The Social Grid

Leveraging the Power of the Web and Focusing on Development Simplicity

Tony Hey
Corporate Vice President of Technical Computing at Microsoft

successful standards efforts

- TCP/IP versus ISO Protocols
  - ISO Committees disconnected with operational world
  - IETF solved and responded to real operational problems
- MPI community message-passing standard
  - Joint effort between research community and the major vendors (IBM, Meiko, Intel, TMC, Ncube)
- Web
  - Pioneered by research community
  - HTML and HTTP were a pragmatic compromise

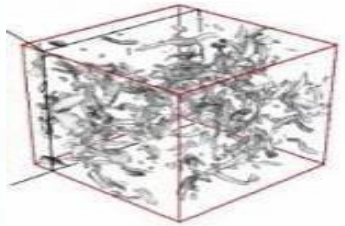
the grid community

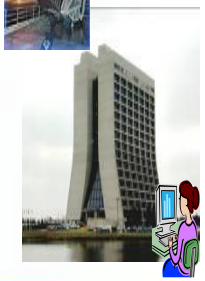
- Initiated by the research community
  - Distributed, collaborative computing
  - HPC Centers
- Successful Grids in industry
  - Finance, Pharma, Oil & Gas
  - Intra-Grids in Business
- Cloud Service providers
  - Google, Amazon, eBay, Yahoo and Microsoft

emergence of a new science paradigm

- Thousand years ago – **Experimental Science**
  - Description of natural phenomena
- Last few hundred years – **Theoretical Science**
  - Newton's Laws, Maxwell's Equations…
- Last few decades – **Computational Science**
  - Simulation of complex phenomena
- Today – **eScience or Data-centric Science**
  - Unify theory, experiment, and simulation
  - Using data exploration and data mining
    - Data captured by instruments
    - Data generated by simulations
    - Data generated by sensor networks
  - ➢ Scientist analyzes databases/files
  - ➢ Commercial and academic applications

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





(With thanks to Jim Gray)

motivation

- Grid middleware and Web Services are becoming very complex and hard to use

➢ Need to go back and look at needs of users

- Recognize the transition to data-centric eScience
- Provide support for domain scientists
- Focus on user and application requirements
- Need for simplicity and ease of development
- Learn lessons from the Web

the web

- Has tremendous momentum
- The browser is the universal canvas for the delivery of information and functionality
- Web protocols, technologies, and middleware are well supported by the IT industry
- Easy to use collection of ‘Web 2.0’ technologies is maturing
- Emergence of “Software-as-a-service”

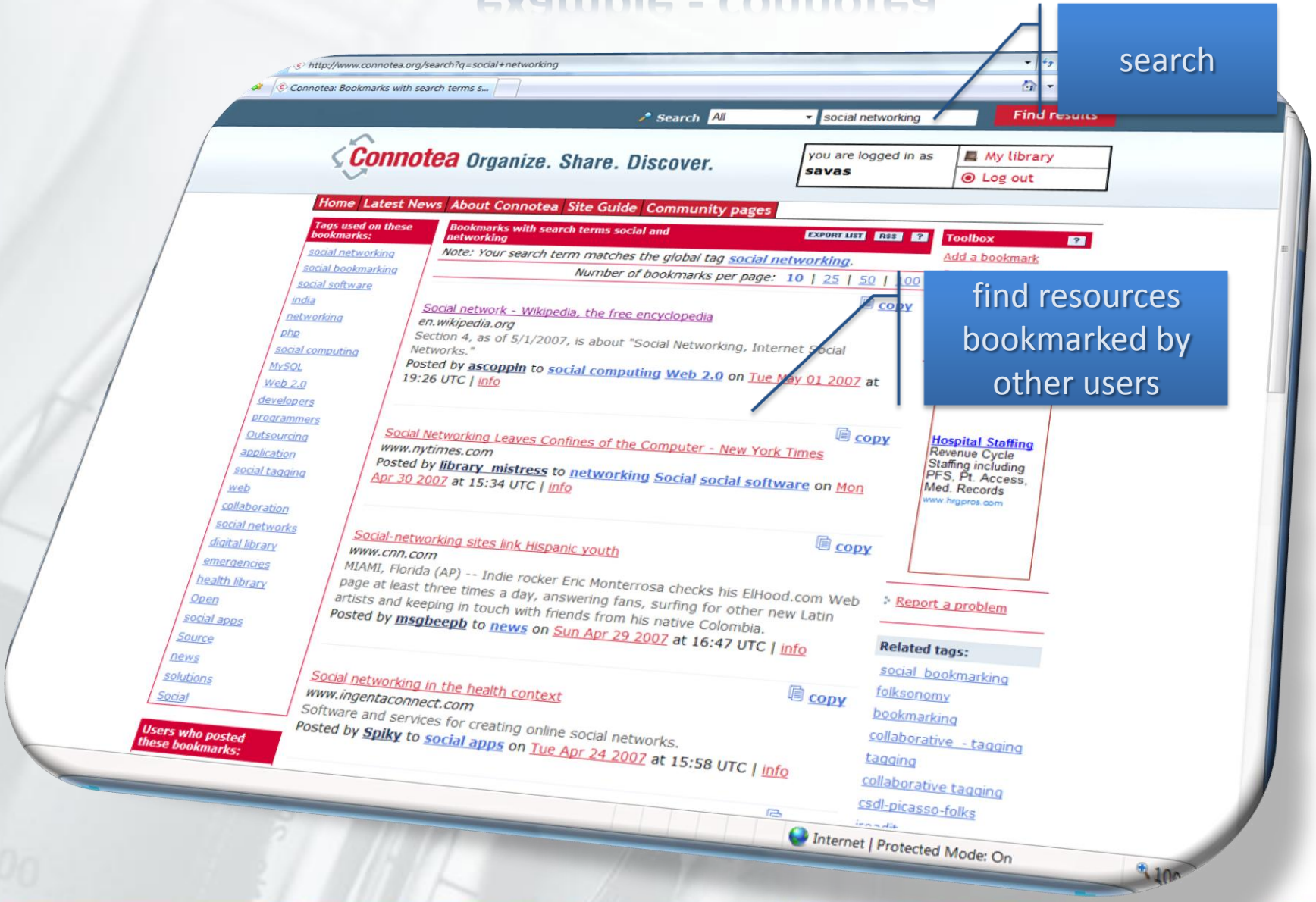
users are both consumers and producers

- Users generate content on the Web
  - Blogs, wikis, photographs, videos, etc.
  - They do not have to know HTML
- They form communities
  - Social networks, virtual worlds
- They interact, collaborate, share
  - Instant messaging, web forums, content sites
- They consume information and services
  - Search, annotate, syndicate

