

# Building an international, online cancer research community

Stuart Bell

Informatics Initiative

National Cancer Research Institute, UK

# The NCRI Partnership



# NCRI

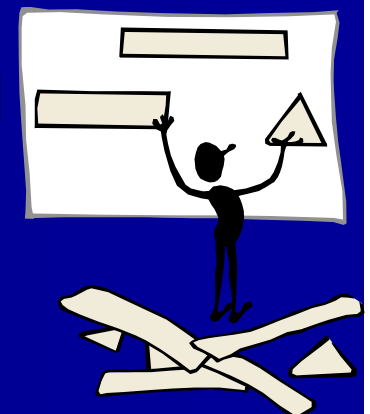
- National Cancer Research Institute
- Partnership of 21 main funders of cancer research in the UK
- Created an 'Informatics Initiative'
- Remit: to make better use of the vast amounts of data generated
  - Reduce duplication of research effort
  - Streamline the research process
  - Form a real, productive and sustained online research community

# What is informatics?

- Process of **collecting, managing and sharing information**
- Information can be used for treatment, planning, management and research, to improve healthcare
- Governments, industry and research charities are developing informatics systems to allow data to be used to its full potential

# The Informatics Challenge

- Think of all the types of data relevant to cancer
  - biology, animal models, genetics, clinical, epidemiology, social and demographic etc.
- Data volume
  - the sheer scale of data
  - distribution
  - indexing
- Data standards
  - not all data contains all elements needed
- **Much research data isn't accessible**



# Cancer Informatics

- General: using technology to make use of information relating to cancer cause, prevention and cure
  - example: Looking for occurrences of a DNA base sequence within 3 billion base pairs of the Human Genome
- NCRI Informatics: using informatics to maximise the impact of cancer research
  - coordination and guiding role
  - primary concern is data sharing
  - main obstacles: willingness, preparation, ability

# IT in science / cancer

- “The online communication of science is an ethical need”

Prof. Umberto Veronesi,  
European Institute of Oncology

- The EU Commission is pushing this
- Merck are establishing a large informatics programme
- Cyberinfrastructure for Biological Sciences\*
  - \* Stein, L.D. Nat Rev Genet. 2008 Sep;9(9):678-88.

# Why? The scale of the data

- Data generated:
  - in the lab
  - during trials of new drugs
  - from studies about the health of an entire population
- Every year approx. **250,000 patients** in Phase I, II and III trials.
  - includes industry and charity-sponsored studies
- Assume each patient gives at least 150 individual data elements then every year approx. **37,500,000** individual data items

# The scale of the problem

- 10 million doctors worldwide, and only 30,000 subscribers to The Lancet
- BUT, computers are ubiquitous around the world, the prospective audience is enormous, and e-health has huge potential
- A sustained, useable environment for online research is needed

# What *was* thought to be an online community?

- Previous attempts at communities were typically:
  - advert sites for relevant congresses
  - virtual meeting spaces
  - “groupware tools”
    - e.g. to maintain shared diaries
  - collection of links
  - news feeds
- None of these approaches are suitable because they aren't of enough value

# What *should* an online community be?

- A dynamic, self-sustaining resource of direct relevance to the day job of a cancer researcher/clinician?
- Google is closer to what's needed but is too 'amateur'
- We need a "Google for cancer research" which might have features such as:
  - ability to ask scientific/clinical questions
  - ability to structure complex, detailed queries
  - ability to collate and download useable results
  - access to published papers, laboratory data, data in public databases, images, anonymised patient data, labs/clinic working in this area etc
  - maintain patient confidentiality

# International scope

- US NCI's Centre for Bioinformatics
- UK National Health Service
- European Bioinformatics Institute
- European Organisation for Research and Treatment on Cancer
- Expanding links to Europe, China, India and Australia
- We want to include all

