

The NIF LinkOut Broker: A Web Resource to Facilitate Federated Data Integration using NCBI Identifiers

Luis Marengo · Giorgio A. Ascoli ·
Maryann E. Martone · Gordon M. Shepherd ·
Perry L. Miller

Published online: 31 October 2008

© The Author(s) 2008. This article is published with open access at Springerlink.com

Abstract This paper describes the NIF LinkOut Broker (NLB) that has been built as part of the Neuroscience Information Framework (NIF) project. The NLB is designed to coordinate the assembly of links to neuroscience information items (e.g., experimental data, knowledge bases, and software tools) that are (1) accessible via the Web, and (2) related to entries in the National Center for Biotechnology Information's (NCBI's) Entrez system. The NLB collects these links from each resource and passes them to the NCBI which incorporates them into its Entrez LinkOut service. In this way, an Entrez user looking at a specific Entrez entry can LinkOut directly to related neuroscience information. The information stored in the

NLB can also be utilized in other ways. A second approach, which is operational on a pilot basis, is for the NLB Web server to create dynamically its own Web page of LinkOut links for each NCBI identifier in the NLB database. This approach can allow other resources (in addition to the NCBI Entrez) to LinkOut to related neuroscience information. The paper describes the current NLB system and discusses certain design issues that arose during its implementation.

Keywords Data integration · Entrez LinkOut · Gateway · GUID · Neurodatabases · PubMed

L. Marengo (✉) · P. L. Miller
Department of Anesthesiology, Center for Medical Informatics,
Yale University School of Medicine,
New Haven, CT 06520-8009, USA
e-mail: luis.marengo@yale.edu

G. M. Shepherd
Department of Neurobiology,
Yale University School of Medicine,
New Haven, CT 06520, USA

P. L. Miller
Department of Molecular, Cellular, and Developmental Biology,
Yale University School of Medicine,
New Haven, CT 06520, USA

G. A. Ascoli
Center for Neural Informatics, Structure, and Plasticity,
Molecular Neuroscience Department,
Krasnow Institute for Advanced Study, George Mason University,
Fairfax, VA 22030, USA

M. E. Martone
Department of Neurosciences, University of California,
San Diego, CA 92092, USA

Introduction

This paper describes a LinkOut broker that has been built as part of the multi-institutional Neuroscience Information Framework (NIF) project (Gardner et al. 2008), an NIH Neuroscience Blueprint initiative. The goal of the NIF LinkOut Broker (NLB) is to coordinate the assembly of neuroscience information accessible via the Web (e.g., experimental data, knowledge bases, and software tools) that can be linked to entries in the National Center for Biotechnology Information's (NCBI's) Entrez system. Each link is associated with an NCBI identifier (ID) which uniquely identifies the Entrez object (e.g., publication, gene, protein) to which it relates.

The NCBI's collection of health sciences databases in Entrez represents a great resource for the biosciences. Their LinkOut service (Schott 2004; NCBI 2007) allows a user to link directly from NCBI entries to information outside of Entrez. For this information to be accessible from Entrez, each external resource must provide the NCBI with a list of links to items within that resource that relate to entries (e.g.,

publications, genes, etc.) in Entrez. When a user browsing Entrez finds an entry of interest (e.g., a PubMed publication) for which external information is available, the Entrez LinkOut service presents links that allow the user to access that information directly. Externally linked information may include full text articles, datasets, high quality images, tools, etc.

While working on the problem of data interoperability within the neurosciences, members of the Interoperability Subcommittee of the Society for Neuroscience Neuroinformatics Committee foresaw the value of leveraging this type of LinkOut service using a brokered approach. This approach has been implemented in the NLB. In the brokered approach, individual neuroscience resources do not need to interact directly with the NCBI's LinkOut personnel. Rather, the NLB system collects links from a set of neuroscience resources, consolidates those links, and submits them to Entrez. This approach provides the following advantages.

- *Coordination by personnel familiar with the field* In the absence of the broker, the NCBI must verify the authenticity and value added by a new resource that wishes to provide external links. NIF project members are familiar with a range of neuroscience resources, and can therefore facilitate the process of identifying useful resources to include, as well as the integration of the information in those resources into the LinkOut approach.
- *Better organized maintenance of the data relationships* Having a single curated database of links for neuroscience data facilitates the collection and maintenance of this information and its inter-relationships. It facilitates both the process of passing links to the NCBI and the task of maintaining all the links in a organized fashion.

