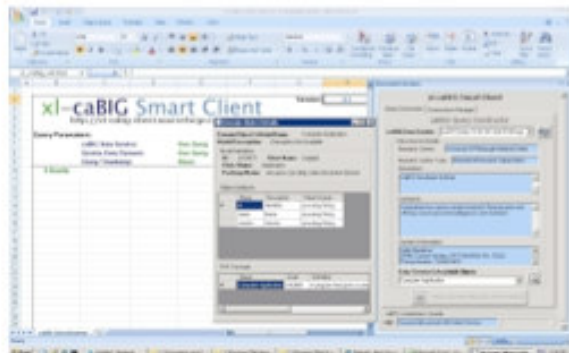
- **caBIG** is a \$60 Million (3yrs) project that will give scientists access to orders of magnitude more related data from researchers around the world.
- What tools do scientists need to browse, query, and analyze this data and make meaningful deductions?
  - Using **Microsoft Excel** and leveraging their familiarity with it's statistical analysis and visualization.
  - Using **Windows Mobile 5.0** to gain on-the-spot hypothesis testing; thus replacing the proverbial "restaurant napkin sketches"
- How could we develop such applications? Using the .NET Framework
  - Visual Studio Tools for Office
  - Compact Framework for Mobile Devices

# caBIG dotNET bridges caGRID and



- 
- caGRID API is written in Java leveraging many Java projects: Globus Toolkit 3.2, OGSA-DAI 5.0, Tomcat, Ant
  - caBIG dotNET is our project bridging the gap between caBIG and .NET
  - It is a C# Client API based on the Hessian Binary Protocol that connects with a corresponding Java Servlet to provide exposition of the caBIG API
  - BSD License; SourceForge Website and CVS Repository <http://xl-cabig-client.sourceforge.net/>