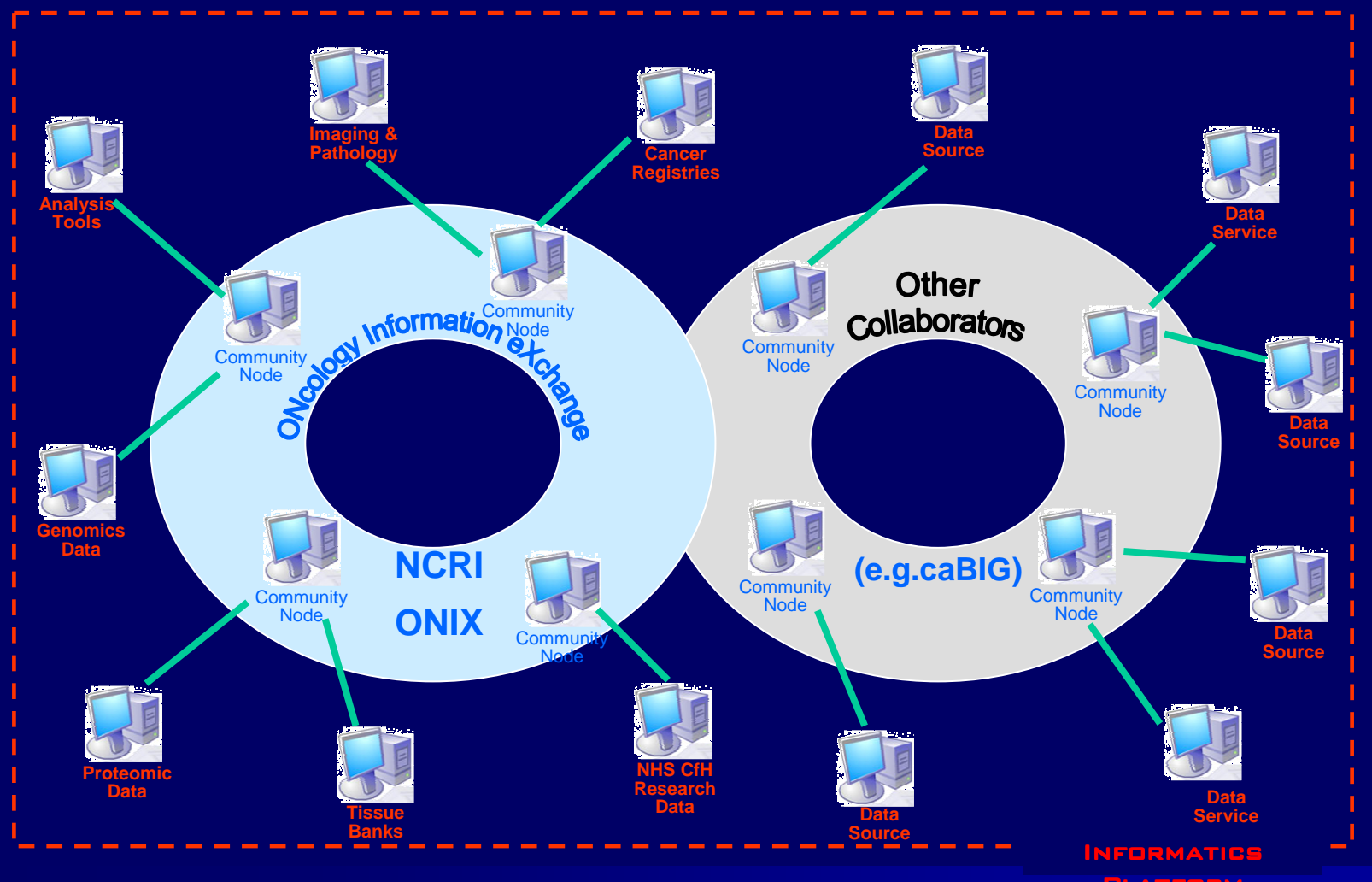
# ONcology Information eXchange (ONIX)

- Online portal to allow discovery and searching of multiple, geographically-disparate resources
- Data from 'omic' through clinical to population and economic and social
- Linked to cancer patient registry data
- Search many datasets simultaneously

# ONIX

- Sophisticated filtering
- Results are downloadable
- Find collaborators
- Open-source software
- Free to register
- Acts as the 'glue', bringing together, but not holding, information

# What does ONIX look like?



# ONIX: resource catalogue

NCRI Oncology Information Exchange - Windows Internet Explorer  
http://ncri-onix.org.uk/portal/#PS103