scientists today…

- Annotate, share, discover data
- Collaborate, exchange ideas over the Web
- Create communities, social networks
- Use workflow tools to compose services

example - connotea

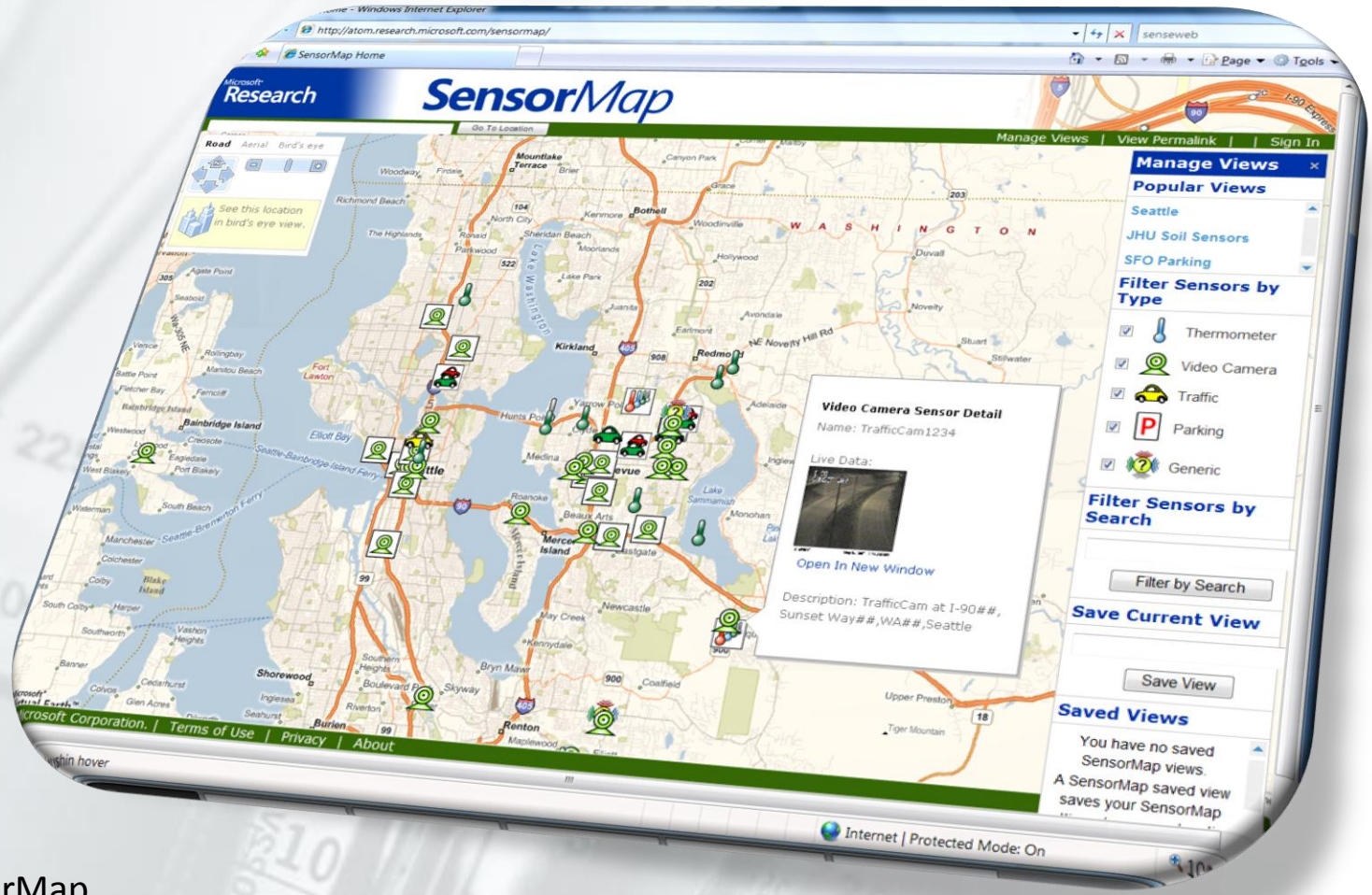


the web application platform

- Service composition as a way to build distributed applications
  - Service-orientation

Yahoo pipes
Aggregate news about “Grid Computing”