The NLB collects links to information in various neuroscience resources as described above, and stores those links in a database. Each link stored in the NLB database is associated with an NCBI ID, indicating the specific Entrez entry to which it relates. This NLB database can be utilized in several different ways.

1. One approach is to simply forward information from the NLB database to the NCBI, so that it can be incorporated into the Entrez LinkOut service, thereby allowing Entrez users to LinkOut directly to each item. This approach is currently operational.
2. A second approach, which is currently operational on a pilot basis, is for the NLB Web server, upon request, to create dynamically its own Web page of LinkOut links for each NCBI ID in its database. Linking out from the NCBI would involve first linking to a Web page of neuroscience links constructed by the NLB. From there

the user could link to each resource as desired. This approach involves one extra level of linking compared to the direct NCBI LinkOut described in (1) above. One potential advantage of this approach is that the NLB team would be responsible for organizing and maintaining the collection of neuroscience links, which is quite dynamic, reflecting the rapid growth and evolution of the field.

If the NCBI utilized this approach, for example, it would only need to maintain a list of NCBI IDs for which neuroscience links were stored in the NLB. It would not need to maintain a list of the links themselves.

3. Both of the above approaches could also be used to link *from* other resources (i.e., not just from the NCBI) to related neuroscience information. For example, two SenseLab (Shepherd et al. 1997; Miller et al. 2001) databases (NeuronDB and ModelDB) could be linked, in the context of data they display to their users, to related information in other neuroscience databases using the LinkOut approach. For this purpose, it would be particularly helpful for NLB to provide a dynamic page of links for each relevant NCBI ID, as outlined in option (2) above. Indeed, this general approach could help integrate a great deal of neuroscience data in a centrally organized fashion.

NLB System Overview

The NIF initiative as a whole is derived in part from an earlier project, sponsored by the Society for Neuroscience, to create a searchable database of neuroscience databases, the Neuroscience Database Gateway (NDG; <http://ndg.sfn.org>). Using the NDG as a test bed, an initial version of the LinkOut Broker was developed. The current NLB builds directly on this previous work.

Figure 1 illustrates schematically the various ways in which the NLB can be used (these capabilities are described in more detail in the next section of the paper).

1. Data are collected from a set of federated resources and stored in the NLB repository using *disco.ndg* messages encoded in XML (as described below). The appropriate set of NIF databases to be included is discovered by querying the NIF Database Registry catalog using an specific BrainML (BrainML.org 2008) query (<http://soma.med.yale.edu:8080/lb/nifcat.do>). Once these resources have been identified, NLB uploads the LinkOut data stored at each site.
2. All data collected as described above is submitted to the Entrez LinkOut system on a regular basis via FTP, encoded in an XML format specified by Entrez. Entrez in turn uses this information to create Web pages with