# xl-caBIG Smart Client

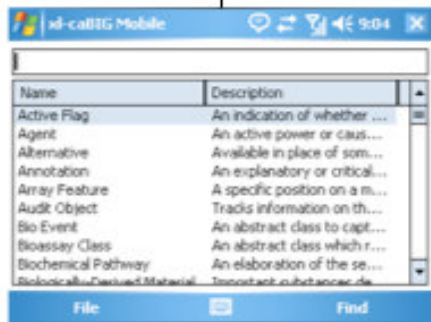


dynamic linking

## caBIG dotNET (C# Client API)



dynamic linking



# xl-caBIG Smart Client MOBILE

# HTML View (JSP/Java)



caucho Hessian Protocol

## caBIG dotNET (Java Servlet)



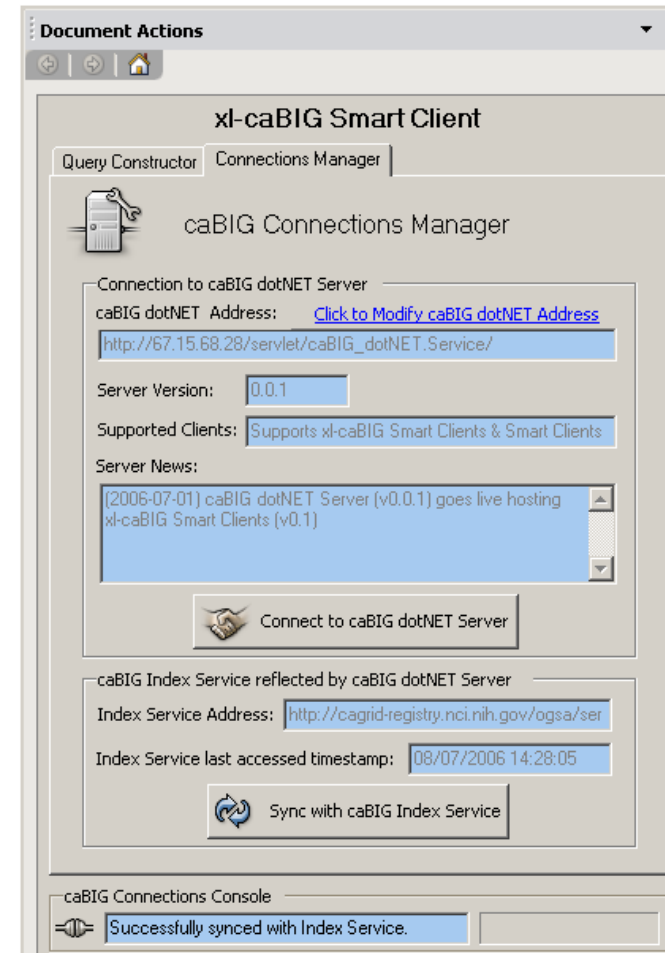
caucho Hessian Protocol



# Console Manager (Java)

# xl-caBIG Smart Client

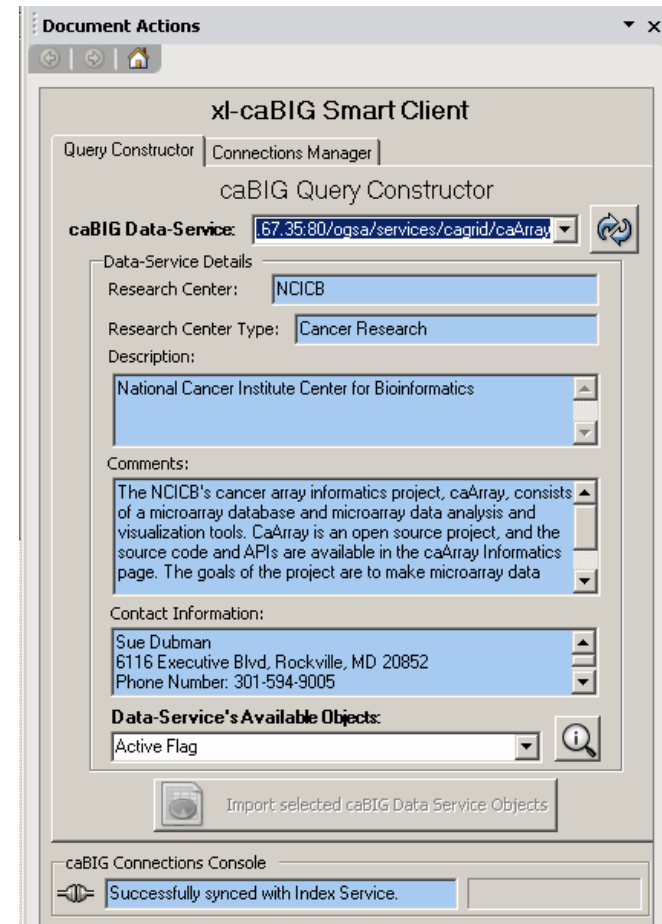
- Live demo



Microsoft Faculty Summit 2006

# xl-caBIG Smart Client

- Live demo



Microsoft Faculty Summit 2006

# xl-caBIG Smart Client

- Live demo

**Domain Object Details**

**Domain Object's Model Name:** Candidate Disease Gene  
**Model Description:** *Description Not Available*

Model Identifiers

**ID:** 2223317    **Short Name:** C19389  
**Class Name:** Target  
**Package Name:** gov.nih.nci.cabio.domain

Object Attributes

Name	Description	Value Domain
id	Identifier	java.lang.Long
type	Type	java.lang.String
name	Name	java.lang.String

EVS Concepts

Name	Code	Definition
Candidate Disease Gene	C19389	A gene proposed to hav

# xl-caBIG Smart Client

Future Directions (July-October)

- The Smart-Client's Document Actions Pane will allow users to graphically compose a query relating multiple Domain Objects.

The image displays three screenshots of the 'Domain Object Details' window, illustrating the graphical composition of a query relating multiple domain objects. Each window shows the details of a specific domain object, including its model name, description, identifiers, attributes, and EVS concepts.