mashups: composing data and functionality



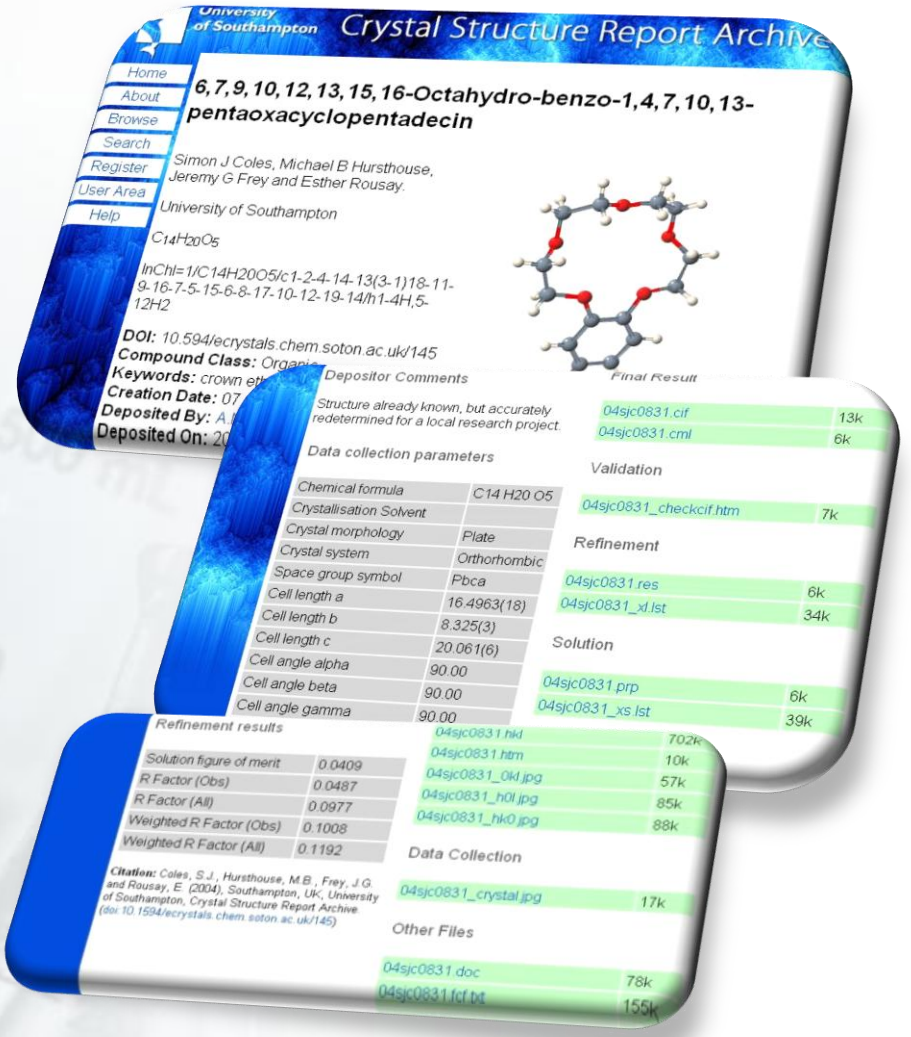
SensorMap

Functionality: Map navigation

Data: sensor-generated temperature, video camera feed, traffic feeds, etc.

the web application platform

- Services on the cloud
  - Blogging
  - Data processing/transformation
  - Content upload, sharing, discovery
  - Storage, computation, messaging



http://ecrystals.chem.soton.ac.uk

Thanks to Jeremy Frey

amazon web services: simple storage service (s3)

- S3 is storage for the Internet
  - Designed to make web-scale computing easier for developers
- Provides a simple Web Services interface to store and retrieve any amount of data from anywhere on the Web
  - 'CRUD' philosophy – Create, Read, Update and Delete operations
- Uses simple standards-based REST and SOAP Web Service interfaces
  - Built to be flexible so that protocol or functional layers can easily be added

