

---

# Simbios - Promoting Open Science in Biocomputational Research

---

Joy P. Ku, Ph.D.

Director of Dissemination

Simbios, an NIH National Center for Biomedical Computing

# Open Access

---

- Open Access = ability to freely read and use scientific and scholarly material
- Many recent successes on the open access front in 2008

Open access does not address an important issue in biocomputational research

# Example: Why Open Access Is Not Enough for Biocomputational Research

## Muscle contributions to support and progression over a range of walking speeds

May Q. Liu<sup>a</sup>, Frank C. Anderson<sup>b</sup>, Michael H. Schwartz<sup>c</sup>, Scott L. Delp<sup>a,b,\*</sup>

<sup>a</sup> Department of Mechanical Engineering, Clark Center, Room S-321, Stanford University, Mail Code 5450, 318 Campus Drive, Stanford CA 94305-5450, USA

<sup>b</sup> Department of Bioengineering, Clark Center, Room S-321, Stanford University, Mail Code 5450, 318 Campus Drive, Stanford CA 94305-5450, USA

<sup>c</sup> Center for Gait and Motion Analysis, Gillette Children's Specialty Healthcare, St. Paul, MN, USA

### ARTICLE INFO

Article history:  
Accepted 16 July 2008

Keywords:  
Pediatric gait  
Forward dynamic simulation  
Muscle function  
Walking speed  
Propulsion

### ABSTRACT

Muscles actuate walking by providing vertical support and forward progression of the mass center. To quantify muscle contributions to vertical support and forward progression (i.e., vertical and fore-aft accelerations of the mass center) over a range of walking speeds, three-dimensional muscle-actuated simulations of gait were generated and analyzed for eight subjects walking overground at very slow, slow, free, and fast speeds. We found that gluteus maximus, gluteus medius, vasti, hamstrings, gastrocnemius, and soleus were the primary contributors to support and progression at all speeds. With the exception of gluteus medius, contributions from these muscles generally increased with walking speed. During very slow and slow walking speeds, vertical support in early stance was primarily provided by a straighter limb, such that skeletal alignment, rather than muscles, provided resistance to gravity. When walking speed increased from slow to free, contributions to support from vasti and soleus increased dramatically. Greater stance-phase knee flexion during free and fast walking speeds caused increased vasti force, which provided support but also slowed progression, while contralateral soleus simultaneously provided increased propulsion. This study provides reference data for muscle contributions to support and progression over a wide range of walking speeds and highlights the importance of walking speed when evaluating muscle function.

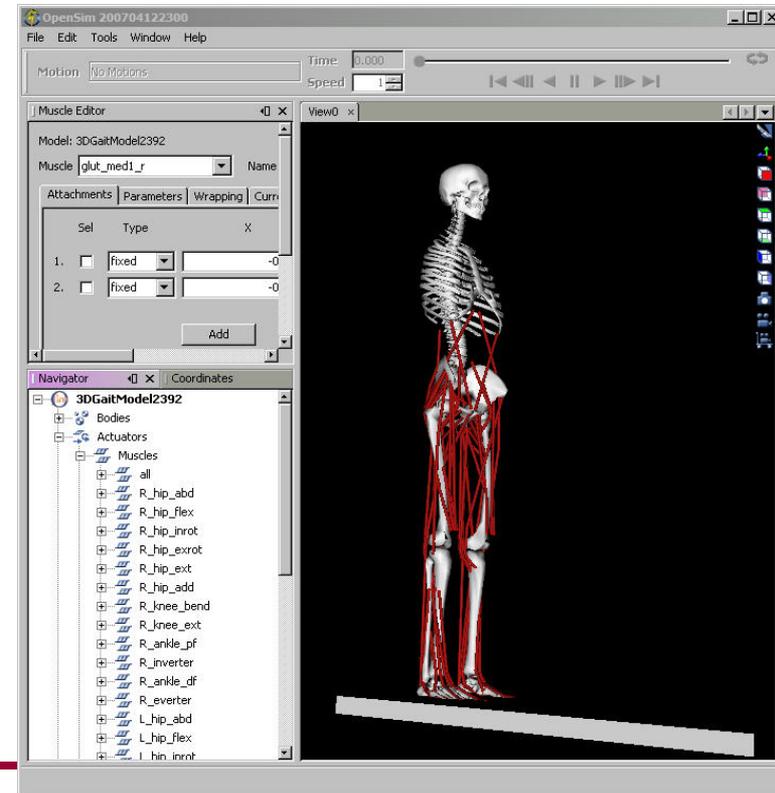
© 2008 Published by Elsevier Ltd.

