BEST PRACTICES IN SYSTEMS IMPLEMENTATION WORKSHOP

IRVINE, CA | OCTOBER 22-23, 2018

DAY 1

OVERVIEW:

The meeting began with introductions from each participant in the workshop and addresses from the four hosts: Michael Yassa, Stephanie Albin, Satrajit Ghosh, and Agnes McMahon.

Agnes McMahon and Stephanie Albin summarized the previous Inventory meeting. **Agnes McMahon** began by recounting the progress that has been made by the IBI since its inception. She explained that, in 2016, a grand challenge was posed to around 75 scientists around the world, with the ambition to address the "biggest challenges in brain research." After this group expanded to over 400 neuroscientists and a growing number of large-scale, nationally sponsored brain projects, a need for coordination was identified and the IBI began taking form. Much of 2017 was spent developing the main ideas behind the IBI, culminating with the Declaration of Intent announced at the Australian Academy of Science in Canberra. By May of 2018, the goals of the initiative were drafted and later refined in July. The spirit of the IBI is to move away from the "siloed era" of science and progress towards the four aspirational goals of the IBI:

1. To promote coordination and leadership
2. To transcend borders
3. To share and disseminate knowledge
4. To shape the future

A vast amount of effort was put into the IBI in 2018 to lay out a framework for which to accomplish these goals. Namely, an organizational structure was established, constituted by representatives from the seven projects across the world and the creation of the working groups: Data Standards and Sharing, Neuroethics, Inventory of Global Brain Projects, Education and Training, Communication and Outreach, and Tool and Technology Dissemination. McMahon stated that the urgency of the current meeting is a result of strong, positive momentum from the last meeting in July.

Stephanie Albin established the context and goals of the Inventory working group specifically. It was initially set up, with the interests of funding agencies in mind, to catalog

the activities of each national brain project to be analyzed for common research themes, niches, and gaps open to be filled. At its inception, the inventory was a single Excel spreadsheet with separate tabs for each participating country. Examples of column labels include: goals, timeframes, funding, and model organisms.

Later in 2017, those involved sought to make this resource useful to more groups that merely funders, with the hope that this shift would foster a higher level of involvement from the community. A workshop was set up in July of 2018 to identify the "blue sky" use cases for neuroscientists, funders, project leads, and stakeholders. Following this, a survey went out to all participants to gather information on the top use cases:

1. Identifying and maintaining collaboration
2. Resource awareness and training
3. Identifying potential research topics and areas of interest
4. Facilitating data use and reuse
5. Transparency in funding and funding opportunities

It was found that participants were most interested in finding collaborations, identifying tools (both published and unpublished), as well as locating training opportunities.

The workshop proceeded with a series of information blitzes on existing platforms and expertise. This was originally intended to take five minutes per person, but expanded to constitute most of the afternoon and took the place of the originally scheduled group discussion.

INFORMATION BLITZES

Note: links and descriptions of resources mentioned can be found at the end of this document.

PRESENTER: BASTIAN GRESHAKE-TZOVARAS (DATA SHARING UNDER A PARTICIPANT-CENTERED PARADIGM)

The first of the information blitzes centered around the management of personal data. Greshake-Tzovaras is Director of Research at Open Humans, a service that allows users to upload, connect, and privately store their genetic, activity tracking, and social media data. In an era when more people than ever before possess a wealth of arguably underutilized personal behavioral and genetic information, Open Humans crowdsources data collection and allows users to selectively join research projects by donating their de-identified data. Knowing that putting personal information in the public domain has its way of inciting uneasiness, Greshake-Tzovaras has found success by storing data on a private cloud platform

and requiring research to request access to datasets. Individuals can then decide whether to share with specific researchers, and complete the consent process according to the institutional requirements of the inquiring research group. He would be interested in applying his expertise to support neuroscientific research by potentially storing neuroimaging data. However, as Open Humans expands, he mentioned a need to develop their sustainability infrastructure and hopes that participating academic researchers would help fund storage and bandwidth in the future.