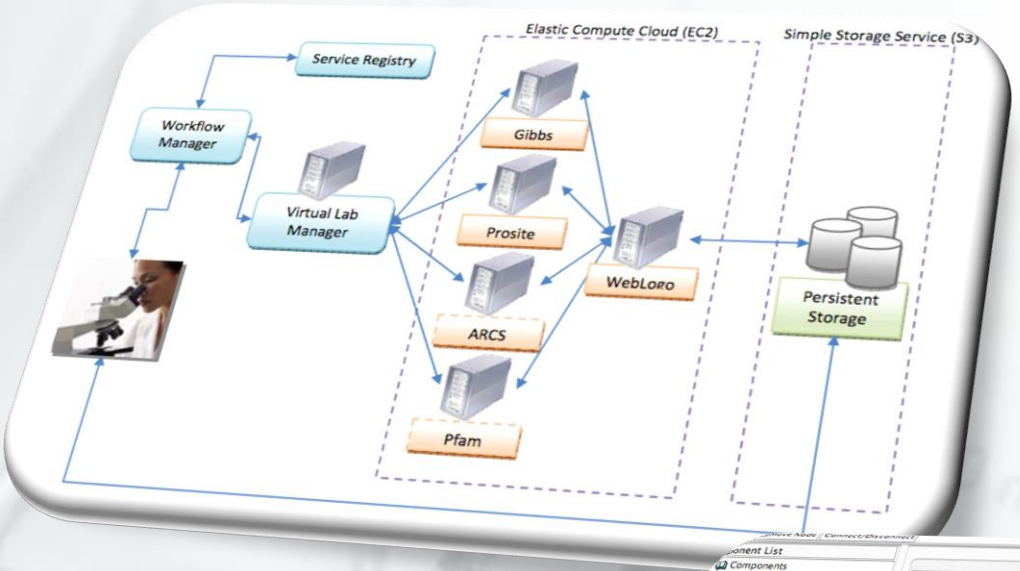
amazon s3 functionality

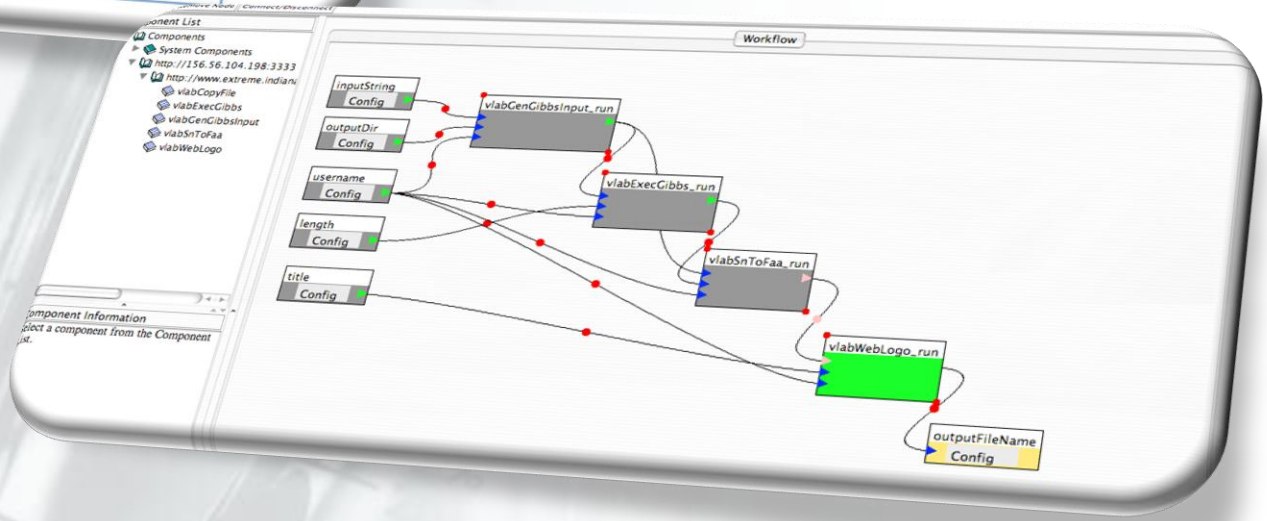
- Intentionally built with a minimal feature set
  - Write, read, and delete objects containing from 1 byte to 5 gigabytes of data each
- Can store unlimited number of objects
  - Each object is stored and retrieved via a unique, developer-assigned key
- Authentication mechanisms provided
  - Objects can be made private or public, and rights can be granted to specific users
- Default download protocol is HTTP
  - BitTorrent protocol interface is provided to lower costs for high-scale distribution

amazon web services: elastic compute cloud (ec2)

- Compute on demand service that works seamlessly with their S3 storage service
- Create Amazon Machine Image (AMI) containing application, libraries and data
- Use EC2 Web Service to configure security and network access
- Use EC2 to start, terminate and monitor as many instances of your AMI as you want
  - Each instance has:
    - 1.7 GHz x86 Processor
    - 1.75 GB RAM
    - 160 GB local disk
    - 250 MB/s network bandwidth
- Used by Catlett and Beckman as capacity computing alternative to TeraGrid 'SPRUCE' capability computing for emergency urgent response

a grad student project using s3 and ec2





Gene Analysis Virtual Lab Experiment

by Jong Youl Choi
at Indiana
(For Beth Plale and Sun Kim)

ogf and standards for grid computing

- Protocol-based integration
  - Based on the Web Services stack
    - WS-* stack has become very complex
    - Danger that Grid-* stack will become too complex
  - Much effort on systems management specifications
  - HPC Basic Profile as the example of a domain-specific specification with industry adoption
- Standards adoption
  - Why are major Grid users/software providers not more engaged with OGF?

grids in industry

- Google, Amazon, Yahoo, eBay and Microsoft are the major ‘Cloud Platform’ providers
  - All have infrastructures of hundreds of thousands of servers
  - Many large data centers, distributed across multiple continents
  - Have developed proprietary technologies for job scheduling, data sharing and management
  - Care about power consumption, fault tolerance, scalability, operational costs, performance, etc.
  - They are living the “Grid dream” on a daily basis

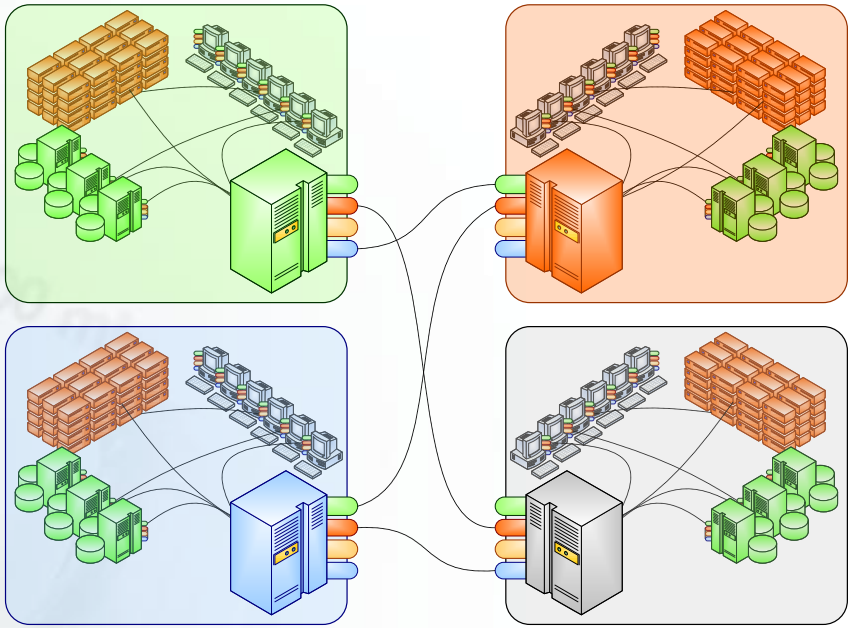
google as an example

- Estimated 450,000* servers distributed around the world

*source: Wikipedia

- Google File System - highly distributed, resilient to failures, parallel, etc.
- Schedulers and load balancers for the distribution of work
  - Use their ‘Map-Reduce’ middleware as parallel computational model

- What is financial/competitive advantage for them to adopt standards for these in-house developed solutions?



industry and standards?