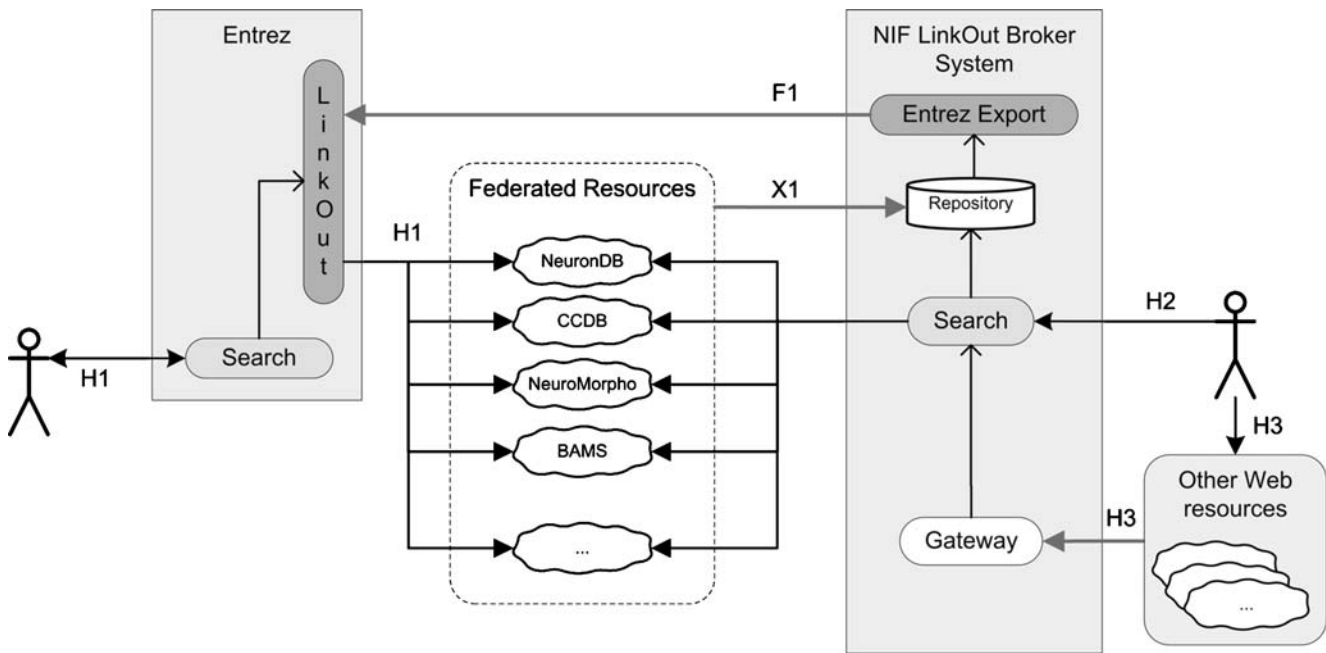


Fig. 1 NLB functional architecture. The NLB imports LinkOut data from federated resources using the *ndg.disco* protocol (*X1*), and stores that information in the NLB’s database repository. Entrez LinkOut data is exported to NCBI via FTP (*F1*). An NCBI user searching Entrez data uses the imported NLB information to navigate directly to individual NLB resources (*H1*). Alternatively, an NLB user can search

NLB federated data and navigate to individual NLB resources (*H2*). In addition, users of other Web resources can be connected to NLB’s gateway page to access NLB resources (*H3*). (The initial letters *F*, *X*, and *H* in the *arrow names* correspond to the technology used: FTP, XML and HTML respectively)

- hyperlinks to the resources containing related data for each relevant Entrez entry.
- 3. The NLB information can also be accessed *from* other Web resources via NLB Gateway pages, which are created dynamically for each NCBI ID as described below.
- 4. In addition, a pilot search interface allows any user to access the NLB server directly to search for links on topics of interest.

The underlying architecture of the NLB system includes the following components.

- The core system contains a Web server application, a database, and Web services. The server contains (a) import modules to retrieve data from NIF resources, (b) an export module to send data to NCBI, (c) a NLB gateway interface, (d) a NLB search interface, and (e) an administrative interface. Web server code and Web services are implemented in the Java language and run on a Tomcat Web server. The system uses a MySQL database as a backend.
- External servers include the NIF Registry (used to discover the NIF resources supporting LinkOut) and the Entrez LinkOut servers (FTP machines used to upload the NLB data).

- A set communication protocols (described later) link these components operationally.

The current NLB is available at <http://soma.med.yale.edu:8080/lb>.

Using the NLB

This section describes operational aspects of the NLB by explaining how three different types of users interact with the system: regular (information seeking) users, resource developers, and the NLB administrator

Information Seeking Users

Regular (information seeking) users may utilize the NLB in three ways.

1. One method involves using the brokered data sent via the NLB to NCBI Entrez. For example (see Fig. 2), a user interested in the data related to a PubMed publication first locates that article within Entrez (e.g. “Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites,” PubMed ID 11731556). From the main Entrez page of that article, the user follows the LinkOut hyperlink to an Entrez

The screenshot shows the PubMed search results page for the article "Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites" by Golding NL, Kath WL, and Spruston N. The article is listed as "1: J Neurophysiol. 2001 Dec;86(6):2998-3010." The page includes a search bar, navigation tabs (Limits, Preview/Index, History, Clipboard, Details), and a "Links" pop-up menu with options like "Compound (MeSH Keyword)", "Substance (MeSH Keyword)", "Cited in PMC", and "LinkOut". A "Related Links" section on the right lists related articles, such as "Dendritic voltage-gated ion channels regulate the action potential firing mode of [J Neurophysiol. 1999]".