Liu, MQ, Anderson, FC, Schwartz, MH, Delp, SL, "Muscle contributions to support and progression over a range of walking speeds," *Journal of Biomechanics*, Nov. 2008, 41(15): 3243-3252

# Example: Why Open Access Is Not Enough for Biocomputational Research

Reproduction of published results from biocomputational research is often very difficult because of limited access to:

- **Software:** Dr. Liu's work uses OpenSim, software developed over many years by many people
- **Data/Model:** Dr. Liu's model has 23 degrees of freedom, 92 muscles/tendons



# Simbios' Efforts

---

OPEN SCIENCE

Open  
Access

Open  
Source  
(Software)

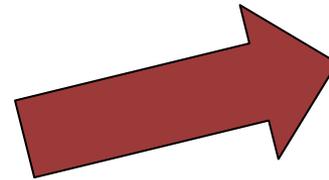
Open  
Data

# What Is Needed: **Effective** Sharing of Research Software, Models, and Data

---

To accelerate progress in the field, there needs to be minimal re-invention

Researchers need to be willing to provide software, models, data, etc. from their research



Researchers need to be willing to use the software, models, etc. that have been made available

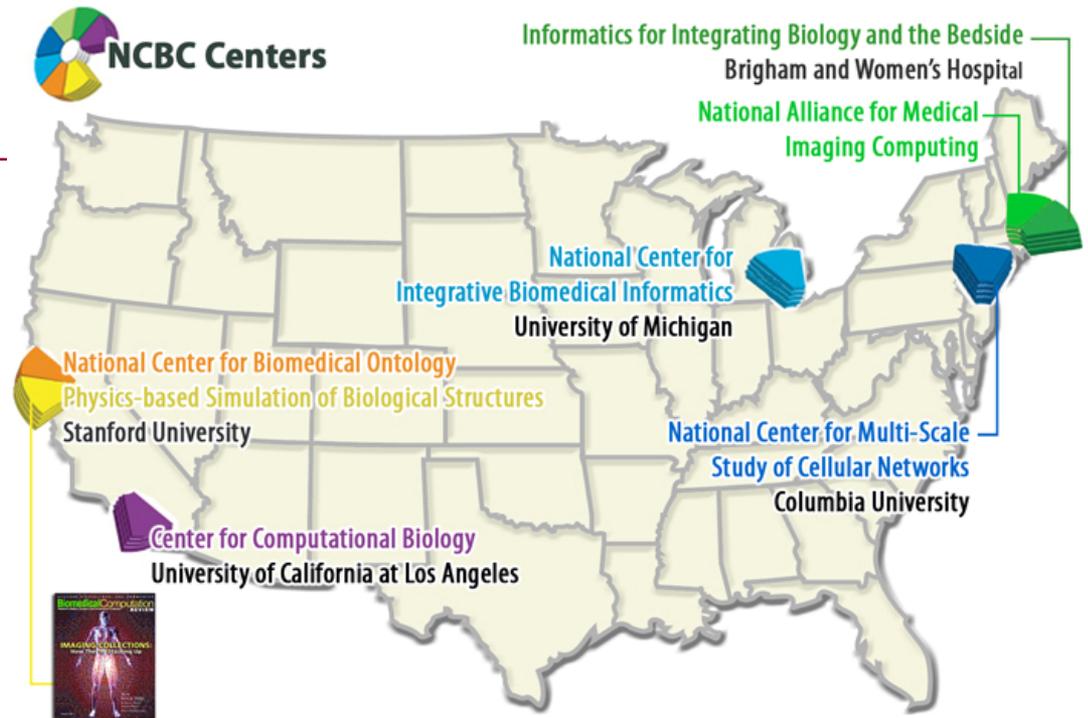
# Outline

---

- Introduction to Simbios
- Using Simtk.org to share software and data
- Challenges in promoting open science