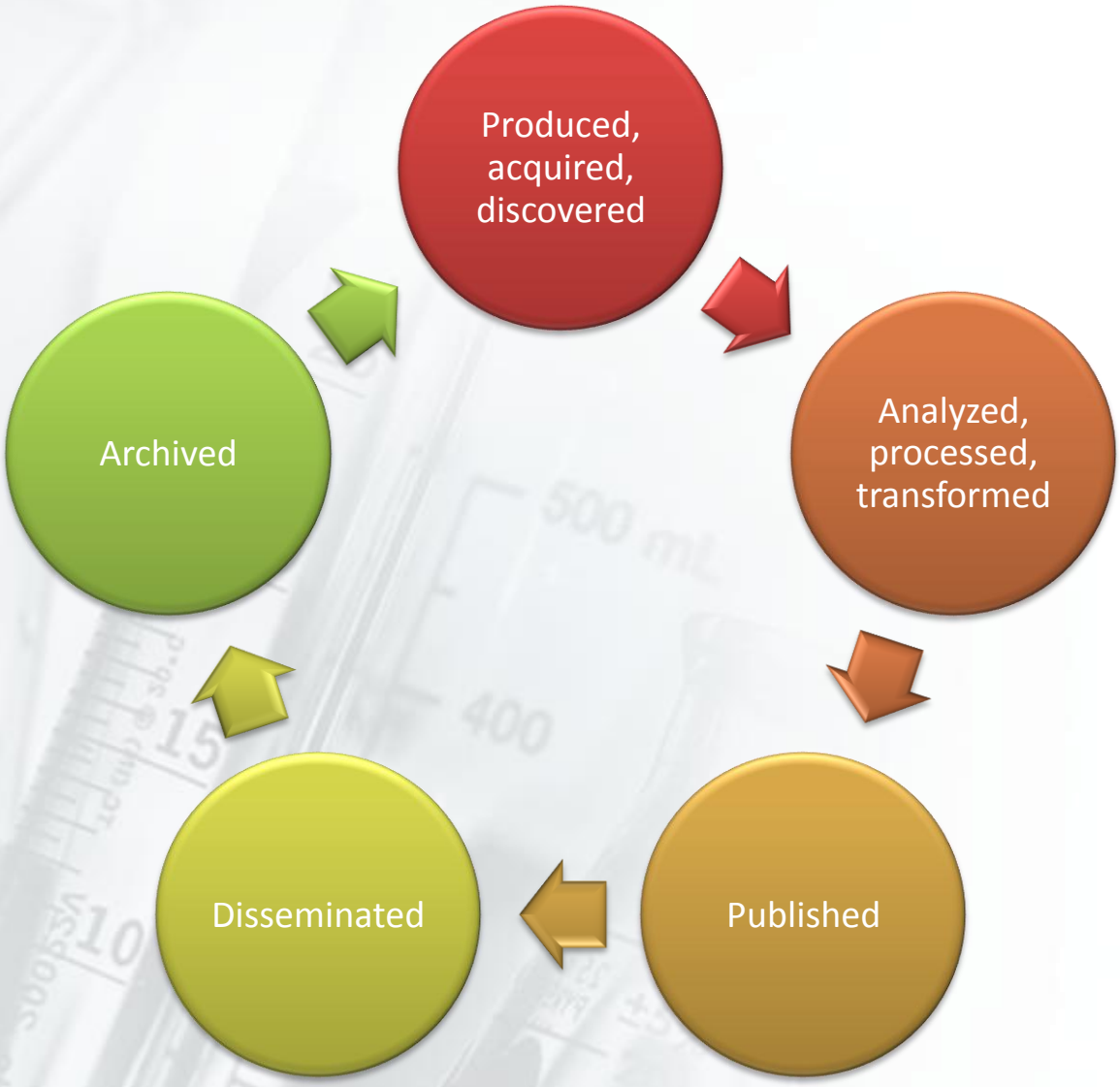
- Why are they not more involved in Grid standards?
  - See competitive advantage in their proprietary infrastructure and proprietary services
- No need for interoperation between companies
  - Will not see map-reduce tasks from Google being offloaded to Amazon's infrastructure any time soon
  - Security and privacy are still major concerns
  - Probably still cheaper for them to maintain their own infrastructure than to offload to someone else's
- There are some possibilities for standardization
  - WS-Management standards could reduce the operational cost of large collections of commodity infrastructure components - motherboards, disks, network hubs, etc.

customer demand can drive interoperability solutions

Example of Security

- Customers want secure access/confidentiality no matter what service they use
- No competitive advantage in doing simple security in a proprietary way
  - Customer demand is forcing Microsoft's Passport, Google's ID, Liberty Alliance, OpenID to interoperate
- More complicated 'VO' scenarios have yet to be proven widely useful
  - Require cross organization trust, credential delegation, id federation
  - Work has already started in this space e.g. WS-Trust, WS-Policy
  - VOs required for some research collaborations
  - Perhaps B2B scenarios like Supply Chain Management can provide commercial driver for their wide adoption