The screenshot shows the LinkOut resources for the article "Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites" (PMID: 11731556). The page displays a message: "The following LinkOut resources are supplied by external providers. These providers are responsible for maintaining the links. What does the icon mean?" Below this, there is a section titled "Miscellaneous" containing a "FREE Neuroscience Information Framework" box. This box lists several resources: "ModelDB: Computational model", "NeuronDB: Neuronal properties", and "neuromorpho.org: Neuronal reconstruction" with various identifiers like (050103b), (060404a zd), (060804a zd), (061004c zd), (062204b zd), and (072903a).

Fig. 2 The PubMed page for the “Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites” paper (PubMed ID 11731556). By following the LinkOut link (from the “Links” pop-up menu), the user is taken to a second page (shown below). Here data

sent by the NLB is grouped within the Neuroscience Information Framework box. Each of these links leads to a resource Web page related to this paper: NeuronDB neuronal property data, a ModelDB computational model, and Neuromorpho.org neuronal reconstructions

page that displays external LinkOut resources for that article. Among these hyperlinks we find a group of links grouped under the title “Neuroscience Information Framework.” These include a link to a computational model in ModelDB, a link to neuronal property data in

NeuronDB, and several links to neuronal reconstructions in neuromorpho.org.

2. The second method involves connecting to a Web page at the NLB gateway from a neuroscience resource that has implemented the ability to link to the NLB. Each

NCBI ID links to a different, dynamically generated page. This approach is illustrated in Fig. 3 and is explained in detail in the next section.

- The third method involves using the NLB's search interface to find NLB links related to a specified topic. For example, Fig. 4 shows the results of a search for “low-threshold calcium currents” that returns one link to neuronal property data in NeuronDB, one link to a computational model in ModelDB, and several links to neuronal reconstructions in neuromorpho.org.

Resource Developers

Resource developers (the personnel in charge of maintaining each resource) may incorporate NLB's functionality into their applications by implementing a simple protocol that we have named “ndg.disco”. This simple format provides explicit information about the data links in their Web resources. The current format specification is defined at <http://ndg.sfn.org/interop/protocols/disco/versions/v2/disco.xsd>. Figure 5 shows an example of what an ndg.disco file looks like, and also shows how that information is incorporated into a resource to allow use of NLB. The main ndg.disco file is called “disco.xml” and is stored on the root directory of the resource.

Developers of resources who are interested in sharing their Entrez-related data via the NLB also need to

implement a Web feed (implemented as a file or Web script) for their LinkOut data. Some examples of LinkOut Web feeds are http://ccdb.ucsd.edu/disco_entrez_objects.xml (file) and http://senselab.med.yale.edu/NeuronDB/disco_entrez_objects.asp (Web script).

The ndg.disco protocols represent an example of automated resource registration and interoperability protocols which are described in more detail in a companion paper (Gupta et al. 2008) in this issue.

The amount of effort needed to implement LinkOut interoperability depends on the complexity of extracting the LinkOut data from the resource and converting it to the ndg.disco XML format. Resource developers interested in allowing Entrez users to link to their resources using NLB should contact NIF staff for guidance.

Another way that developers connect to NLB is using the NLB gateway. This procedure requires that a neuroscience resource has implemented the ability to link to the NLB. Figure 3 illustrates this capability for a PubMed ID. To display this page, the gateway uses the following Web link format: <http://soma.med.yale.edu:8080/lb/gateway.do?id=PubMed|11731556>

Another example of using the gateway can be seen using the AF1209005 ID for the “Mus musculus VNO olfactory cluster” in the following link: ([http://soma.med.yale.edu:8080/lb/gateway.do?id=Nucleotide|AF129005\[pacc\]](http://soma.med.yale.edu:8080/lb/gateway.do?id=Nucleotide|AF129005[pacc])). The resulting gateway page shows links to three olfactory receptors in the Olfactory Receptor Database.

Fig. 3 The NLB gateway Web page that is generated for same PubMed ID seen in Fig. 2. Data links are grouped by categories and resources. This extra “metadata” that groups these links into

neuroscience-related categories is not available in the Entrez LinkOut links



NIF
Neuroscience Information Framework

LinkOut Broker - Search Interface

[main](#) [help](#)

low-threshold calcium currents Search Cancel

Search results

Resource	Information	Category	Data	NCBI
ModelDB	Low Threshold Calcium Currents in TC cells (Destexhe et al 1998)	Computational model	Data	9570789
neuromorpho.org	HomePage	Neuronal reconstruction	Data	9570789
neuromorpho.org	tc2CO	Neuronal reconstruction	Data	9570789
neuromorpho.org	tcD[1]	Neuronal reconstruction	Data	9570789
NeuronDB	Dendritic Low-Threshold Calcium Currents in Thalamic Relay Cells	Neuronal properties	Data	9570789

Fig. 4 The NLB search page returning NIF resources with links related to “low-threshold calcium currents.” The “Data” link takes the user to the resource data page and the “NCBI” link to its Entrez information. This search interface allows any user to search the NLB via the Web.