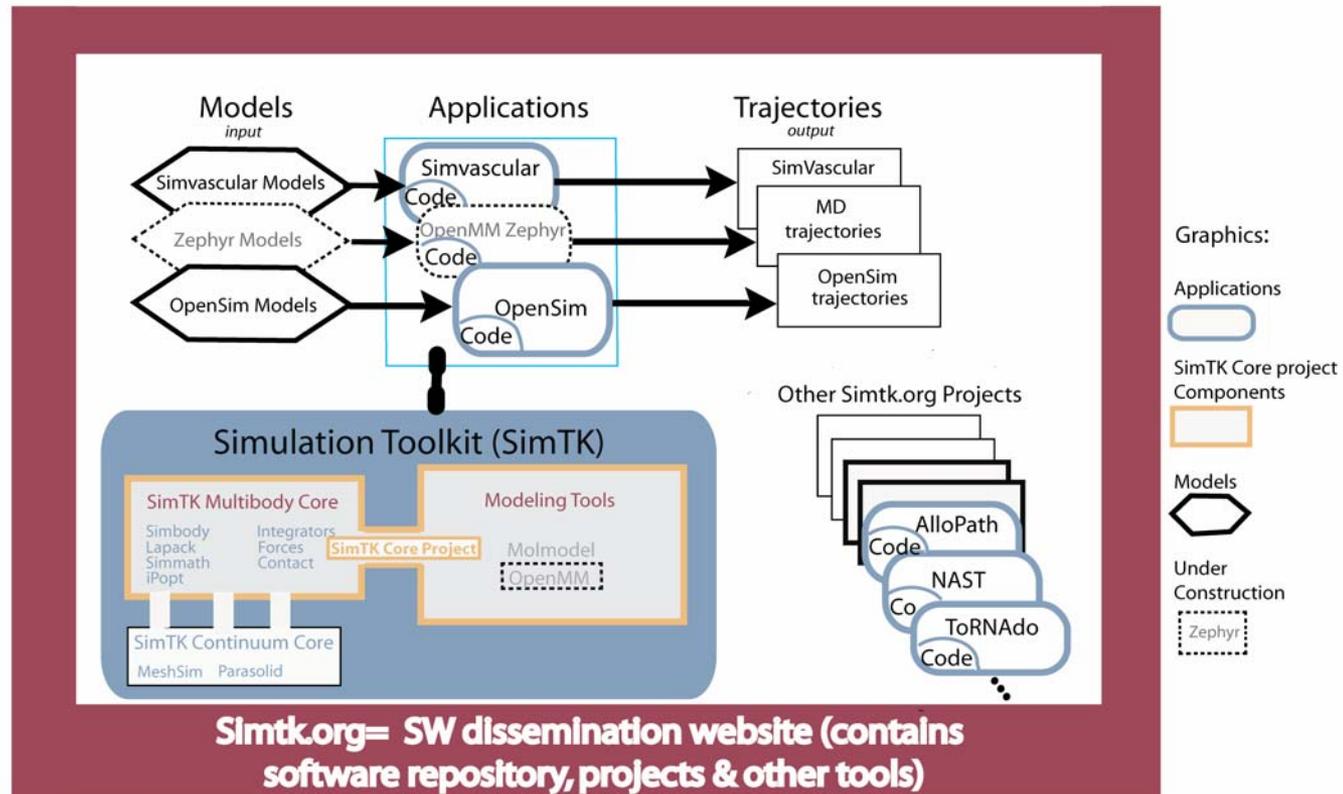
# Simbios

- One of 7 National Centers for Biomedical Computing (NCBCs)



- Funded by the National Institutes of Health (NIH) as part of the Roadmap for Bioinformatics and Computational Biology
- <http://www.ncbcs.org>

# Simbios: Physics-Based Simulation of Biological Structures



Simbios software offerings and organization:

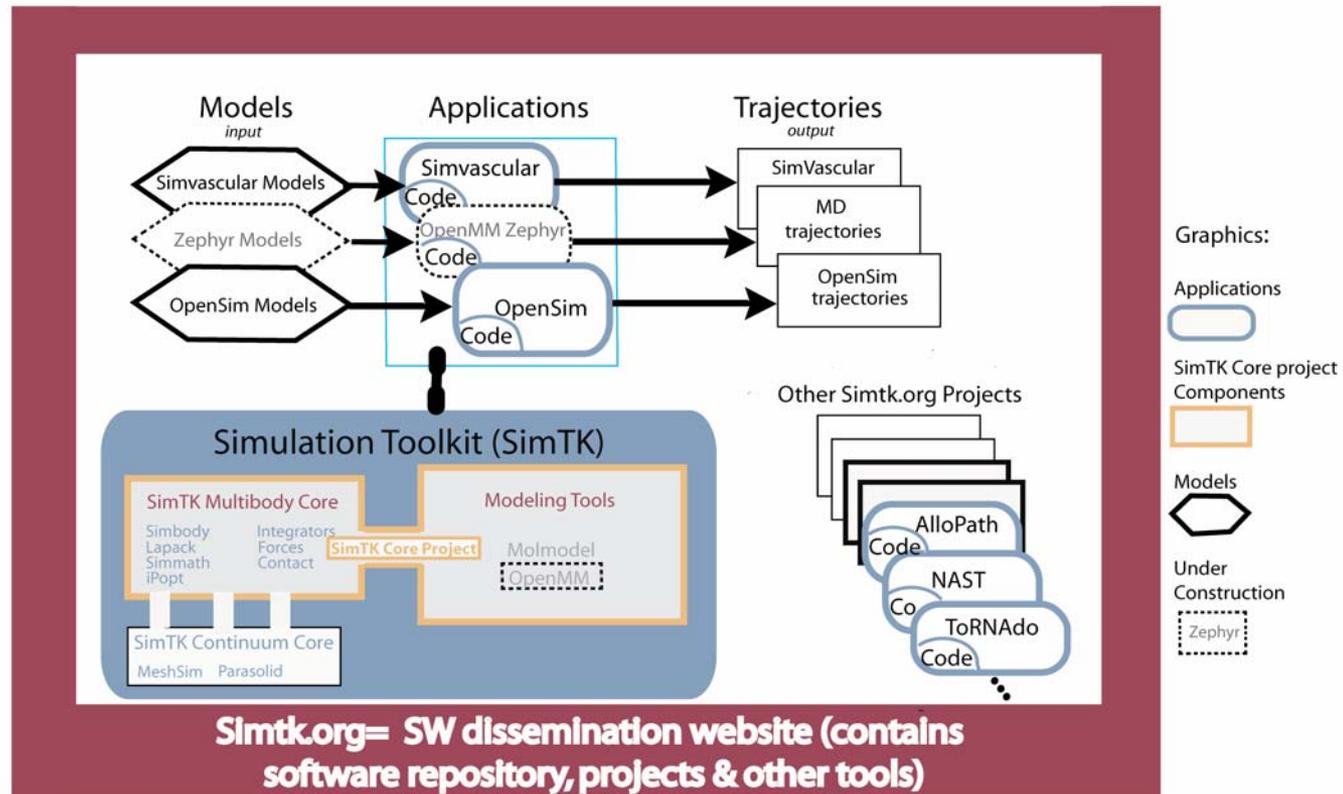
Simtk.org is Simbios's dissemination website for Software and data. The basic unit of work on Simtk.org is a projects. Each project is equipped, among others, with a SVN repository to store code. Projects have different content, some contain applications, (shown as ) others contain models ()—often the input to simulation applications, others contain trajectories -- the output of simulation applications, and others contain yet different data. Simbios's simulation toolkit, shown on the lower left, also resides on simtk.org. The yellow framed part of the SimTK toolkit can be downloaded as one package from the project:simtk.org/home/simtkcore. Simbios's main applications (shown above center) are built upon and use the Simulation Toolkit, others do not. SimVascular uses commercial components for it's toolkit, shown in the lower

# SimTK Core

---

- Toolkit of high-performance computational modules for simulating physics of biological structures
- Written in C++
- Includes code for linear algebra, numerical integration, multibody dynamics (Simbody), ...