scientific data on the web is…



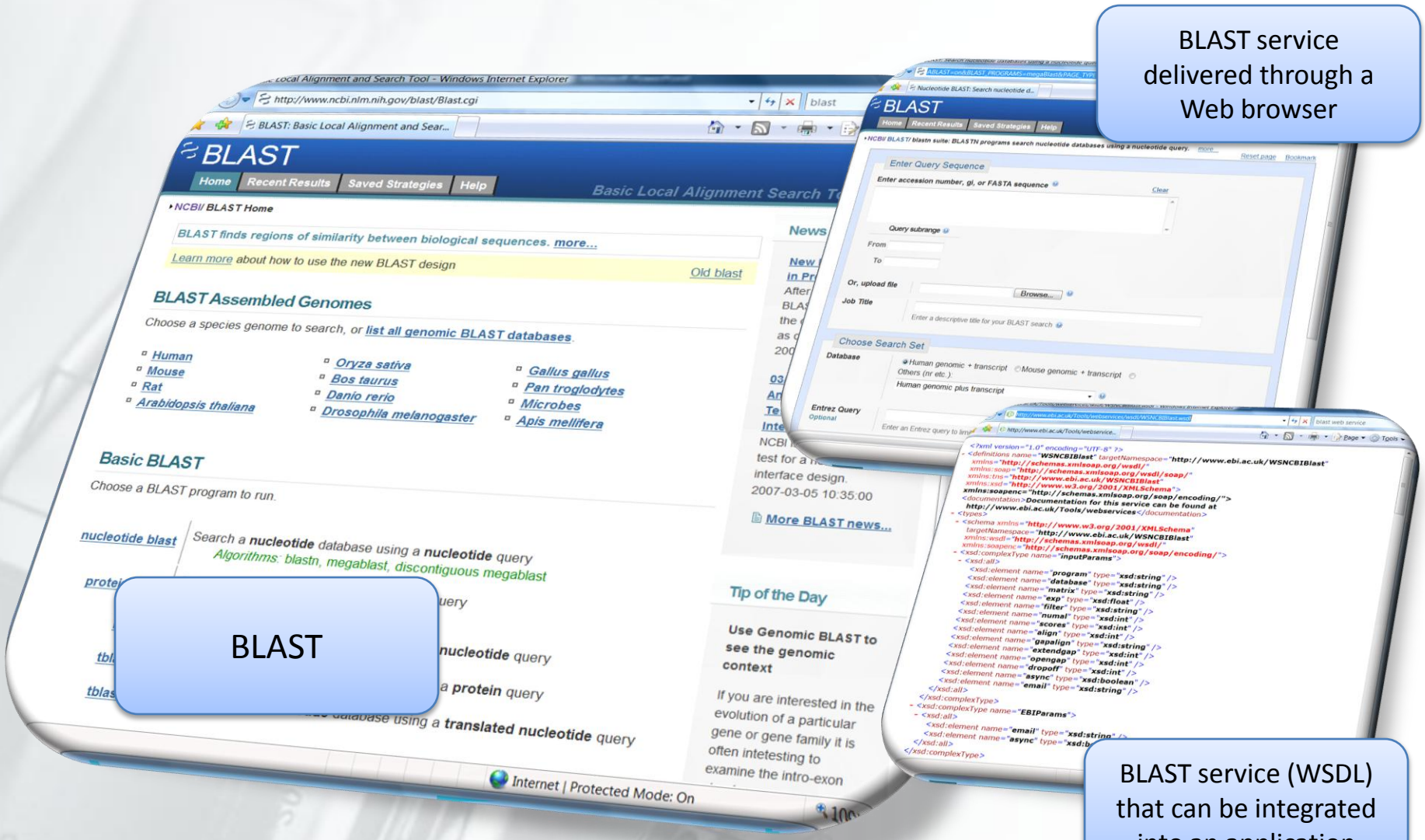
eScience: data is easily accessible

With thanks to
Catharine van Ingen

eScience: data is easily shareable

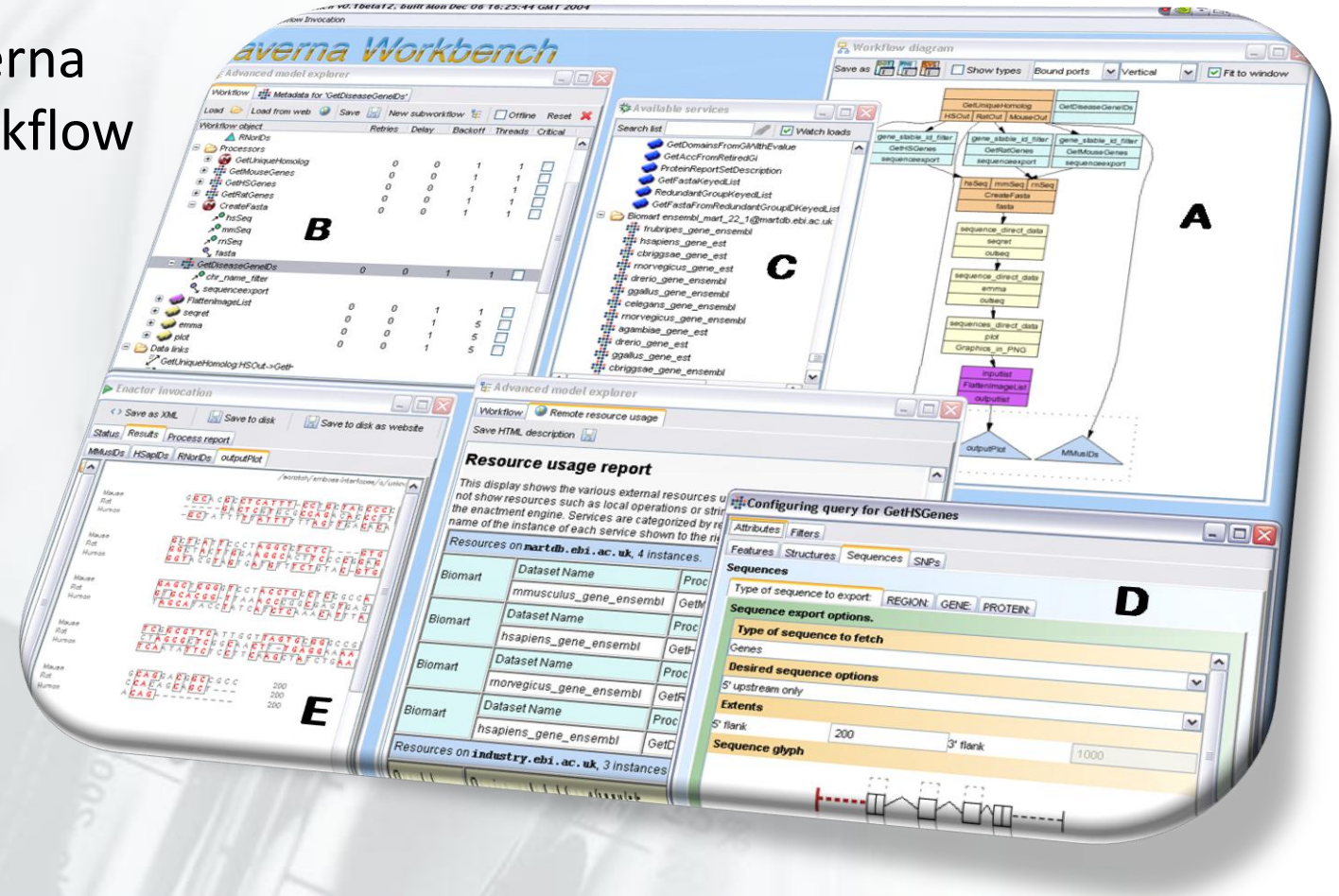
Sloan Digital Sky Server/SkyServer
http://cas.sdss.org/dr5/en/

eScience: services expose functionality



eScience: services can be composed

Taverna Workflow



eScience: data and services can be accessed securely

security on the web

- HTTPS
- XML Encryption
- WS-Security
- Single-sign on
  - Windows Live ID
  - Open ID
  - Google ID
- Virtual Organizations through portals
  - Aggregation of services into one portal, with single authentication