The user may search for simpler keywords such as “dendritic spines” or “Pyramidal cells”, but a more focused search was used in this example to allow results from multiple databases to be displayed in a single page

Linking to the gateway interface in this fashion is not restricted to users of NIF resources. Any application developer can incorporate linkage to NLB neuroscience information by creating a URL composed using the template as described below.

http://soma.med.yale.edu:8080/lb/gateway.do?id={entrez_db}{entrez_object_id}{entrez_query}

- Replace {entrez_db} with any of the commonly known Entrez database names: (e.g.: “PubMed”, “Nucleotide”, etc.). The list is available on the Entrez site.
- Replace either the {entrez_id} with the object_id given in the referred Entrez database, or the {entrez_query} with an Entrez query string. If both {entrez_id} and {entrez_query} are present, {entrez_query} will be ignored by the system. (Details as to how to construct an entrez_query can be found at the Entrez site.)

The NLB Administrator

The NLB also provides a Web-based management console that is available in read-only mode to the public (see Fig. 6). An authenticated administrator (a user with an administrator account that allows read/write access) can use this interface to update and coordinate the NLB contents.

In the main administrative page, clicking the update button first queries the NIF Database Registry to locate neuroscience resources that contain NLB links. The NLB then uses this list to query independently each of these resources to extract their Entrez LinkOut information (stored locally). This information is then stored in the NLB database and can be viewed from the administrative interface. In addition, the interface will generate LinkOut data for each resource which can be uploaded via FTP to NCBI’s LinkOut service. This process is currently performed manually upon request of the database owner. In future NLB versions, the LinkOut import process could be performed automatically, e.g., every 24 h. The upload process to NCBI could also be performed automatically if data has changed from a previous uploaded version.

Current Status

At the time of this publication the NLB is providing roughly 22,000 neuroscience links to Entrez users. These are distributed in the following databases:

- ORDB has 15,561 links ((Crasto et al. 2002); <http://senselab.med.yale.edu/ORDB>).

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
- <disco version="1.1" date="" about="http://ndg.sfn.org/interop/protocols/disco/#1.1">
  <resource_base_location url="http://ccdb.ucsd.edu/" />
  <technical_contact name="Willy WaiHo Wong" email="wawong@sdsc.edu" />
  <resource_description url="/resource_description.xml" format_type="ndg.source-description.1.0" />
  <db_schema url="/disco_dbschema.jsp" format_type="ndg.disco.schema.1" />
  <terminology url="/disco_lexicon.jsp" format_type="ndg.disco.lexicon.1" />
  <interoperability url="/index.html" format_type="GIDI" />
  <entrez_oid url="/disco_entrez_objects.xml" format_type="disco.entrez_oid.1" />
</disco>