# Simbios: Physics-Based Simulation of Biological Structures



Simbios software offerings and organization:

Simtk.org is Simbios's dissemination website for Software and data. The basic unit of work on Simtk.org is a projects. Each project is equipped, among others, with a SVN repository to store code. Projects have different content, some contain applications, (shown as ) others contain models ()--often the input to simulation applications, others contain trajectories -- the output of simulation applications, and others contain yet different data. Simbios's simulation toolkit, shown on the lower left, also resides on simtk.org. The yellow framed part of the SimTK toolkit can be downloaded as one package from the project:simtk.org/home/simtkcore. Simbios's main applications (shown above center) are built upon and use the Simulation Toolkit, others do not. SimVascular uses commercial components for it's toolkit, shown in the lower

# Simtk.org

---

- Original purpose: necessary infrastructure to support Simbios' code development and distribution of code and data
- Grew to host software and data from other researchers



# Outline

---

- Introduction to Simbios
- Using Simtk.org to share software and data
- Challenges in promoting open science

Simtk.org =  
collection of  
projects

Each project has  
its own set of  
customizable  
web pages

Statistics:  
290 projects

The screenshot shows a web browser window displaying the SimTK website search results. The browser's address bar shows the URL: [https://simtk.org/search/?type\\_of\\_search=soft&words=&Search=+Go+](https://simtk.org/search/?type_of_search=soft&words=&Search=+Go+). The website header includes navigation links for Home, About SimTK, and How to Contribute, along with a search bar and a 'Go' button. The main content area is titled 'Search Results: (refine with categories)' and shows 'No Keyword; Current Categories: All Categories;'. On the left, there are sections for 'Categories' (listing 'You searched' with 'No Keyword' and 'All Categories'), 'Refine Search' (with options to restrict by DBP, Download Content, and Applications), and 'Restrict by Download Content' (listing 'Applications', 'Computational Tools and Libraries', 'Data Sets', 'Developer Tools', 'Models', 'Publicly Available Downloads', 'SimTK Components', and 'Web Site'). The main search results are organized into three columns: 'Projects', 'People', and 'People'. The 'Projects' column lists 'OpenSim' (a software for modeling, simulating, controlling, and analyzing the neuromusculoskeletal system), 'SimTKcore' (a separate project providing 'one stop shopping' for SimTK Core software and support), 'Neuromuscular Models Library' (a set of models developed by the Neuromuscular Biomechanics Lab at Stanford), and 'LAPACK linear algebra library (with BLAS)' (high quality, high performance numerical methods for matrix arithmetic and solving linear algebra problems). The 'People' column lists several individuals affiliated with projects, including Adam Beberg, Ahmet Erdemir, Ajay Seth, AJ Rader, Alain Laederach, Alan Yen, Alberto Figueroa, Alena Shmygelska, Alex TenEyck, Allison Arnold, Amanda Knutson, Andres Kriete, Anita Vasavada, Ankur Dhanik, An Nguyen, Antal Novak, Anthony Sherbondy, Archana Sangole, Austin Mosley, Ayman Habib, Balaji Srinivasan, and Bart Adams. The browser window also shows various toolbars and a status bar at the bottom indicating 'Internet' and '100%' zoom.



# Simtk.org Features

---

- Established infrastructure
  - Polished look
  - Collection of statistics of usage
  - Tools for supporting users
    - User forum
    - Mailing lists
    - Wiki
  - Tools for supporting developers
    - svn code repository
    - Nightly builds & dashboards
    - Backups
- Form-based data entry
- Few requirements to establish a project

SimTK - the Simulation Toolkit, part of the Simbios project - Windows Internet Explorer

https://simtk.org/register/projectinfo.php

File Edit View Favorites Tools Help

Google G Go 96 blocked Check Autolink Autofill Settings SnagIt

SimTK - the Simulation Toolkit, part of the Simbios proj...