**Window 1: Candidate Disease Gene**

- Domain Object's Model Name: Candidate Disease Gene
- Model Description: Description Not Available
- Model Identifiers: ID: 2223317, Short Name: C19389, Class Name: Target, Package Name: gov.nih.nci.cabio.domain
- Object Attributes:

Name	Description	Value Domain
id	Identifier	java.lang.Long
type	Type	java.lang.String
name	Name	java.lang.String
- EVS Concepts:

Name	Code	Definition
Candidate Disease Gene	C19389	A gene proposed to hav

**Window 2: Clone Gene Relative Location Object**

- Domain Object's Model Name: Clone Gene Relative Location Object
- Model Description: Description Not Available
- Model Identifiers: ID: 2223343, Short Name: C37925.C45377, Class Name: CloneRelativeLocation, Package Name: gov.nih.nci.cabio.domain
- Object Attributes:

Name	Description	Value Domain
id	Identifier	java.lang.Long
type	Type	Gene Relative Location Type
- EVS Concepts:

Name	Code	Definition
Gene Relative Location Object	C45377	The location of a f
Clone	C37925	Clone; a group of s

**Window 3: Gene Relative Location Object**

- Domain Object's Model Name: Gene Relative Location Object
- Model Description: Description Not Available
- Model Identifiers: ID: 2223344, Short Name: C45377, Class Name: GeneRelativeLocation, Package Name: gov.nih.nci.cabio.domain
- Object Attributes:

Name	Description	Value Domain
id	Identifier	java.lang.Long
type	Type	java.lang.String
- EVS Concepts:

Name	Code	Definition
Gene Relative Location Object	C45377	The location of a f

Blue arrows indicate relationships between the 'id' attribute of one object and the 'id' attribute of another, showing how they are linked in a query.

Microsoft Faculty Summit 2006

# xl-caBIG Smart Client

Future Directions (July-October)

---

- This query will be transcribed into **caBIG** Common Query Language (XML) and executed against data-services via **caBIG dotNET**

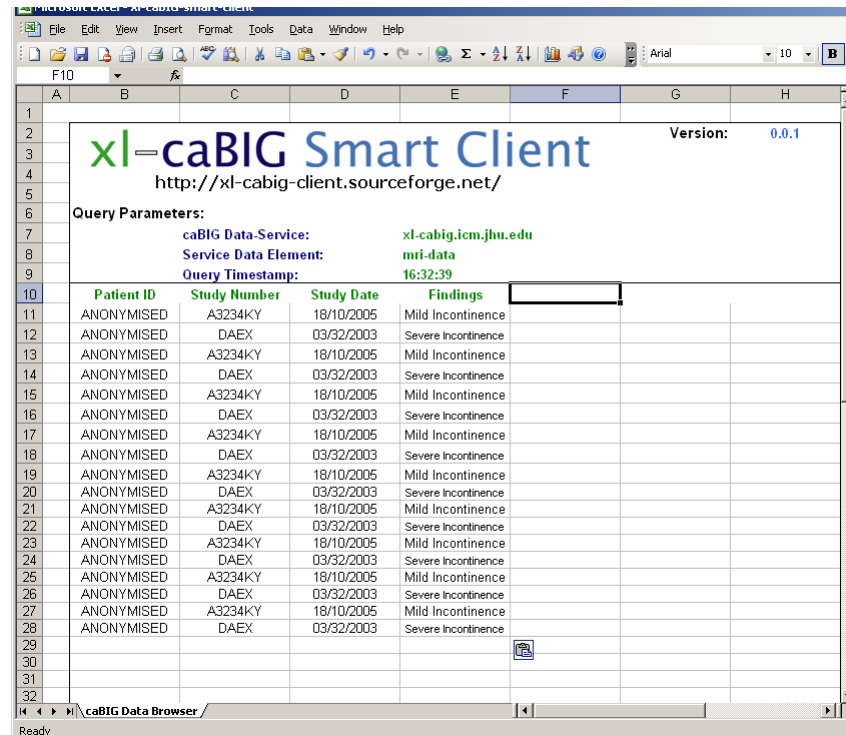
```
<caGridXMLQuery name="caArrayQuery">  
  <criteria name="gov.nih.nci.mageom.domain.Protocol.Protocol">  
    <criteria name="identifier" condition="EQUAL_TO value="P-MEXP-1963"/>  
  </criteria>  
</caGridXMLQuery>
```