- <disco version="1.1" about="http://ndg.sfn.org/interop/protocols/disco/#1.1">
- <entrez_oid_list>
  <oid db="PubMed" oid="15988042" linkcategory="Electron microscopy product"
    linkname="Development of a model for microphysiological simulations: small
    nodes of ranvier from peripheral nerves of mice reconstructed by electron
    tomography" linkurl="http://ccdb.ucsd.edu/sand/main?
    event=displayAllProjectProds&mpid=48&ptype=sproject" />
  <oid db="PubMed" oid="12794746" linkcategory="Light microscopy product"
    linkname="Examination of the relationship between astrocyte morphology and
    laminar boundaries in the molecular layer of adult dentate gyrus"
    linkurl="http://ccdb.ucsd.edu/sand/main?
    event=displayAllProjectProds&mpid=3572&ptype=sproject" />
  <oid db="PubMed" oid="11826104" linkcategory="Electron microscopy product"
    linkname="Ultrastructure of a somatic spine mat for nicotinic signaling in
    neurons" linkurl="http://ccdb.ucsd.edu/sand/main?
    event=displayAll&mpid=3620" />

```

Fig. 5 The top screen shows the *disco.xml* file for the Cell Centered Database (CCDB). *Highlighted* are portions of the file required to interoperate with the NLB. The lower screen shows a portion of

CCDB's LinkOut XML which illustrates how PubMed IDs are linked to URLs. This linkage information is stored in a file ("*./disco_entrez_objects.xml*") pointed to by the *disco.xml* file

- Neuromorpho.org has 4,697 links ((Ascoli et al. 2007); <http://neuromorpho.org/neuroMorpho>),
- Brain Architecture Management System has 326 links ((Bota et al. 2005); <http://brancusi.usc.edu/bkms>),
- Cell Centered Database (CCDB) has seven links ((Martone et al. 2008); <http://ccdb.ucsd.edu/CCDBWebSite>),
- Internet Brain Volume Database has 352 links (<http://www.cma.mgh.harvard.edu/ibvd>),
- ModelDB has 401 links ((Hines et al. 2004); <http://senselab.med.yale.edu/ModelDB>), and
- NeuronDB has 490 links (<http://senselab.med.yale.edu/NeuronDB>)

Discussion

It has become a common practice to create links to other databases using their unique identifiers. It is also common

to see in many bioscience databases an Entrez ID for publications, genes, proteins, etc. While links to PubMed are rather easy to maintain (since URL paths are relatively stable and forwarding scripts are created in PubMed to redirect old referencing URLs to new ones), links to objects in many other databases, particularly research databases, are considerably more likely to change over time.

Whenever a database changes its URL referencing scheme, all sites that have links to that database need to update their scripts. In addition, unless the resource creates a redirecting script, all the referencing sites will have their links to that database broken.

Using the Web-based NLB approach described in this paper, one can have a single referencing scheme. Any application need only include links to the NLB gateway, passing in just an NCBI ID. Only the NLB needs to keep track of any changes made by participating databases. In addition, the *ndg.disco* approach allows the local database developers to change their local *ndg.disco* file to reflect any

NIF
Neuroscience Information Framework

LinkOut Broker

Neuroscience Resource List

[main](#)

Synchronized resources from [NIF Registry](#) updated on [2008-07-28 18:52:39]

This information is automatically synchronized with the registry for those sources supporting disco protocol.

Resource Summary

Name	Protocol	Objects	Import Date	Status Note	Export to NCBI
NeuronDB	disco	490	2008-04-29 12:52:37		Export
CoCoDat	disco			No linkOut	
Cell Centered Database	disco	7	2008-04-17 13:20:41		Export
BrainInfo	disco			No linkOut	
ORDB	disco	15621	2008-04-24 13:47:49		Export
neuromorpho.org	disco	4697	2008-07-28 18:52:39		Export
ModelDB	disco	401	2008-04-17 13:21:44		Export
BAMS	disco	326	2008-04-17 13:22:00		Export
IBVD	disco	352	2008-04-17 13:22:15		Export

Fig. 6 The NLB administration console with nine neuroscience resources implementing automated resource interoperability (i.e., the *ndg.disco* protocol). Seven of these resources provide LinkOut data (see the “Status Note” column). For a given resource, selecting the “Name” link will display a summary of the resource. Selecting the “Objects” link will display all imported information from that

resource. Selecting the “Export to NCBI” link will generate the XML files that can be uploaded to the NCBI Entrez LinkOut. The Update button at the *top of the page* will retrieve new resource descriptions (new or updated *ndg.disco* files) from the NIF registry and update them all at once

changes made to their database. The NLB can import the new *ndg.disco* file, and ideally update its URL links automatically.

Another potential advantage of the NLB approach is that it could facilitate the use of back-up URLs. For example, if a resource’s Web site went down, one could readily instruct the NLB gateway to direct users to a back-up Web server for that resource.