---

#### 4. Project Title

The title will be displayed in all search results and should be short and descriptive.  
Restrictions: please limit to 80 characters.

*Samples:*

- 1) ToRNADo RNA 3D Structure Morphing and Visualization Application  
(short acronym plus short descriptive text)
- 2) Myosin Resource Directory  
(just descriptive)
- 3) Cardiovascular Model Repository

**Your Project Title:**

---

#### 5. Primary Content of your SimTK project

(check all that apply)

- Applications
- Computational Tools and Libraries
- Data Sets
- Developer Tools
- Miscellaneous
- Models
- SimTK Components
- Web Site

---

#### 6. Is your project related to any of our driving biological problems?

(check all that apply):

- Cardiovascular Dynamics
- Myosin Dynamics
- Neuromuscular Biomechanics
- Protein Folding
- RNA Folding

---

#### 7. Short Project Purpose (for display in search results)

Please give a brief description of the purpose of your project. This information will be displayed in search results

# Simtk.org Features

---

- Established infrastructure
  - Polished look
  - Collection of statistics of usage
  - Tools for supporting users
    - User forum
    - Mailing lists
    - Wiki
  - Tools for supporting developers
    - svn code repository
    - Nightly builds & dashboards
    - Backups
- Form-based data entry
- Few requirements to establish a project

# Publication Projects on Simtk.org

The screenshot shows the Simtk.org website interface. At the top, there is a navigation bar with links for Home, About SimTK, and How to Contribute. A search bar is present with the text 'Search Simtk.org' and a 'Go' button. On the right side of the navigation bar, there are links for Joy Ku, News, Log Out, Create Project, and My Page. The main content area features a large 'PROJECT' watermark. On the left, there is a sidebar with a navigation menu including Overview, Team, Downloads, Documents, Publications, and Advanced. Below this menu is a section for 'Downloads & Source Code' with a list of download links (GIL01 through GIL12). The main content area displays the project title 'Multiple Speed Walking Simulations' and a 'Publication Overview' section. This section includes a citation: 'Liu, M.Q., Anderson, F.C., Schwartz, M.H. and Delp, S.L. Muscle contributions to support and progression over a range of walking speeds. Journal of Biomechanics. (2008) [Publication](#)'. Below the citation is a 'Description' section with text and four 3D skeletal diagrams of a human figure in various walking poses, labeled 'very slow', 'slow', 'free', and 'fast'. To the right of the description is a 'Project Lead' section with two photos and names: Scott Delp and Chand John, each with a 'Contact' link. Below this are two placeholder icons for Melanie Fox and May Liu, also with 'Contact' links. At the bottom of the page, there is a 'Driving Biological Problems' section with the text 'This project is part of Neuromuscular'. A 'Purpose/Synopsis' section at the very bottom provides a brief overview of the project's database.

citation of publication

description and listing of downloads available



# Simtk.org Features

---

- Established infrastructure
- Forms for entering data
- Limited requirements to establish a project
- Community base
  - Current statistic: ~ 3300 members

# Outline

---

- Introduction to Simbios
- Using Simtk.org to share software and data
- Challenges in promoting open science

# Challenges: Perception of Sharing

---

- Perception that sharing is not beneficial for a researcher
  - Could compromise a researcher's competitive advantage
  - Time and resources involved in sharing with no guaranteed or immediate return on investment