# xl=caBIG Smart Client

## Future Directions (July-October)

- Results of the query (XML) will be interpreted by the Smart Client to populate cells in the Excel Workbook.



Microsoft Excel - xl=caBIG Smart Client

File Edit View Insert Format Tools Data Window Help

F10

xl=caBIG Smart Client Version: 0.0.1  
http://xl-cabig-client.sourceforge.net/

Query Parameters:

caBIG Data-Service: xl-cabig.icm.jhu.edu  
Service Data Element: mri-data  
Query Timestamp: 16:32:39

Patient ID	Study Number	Study Date	Findings
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence
ANONYMISED	A3234KY	18/10/2005	Mild Incontinence
ANONYMISED	DAEX	03/32/2003	Severe Incontinence

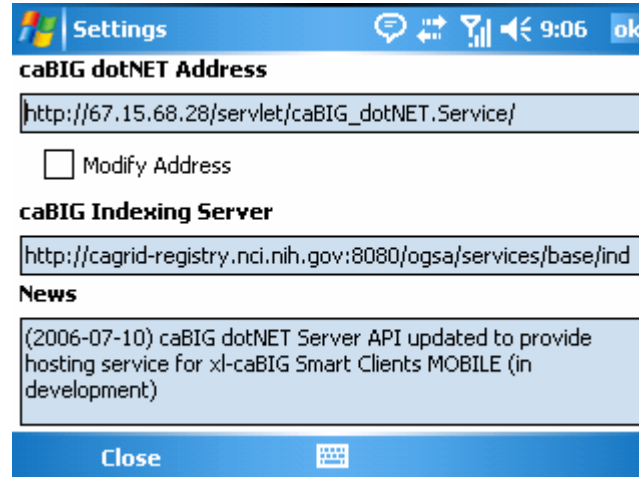
caBIG Data Browser

Ready

Microsoft Faculty Summit 2006

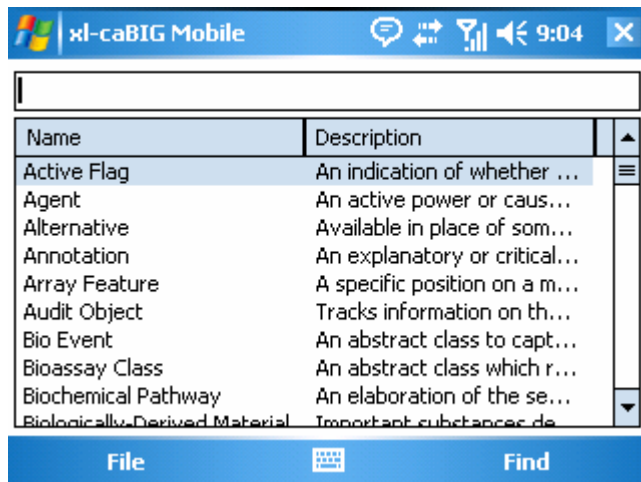
# xl-caBIG Smart Client MOBILE

- xl-caBIG MOBILE uses the Smart Phone's GPRS to connect with the caBIG dotNET service