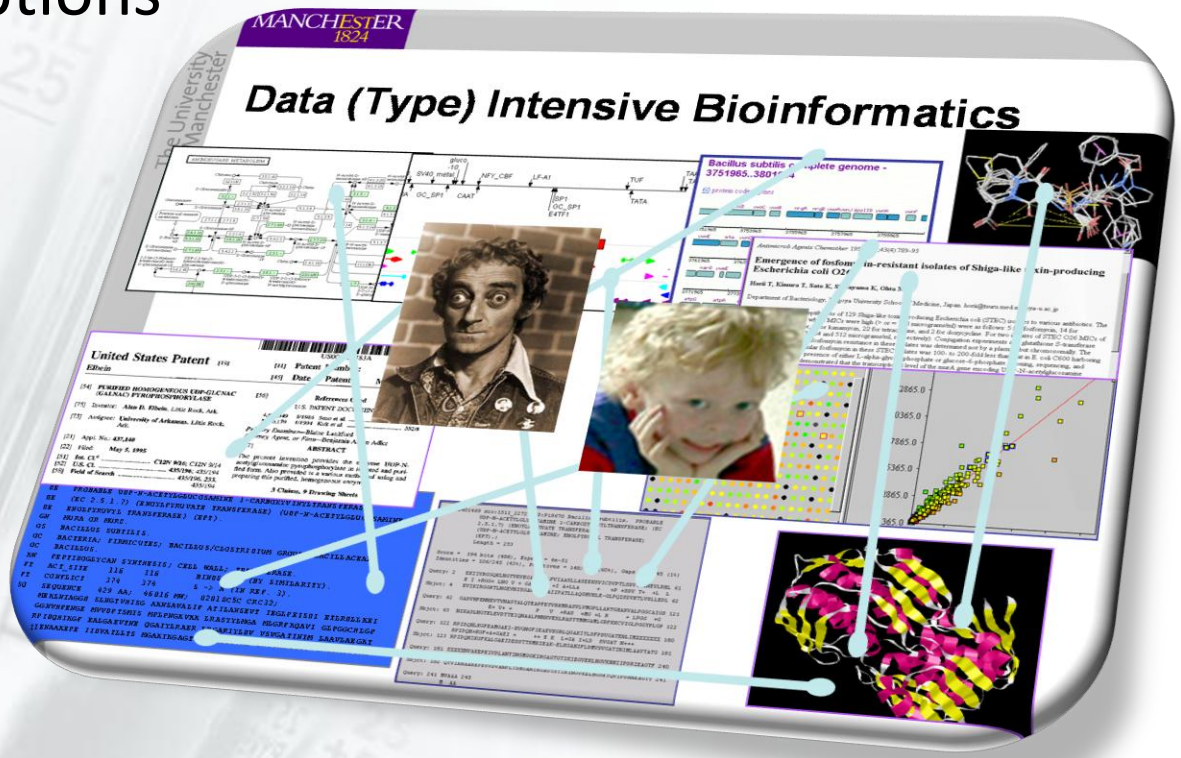
➢ Grid Security?

  - Need to cater for many different user communities

eScience: knowledge can be created/published/archived/discovered

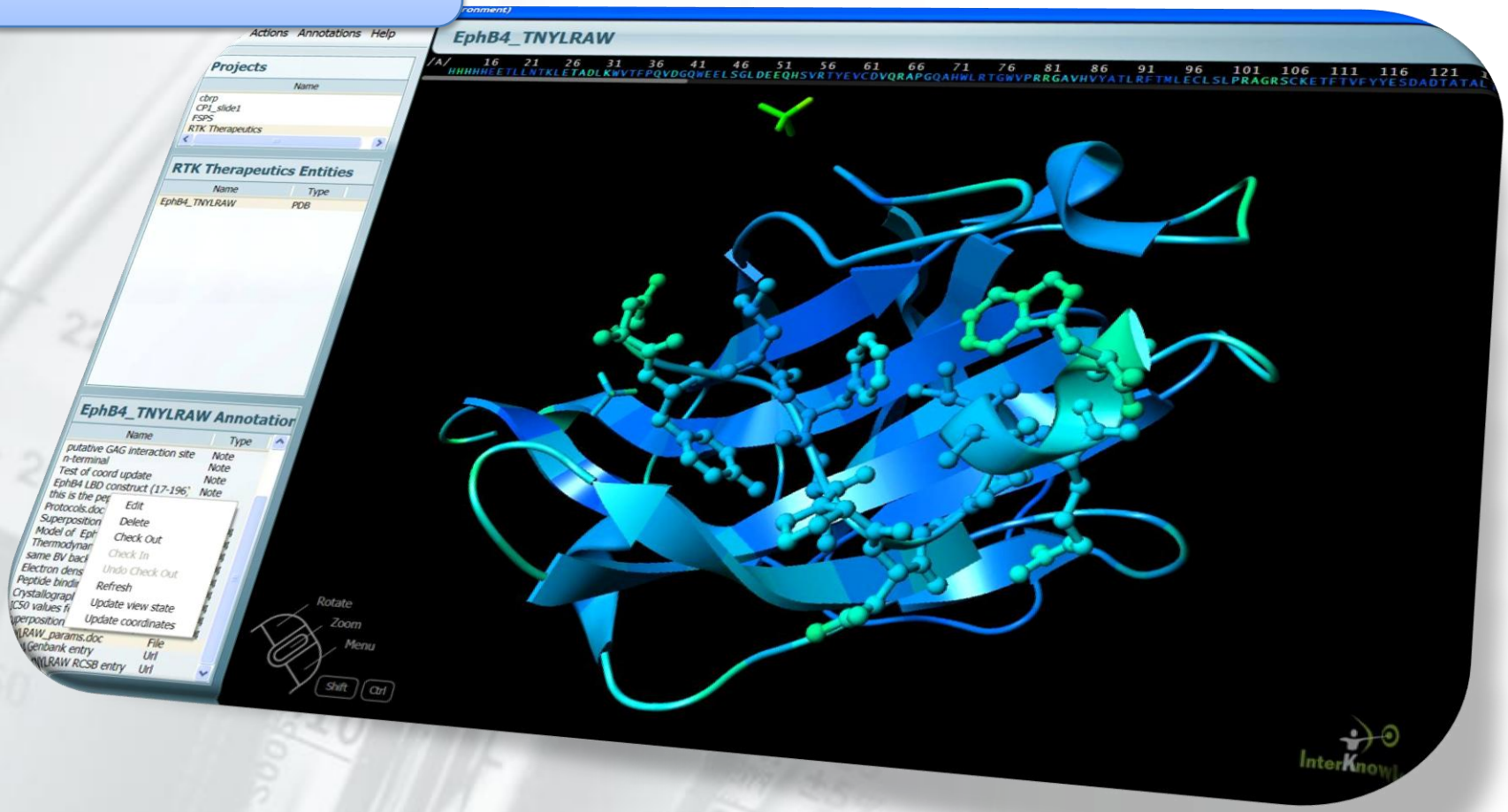
- Semantic relationships between different data
- Semantic descriptions of services
- Annotations
- Provenance
- Repositories
- Ontologies