ONCOLOGY INFORMATION EXCHANGE

Training FAQ Help

Welcome Stuart  
X Log out of ONIX

HOME

RESOURCES  
Cancer InfoMatrix  
Resource Catalogue

SEARCHING

VOCABULARIES

DOCUMENT REPOSITORY

USER OPTIONS

RESOURCE OWNER

Home > Resources > Resource Catalogue

### Resource Catalogue

Catalogue

My filters

First Previous page 1 of 11 [showing results 1 to 10 of 110] Next Last

Resource Name	Interop Level	Country	Project Type	Metadata	Data
<b>BioMart</b>	4	International	Tools	n/a	n/a
<b>COSMIC</b>	4	United Kingdom	Multidiscipli...	n/a	n/a
<b>EMBL-EBI-ArrayExpress</b>	4	European	Multidiscipli...	n/a	n/a
<b>EMBL-EBI-IntAct</b>	4	European	Tools, Data s...	n/a	n/a
<b>EMBL-EBI-PRIDE</b>	4	European	Tools, Data s...	n/a	n/a
<b>HealthAgents</b>	4	European	Semantic tool...	n/a	n/a
<b>Human Protein Atlas (HPA)</b>	4	Sweden	International...	n/a	n/a
<b>IARC TP53 Database</b>	4	International	Data source	n/a	n/a
<b>ICGC</b>	4	International	Consortium	n/a	n/a
<b>Insight Toolkit</b>	4	International	Consortium, I...	n/a	n/a

Terms of Use | Contact Us | Privacy | Site Map | © ONIX 2009 | v1.0.1

Designed & maintained by NCRI INFORMATICS INITIATIVE

PS103

Internet | Protected Mode: On

100%

09:53

# ONIX: search facility

The screenshot shows the ONIX (Oncology Information Exchange) search facility in a Windows Internet Explorer browser. The browser's address bar displays the URL <http://ncri-onix.org.uk/portal/#PS116>. The page features a green header with the ONIX logo and navigation buttons for Training, FAQ, and Help. A left sidebar contains a navigation menu with categories like HOME, RESOURCES, SEARCHING, VOCABULARIES, DOCUMENT REPOSITORY, USER OPTIONS, and RESOURCE OWNER. The main content area is titled "Keyword Search" and includes a search input field, a "Search" button, and a list of searchable databases with checkboxes. The footer contains links for Terms of Use, Contact Us, Privacy, and Site Map, along with copyright information for ONIX 2009 and the NCRI Informatics Initiative.

NCRI ONCOLOGY INFORMATION EXCHANGE

Training FAQ Help

Welcome  
Stuart  
X [Log out of ONIX](#)

HOME  
RESOURCES  
SEARCHING  
Simple Search  
**Keyword Search**  
Investigations  
Published Investigations  
VOCABULARIES  
DOCUMENT REPOSITORY  
USER OPTIONS  
RESOURCE OWNER

Home > Searching > Keyword Search

### Keyword Search

Query

Search

[Select All](#) [Deselect All](#)

- [Entrez Protein](#)
- [NCBI Cancer Chromosomes](#)
- [NCBI GEO Datasets](#)
- [Entrez OMIM](#)
- [NCBI Probe](#)
- [Entrez PubChem Compound](#)
- [NCBI PubMed](#)
- [Entrez SNP](#)
- [Entrez Genome](#)
- [Entrez Unigene](#)

Terms of Use | Contact Us | Privacy | Site Map | © ONIX 2009 | v1.0.1

Designed & maintained by NCRI INFORMATICS INITIATIVE

PS116

Internet | Protected Mode: On 100%

09:55

# ONIX: search results

The screenshot shows the NCRI Oncology Information Exchange search results page. The search string is 'pten'. The results are displayed in a table with columns for Resource Name, Num of Results, View, and Metadata Filter. The table lists various resources such as NCBI Cancer Chromosomes, NCBI GEO Datasets, Entrez Protein, Entrez OMIM, NCBI Probe, Entrez PubChem Compound, NCBI PubMed, Entrez SNP, Entrez Unigene, and Entrez Genome, along with their respective result counts and links to view details and filter metadata.