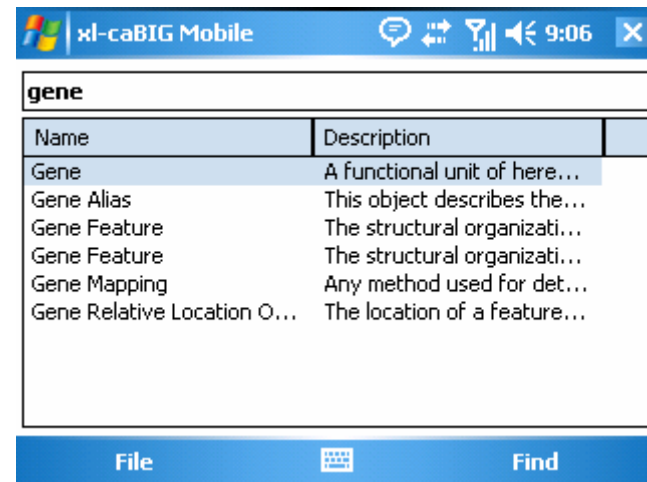
# xl-caBIG Smart Client MOBILE

- Scientists use their **Windows 5.0** Mobile Device to browse and search for cancer concepts of interest



The screenshot shows the xl-caBIG Mobile application interface. The title bar reads "xl-caBIG Mobile" and includes a search icon, a keyboard icon, a signal strength icon, and a time display of 9:04. Below the title bar is a search input field. The main content area displays a table with two columns: "Name" and "Description". The table lists various biological concepts, with "Active Flag" selected. A bottom navigation bar contains "File", a keyboard icon, and "Find".

Name	Description
Active Flag	An indication of whether ...
Agent	An active power or caus...
Alternative	Available in place of som...
Annotation	An explanatory or critical...
Array Feature	A specific position on a m...
Audit Object	Tracks information on th...
Bio Event	An abstract class to capt...
Bioassay Class	An abstract class which r...
Biochemical Pathway	An elaboration of the se...
Biologically-Derived Material	Important substances de...

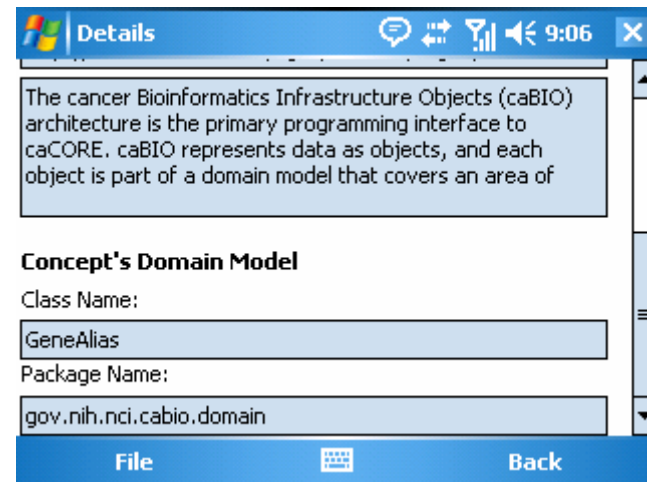
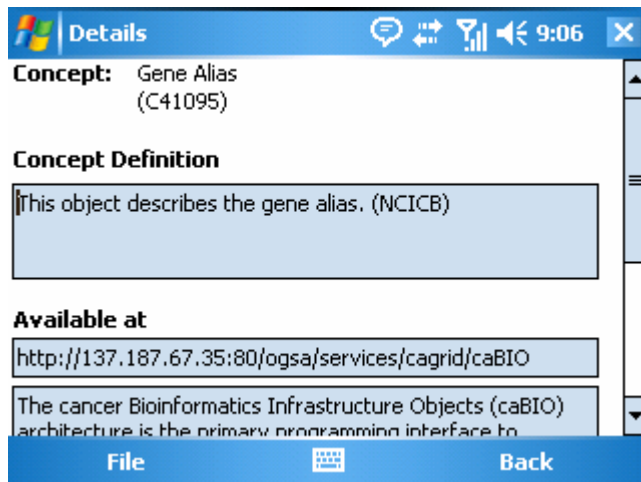


The screenshot shows the xl-caBIG Mobile application interface with the search term "gene" entered in the search input field. The title bar reads "xl-caBIG Mobile" and includes a search icon, a keyboard icon, a signal strength icon, and a time display of 9:06. The main content area displays a table with two columns: "Name" and "Description". The table lists various biological concepts related to genes, with "Gene" selected. A bottom navigation bar contains "File", a keyboard icon, and "Find".