myGrid

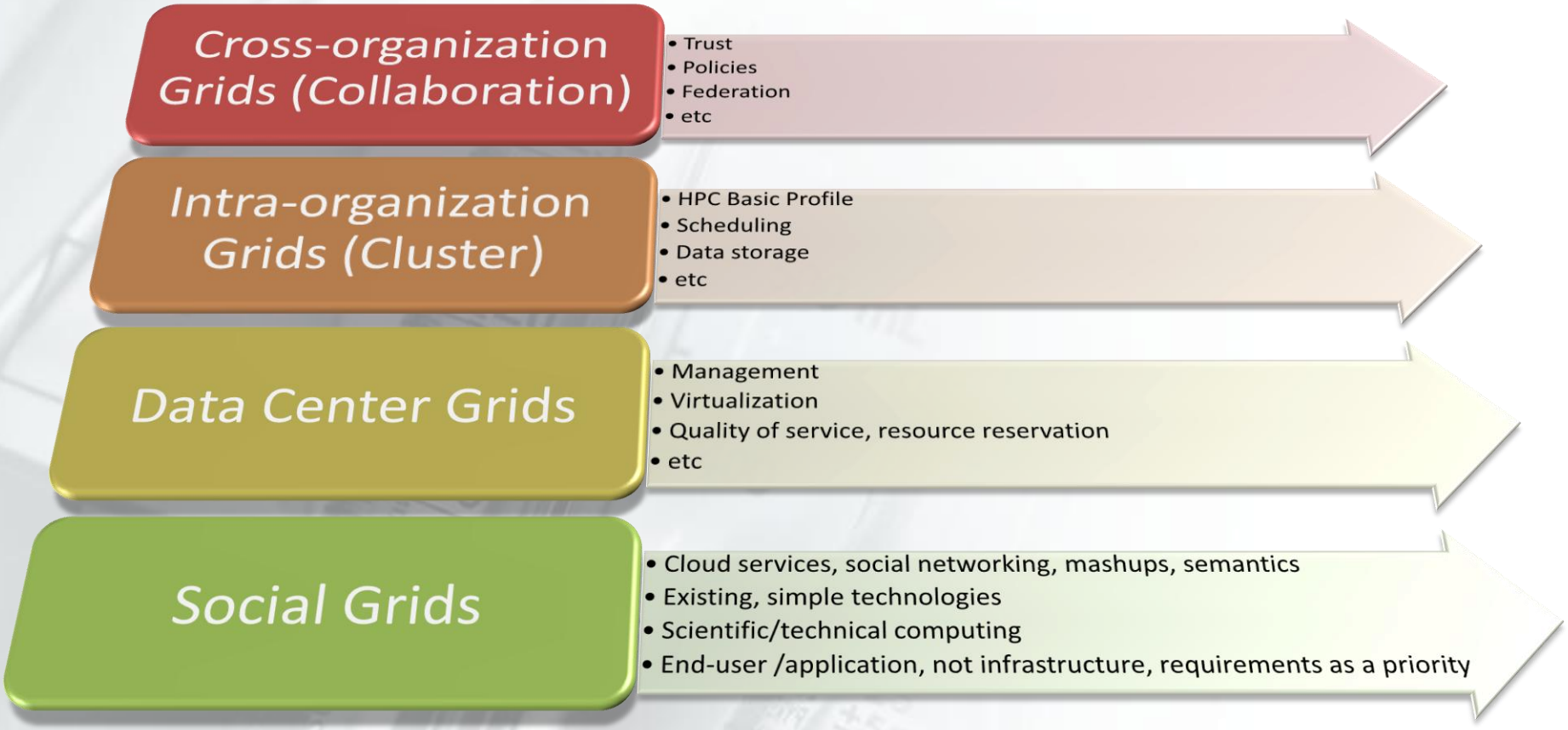


eScience: new software tools can make a difference

SCRIPPS
(see www.microsoft.com/science)



a social grid built around the web?



social grids and the web

- Focus on **solutions** for scientific/technical computing and **not just on infrastructure**
- Focus on “data-centric eScience”
  - Help domain experts define formats for representing and annotating domain-specific data
- Keep it simple, build on known Web technologies
  - Solutions that “just work” without the need for complicated middleware platforms
  - Leverage only existing, Web infrastructure (HTTP, XML, simple Web Services, services in the cloud)

call to action

- Need for a new type of Grid
  - Domain-specific
  - Provide support for scientists to manage their domain's data
    - Engage with them to define domain-specific data formats
    - Educate them on the use of the Web as a platform for managing their domain's information
    - Demonstrate the use of semantics to manage knowledge
  - User-driven requirements
  - Scientists are our largest audience, let's help them

Microsoft®

Your potential. Our passion.™