NCRI ONCOLOGY INFORMATION EXCHANGE

Training FAQ Help

Welcome Stuart  
X Log out of ONIX

HOME  
RESOURCES  
SEARCHING  
Simple Search  
Keyword Search  
Investigations  
Published Investigations  
VOCABULARIES  
DOCUMENT REPOSITORY  
USER OPTIONS  
RESOURCE OWNER

Home > Searching > Keyword Search

### Keyword Search

Query Results

Search String : pten Save Marked Result Set

First Previous page 1 of 1 Next Last

Resource Name	Num of Results	View	Metadata Filter
<a href="#">NCBI Cancer Chromosomes</a>	0	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">NCBI GEO Datasets</a>	37	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez Protein</a>	1271	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez OMIM</a>	93	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">NCBI Probe</a>	338	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez PubChem Compound</a>	0	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">NCBI PubMed</a>	3814	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez SNP</a>	1163	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez Unigene</a>	64	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>
<a href="#">Entrez Genome</a>	11	<a href="#">View</a>	<a href="#">▶ Metadata Filter</a>

Terms of Use | Contact Us | Privacy | Site Map | © ONIX 2009 | v1.0.1  
Designed & maintained by NCRI INFORMATICS INITIATIVE PS116

Done Internet | Protected Mode: On 100% 09:56

# Future vision of ONIX

- ONIX public release – summer 2009
- Incremental releases of ONIX with:
  - Increasing functionality (2009, 2010)
  - Steadily increase the “resources” connected
  - Move from passive to pro-active communicators
- Expand beyond the UK, Europe & US
  - Firming up plans for working with other initiatives (RCP, NCI, EORTC etc.)
  - Looking for other initiatives and projects for ONIX

# Some ways of using ONIX

- Researchers & Clinicians
  - Who is working on similar problems?
  - Finding out which resources have got information
  - One search goes across multiple data types
  - Searches can be stored and repeated over time
- Funders & Governance
  - What scientific questions are being asked?
  - What resources are being used?
  - What is the quality of a resource?
- General
  - How is the picture different in other countries?

# Tele-pathology

- Electronic transmission of pathological images
- Case referral
- Provision of emergency services in absence of resident pathologist
- Aid diagnosis
- Training and CPD application

# Tele-pathology

- Imaging and pathology data are an intrinsic part of ONIX
- Imaging & Pathology demonstrator
  - Multiple modality images can be integrated into data sets and databases to extend the range of searchable clinical information
  - Call up slides for comparison
- Viewing images online saves on travel, and can be undertaken at times convenient to the user

# Online analysis

- HealthAgents

- A system to help analyse MRI spectra against uploaded spectra with complete patient histories
- Gives a prognosis compared against other patient outcomes

- Workflow

- IT tools to automate sequential analyses of data

- There exists a wealth of online tools which can be downloaded to facilitate data analysis

# ONIX as a teaching tool

- ONIX is already being used in medical undergraduate teaching courses in the UK e.g. University College London
- We plan to promote this to all UK medical schools
- Helps doctors-in-training understand 'informatics' and shows them how it can benefit daily practice

# Educational Resources

- iPath – digital pathology repository for training
- MERLOT (Multimedia Educational Resource for Learning and Online Teaching) – California State University
- OERC (Open Educational Resources for Cancer) – INCTR initiative

# Informatics Training

- Many NCRI Partners have set aside funding for interdisciplinary training
  - Designed to allow periods of study in areas not directly related to present career paths
- NCICB – e-learning resource for caBIG
- ECDL/ICDL Health
  - Informatics training resource for doctors, nurses and other health professionals
  - Recognises skill barrier in uptake of informatics resources

# Conclusion

- Informatics is of growing significance
- Helps manage the wealth of (un)available data
- Can lead to demonstrable benefits
  - Time saving
  - Cost saving
  - Streamlining of research activities
  - Educational applications
  - Assist in diagnosis
  - Share technical expertise and resources
  - Receive the benefits of research investment of others

# Conclusion

- Systems such as ONIX can provide a real benefit to researchers and clinicians in any part of the world
- We hope it will become an international and permanent 'research community'
- The principles of ONIX can be applied to any disease area
- Welcome suggestions for ONIX resources

# The Future

“The future is here. It is just not evenly distributed”.

Prof. Sir Muir Gray  
Chief Knowledge Officer  
UK National Health Service

# Thank You

[stuart.bell@ncri.org.uk](mailto:stuart.bell@ncri.org.uk)

<http://www.cancerinformatics.org.uk/>