Name	Description
Gene	A functional unit of here...
Gene Alias	This object describes the...
Gene Feature	The structural organizati...
Gene Feature	The structural organizati...
Gene Mapping	Any method used for det...
Gene Relative Location O...	The location of a feature...

# xl-caBIG Smart Client MOBILE

- Once users have found a cancer concept they're interested in, xl-caBIG MOBILE gives further information about the concept and where on *caBIG* this concept type is available



# xl-caBIG Smart Client MOBILE

---

- xl-caBIG MOBILE running on Cingular 8125 Smart Phone



Microsoft Faculty Summit 2006

# xl-caBIG Smart Client MOBILE

- xl-caBIG MOBILE is easy to use and very portable
- What Bill Buxton says about TREK's design might be true but for some tasks handlebars just get in the way



# xl-caBIG Project Team

CO-PIs *Katarzyna Macura MD PhD* and  
*Robert Macura MD PhD*



Lead Programmer: *Tom Macura BS BS*  
(caBIG dotNET, Smart Client)



Programmer: *Wiktor Macura*  
(Smart Client MOBILE)



Sponsored by: *Microsoft Research Smart Clients for eScience 2005*

Project website: <http://xl-cabig-client.sourceforge.net/>

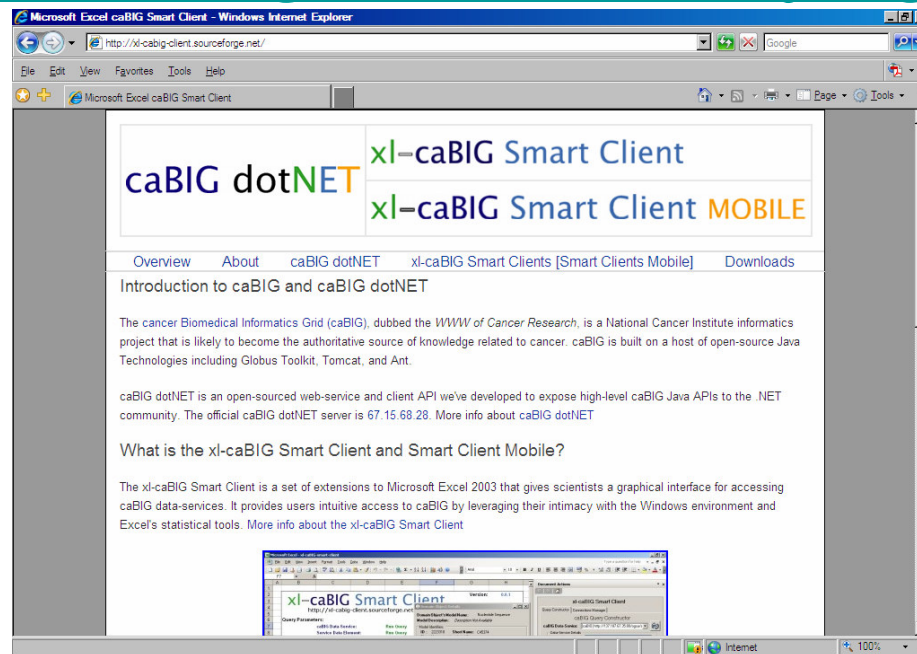


*BSD Open-Source License* 

# Thank You !

- Microsoft Research, Dan Fay, and Simon Mercer

<http://xl-cabig-client.sourceforge.net/>



Microsoft Faculty Summit 2006