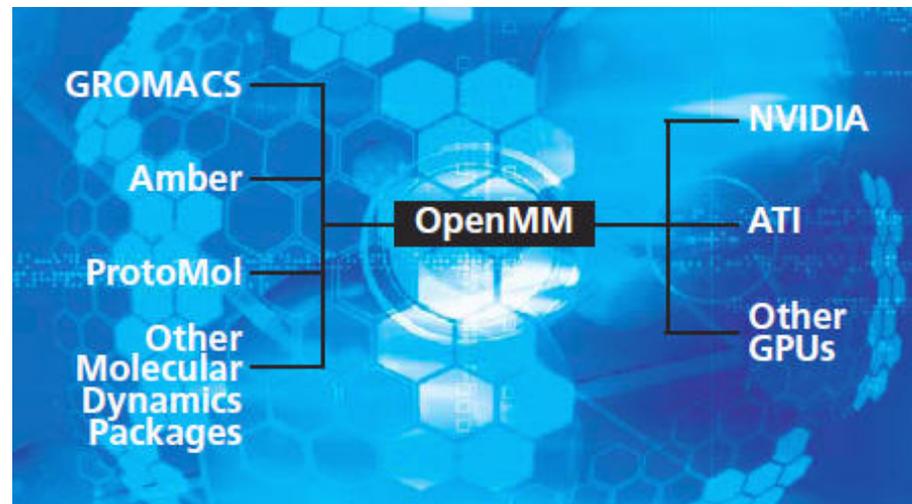
# Challenges: Rewards for Software Development?

---

- Software development is costly; dissemination, maintenance and support even more so
- Devoting time to computational infrastructure projects leads to an initial decrease in concrete research outputs
- Difficult to get funding to support computational infrastructure projects
  - PAR-05-057: Continued Development and Maintenance of Software

# Example: OpenMM

Enables molecular modeling simulations to take advantage of hardware acceleration (e.g., with graphics processing units or GPUs)



What is the likelihood of an individual creating something like OpenMM without specific funding to do so?

# Other Challenges for Open Science

---

- Proprietary rights: Increasing trend of universities to generate income through technology transfer and commercialization
- Patient Data and IRB

# What Is Needed: **Effective** Sharing of Research Software, Models, and Data

---

To accelerate progress in the field, there needs to be minimal re-invention

Researchers need to be willing to provide research results in a useful format



Researchers need to be willing to use the software, models, etc. that have been made available

# Challenges to Attracting Users

---

- Users need a level of confidence in tool, model, and/or data set
- Standards may be required
  - Facilitate sharing
  - Ease of finding needed tool or data

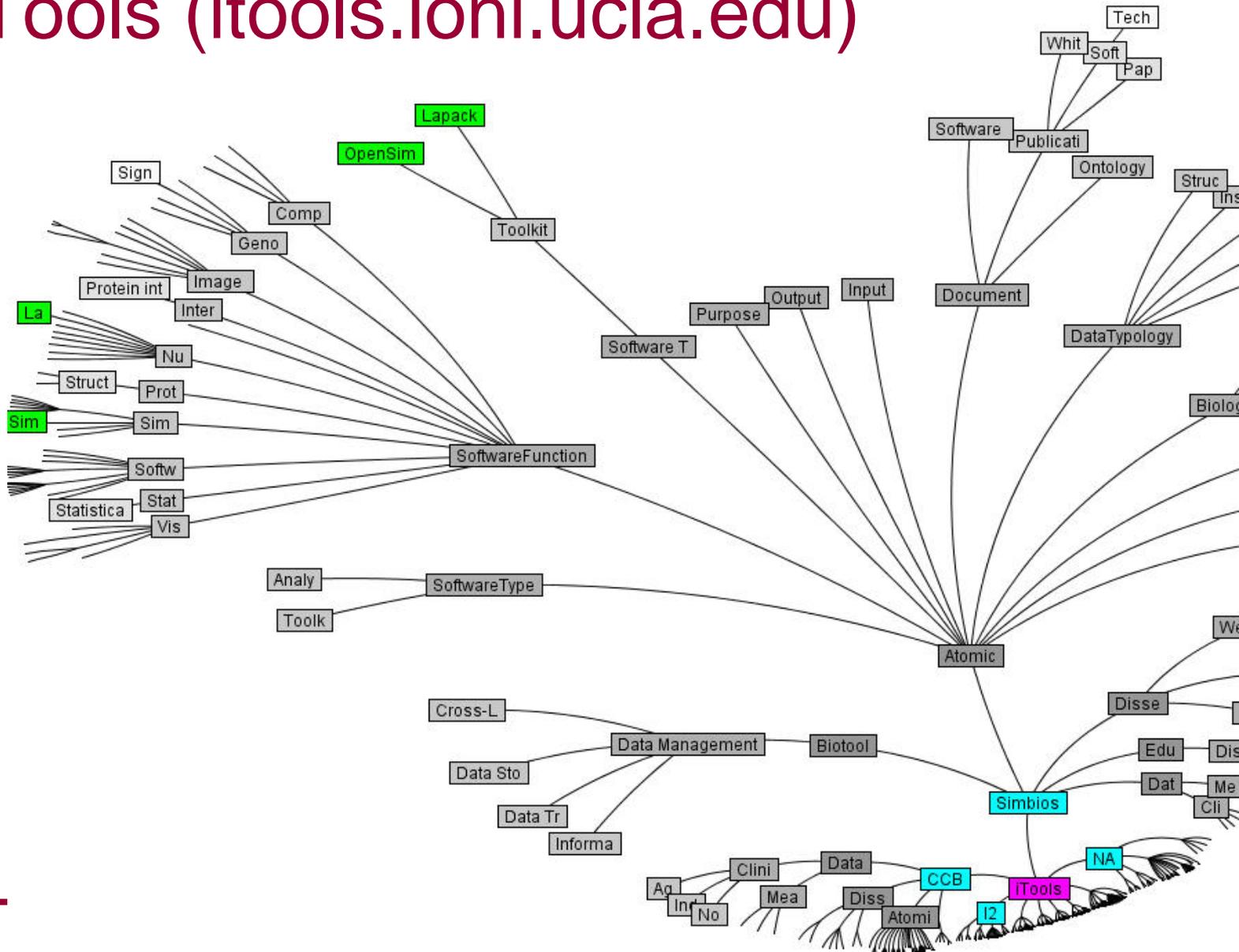
# Biositemaps ([ncbcs.org/biositemaps](http://ncbcs.org/biositemaps))