It is also worth emphasizing that the NLB approach need not be limited to NCBI IDs. This approach has also recently been used to semantically annotate life science data using Life Science Identifiers (LSIDs) (Martin et al. 2005). (Indeed, NCBI IDs and LSIDs are examples of Globally Unique Identifiers (GUIDs), a mechanism used to uniquely identify digital information (Clark et al. 2004)).

In summary, the NLB approach can be applied flexibly in several ways. We believe that it provides an organized paradigm for linking neuroscience information in a fashion that could be of great service to the neuroscientist user. The approach was originally developed to allow LinkOut from

Entrez to neuroscience databases, but the same technique could be used to allow LinkOut in other bioscience domains. Using a dynamic Web-based NLB to provide a set of neuroscience links for each relevant NCBI ID provides an efficient approach to helping interlink a community of interrelated neuroscience resources.

Information Sharing Statement

The LinkOut broker application is freely accessible to the public at <http://soma.med.yale.edu:8080/lb>. The source code is also freely available. Please contact the first author.

Acknowledgments This project has been funded in whole or in part through the NIH Blueprint for Neuroscience Research with Federal funds from the National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN271200577531C. This research was also supported by *NIH grants P01 DC04732 and R01 DA021253, * Volunteer consultant-collaborators and friends, and * The Society for Neuroscience.

We would like to especially acknowledge the work done implementing the automated resource registration and interoperability protocols, including *ndg.disco*, by Mihail Bota at UCLA for the Brain Architecture Management System and by David Kennedy at the Massachusetts General Hospital for the Internet Brain Volume Database.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Ascoli, G. A., Donohue, D. E., & Halavi, M. (2007). NeuroMorpho. Org: a central resource for neuronal morphologies. *The Journal of Neuroscience*, 27, 9247–9251. doi:10.1523/JNEUROSCI.2055-07.2007.
- Bota, M., Dong, H. W., & Swanson, L. W. (2005). Brain architecture management system. *Neuroinformatics*, 3, 15–48. doi:10.1385/NI:3:1:015.
- Brain, M. L. (2008) Brain Markup Language. <http://brainml.org>
- Clark, T., Martin, S., & Liefeld, T. (2004). Globally distributed object identification for biological knowledgebases. *Briefings in Bioinformatics*, 5, 59–70. doi:10.1093/bib/5.1.59.
- Craστο, C., Marengo, L., Miller, P., & Shepherd, G. (2002). Olfactory Receptor Database: a metadata-driven automated population from sources of gene and protein sequences. *Nucleic Acids Research*, 30, 354–360. doi:10.1093/nar/30.1.354.
- Gardner, D., Akil, H., Ascoli, G. A., Bowden, D. M., Bug, W., Donohue, D. E., et al. (2008). The Neuroscience Information Framework: a data and knowledge environment for neuroscience. *Neuroinformatics*, this issue.
- Gupta, A., Bug, W., Marengo, L., Qian, X., Condit, C., Rangarajan, A., et al. (2008). Federated access to heterogeneous information resources in the Neuroscience Information Framework (NIF). *Neuroinformatics*, this issue.
- Hines, M. L., Morse, T., Migliore, M., Carnevale, N. T., & Shepherd, G. M. (2004). ModelDB: A Database to Support Computational Neuroscience. *Journal of Computational Neuroscience*, 17, 7–11. doi:10.1023/B:JCNS.0000023869.22017.2e.
- Martin, S., Hohman, M. M., & Liefeld, T. (2005). The impact of Life Science Identifier on informatics data. *Drug Discovery Today*, 10, 1566–1572. doi:10.1016/S1359-6446(05)03651-2.
- Martone, M. E., Tran, J., Wong, W. W., Sargis, J., Fong, L., Larson, S., et al. (2008). The cell centered database project: an update on building community resources for managing and sharing 3D imaging data. *Journal of Structural Biology*, 161, 220–231. doi:10.1016/j.jsb.2007.10.003.
- Miller, P. L., Nadkarni, P., Singer, M., Marengo, L., Hines, M., & Shepherd, G. (2001). Integration of multidisciplinary sensory data: a pilot model of the human brain project approach. *Journal of the American Medical Informatics Association*, 8, 34–48.
- NCBI. (2007) Entrez LinkOut Service. <http://www.ncbi.nlm.nih.gov/projects/linkout>.
- Schott, M. J. (2004). PubMed enhancements: fulfilling the promise of a great product. *Medical Reference Services Quarterly*, 23, 1–11. doi:10.1300/J115v23n04_01.
- Shepherd, G. M., Healy, M. D., Singer, M. S., Peterson, B. E., Mirsky, J. S., Wright, L., et al. (1997). Senselab: a project in multidisciplinary, multilevel sensory integration. In E. H. Koslow, & F. M. Huerta (Eds.), *Neuroinformatics: An Overview of the Human Brain Project* (pp. 21–56). New York: Erlbaum.