---

```
<desc:Resource_Properties rdf:ID="OpenMM">
  <desc:URL>https://simtk.org/home/openmm</desc:URL>
  <desc:license>MIT License</desc:license>
  <desc:resource_name>OpenMM</desc:resource_name>
  <desc:description>Application programming interface that allows molecular modeling
  <desc:contact>Vijay Pande</desc:contact>
  <desc:keywords>RNA folding, protein folding, molecular dynamics, Gromacs, GPUs</des
  <desc:center>Simbios Center</desc:center>
  <desc:organization>Stanford University</desc:organization>
  <desc:resource_type>
    <BRO:Molecular_Dynamics/>
  </desc:resource_type>
</desc:Resource_Properties>
```

- RDF file to describe biocomputational tools
- Generated by software developer
- Tools, like iTools, can use the biositemap files available on the Web to allow people to search for software and data

# iTools (itools.loni.ucla.edu)



# Biositemaps ([ncbcs.org/biositemaps](http://ncbcs.org/biositemaps))

---

```
<desc:Resource_Properties rdf:ID="OpenMM">
  <desc:URL>https://simtk.org/home/openmm</desc:URL>
  <desc:license>MIT License</desc:license>
  <desc:resource_name>OpenMM</desc:resource_name>
  <desc:description>Application programming interface that allows molecular modeling
  <desc:contact>Vijay Pande</desc:contact>
  <desc:keywords>RNA folding, protein folding, molecular dynamics, Gromacs, GPUs</des
  <desc:center>Simbios Center</desc:center>
  <desc:organization>Stanford University</desc:organization>
  <desc:resource_type>
    <BRO:Molecular_Dynamics/>
  </desc:resource_type>
</desc:Resource_Properties>
```

- RDF file to describe biocomputational tools
- Generated by software developer
- Tools, like iTools, can use the biositemap files available on the Web to allow people to search for software and data

# The Discussion Continues...

---

- National centers, like Simbios, can provide some of the tools needed by biocomputational researchers – these are publicly available
- To promote open science among the research community
  - Simtk.org provides the *means* for sharing research tools and data
  - Ultimately, successful sharing requires a **motivated** researcher
    - External incentives?
    - Role of Simtk.org and other repositories?