PRESENTER: JEFFREY S. GRETHE (MAINTAINING RESOURCE INVENTORIES)

Jeffrey Grethe is director of the Center for Research in Biological Systems at the University of California, San Diego, and is Principle Investigator for the NIDDK Information Network (dkNET). DkNET builds upon the Neuroscience Information Framework (NIF), an infrastructure previously developed, under the leadership of Dr. Grethe, to make resources and data discoverable and accessible. NIF and dkNET were developed at the request of the NIH to provide a simple unambiguous identifier for resources, reagents, cell lines, and tissue samples. This network is built around publishers and funders, and enables text mining of papers and grants. Thus far, grant information is limited to US-funded projects, but Dr. Grethe expressed that they would consider integrating international sources of funding. He also admitted that NIF has struggled with the unavoidable challenge of information curation. As biomedical information on the web is extremely fluid, maintenance of information has thus far been a community effort.

PRESENTER: SUNIL P. GANDHI (USING LAB COURSES TO SPEED DISSEMINATION OF BRAIN TECHNOLOGIES)

Over the last year, Dr. Gandhi has developed and improved upon iDISCO, a mediated clearing and light sheet fluorescence microscopy technique. As a result, he has received an overwhelming amount of interest from the scientific community and requests for collaboration. To meet this interest, Gandhi took an inventive approach: designing a laboratory course to spread methodological knowledge of iDISCO through the paths taken by graduate- and medical school-bound undergraduates. Graduate students volunteered their time to teach this technique, and the engagement was contagious. When interviewed afterwards, students expressed that they wanted to continue to learn more Python to render 3D images of cleared brains. Dr. Gandhi said that this experience taught him more about how people outside of his immediate scientific community consume data and considers this to be an important consideration when designing the inventory.

PRESENTER: DAVID KENNEDY (NITRC AND NEUROIMAGING: COMMUNITY BUILDING BLOCKS FOR AN IBI)

This information blitz presented an existing source of neuroimaging data and tools. The Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC) currently serves the neuroimaging community in three ways: as a place to find software or advertise software in development (resource registry), a platform for data sharing (image repository), and as a high-performance cloud-based computational environment. Neuroimaging can be considered a microcosm of IBI in the sense that it is a field with an abundance of data, a significant amount of which is shared. There are also numerous ways to do data analysis, as well as good software, data standards, and compute environments. Thus, the lessons learned from successes and challenges to NITRC may be applied during the development of the IBI inventory. Users of NITRC can find out what investigators are working on, search publications associated with a resource, and post to public forums. The architecture of NITRC is comprised of knowledge graphs, where projects (both software and datasets) are connected to grant information, people, and publications. All nodes on the knowledge graphs may also link out to sites such as NIF, PubMED, NIH Reporter, and GitHub. A challenge to NITRC has been the lack of centralized management of content, which is currently managed by the resource developers themselves. There is also a need to consider improvements upon the existing sustainability model. Although the NIH typically does not fund infrastructure, NITRC has been kept afloat through grants due to its importance to the neuroimaging community. Dr. Kennedy proposed that institutions or users might be able to help fund the dissemination of their work, but is concerned that this would heighten the barrier to contribute.

PRESENTER: SASKIA DE VRIES (BUILDING THE ALLEN BRAIN OBSERVATORY)

The Allen Brain Observatory is a physiological survey, using two-photon calcium imaging, of the mouse visual cortex developed with the mission to generate a useful public resource. The observatory includes around 65,000 neurons with 456 three-hour experiments from 245 mice. Putting this resource on the internet required a large effort towards optimizing its user interface. During its development, de Vries and the team at the Allen Institute for Brian Science (AIBS) pre-computed various metrics for different cells and created dynamic, intuitive visualizations to accompany their data. However, after its launch, it became clear that many users only engaged with the website to become acquainted with the available measures before downloading the SDK to perform their own analyses. If users are not using the website as expected, it becomes important to understand who the users are, how they've learned about the Brain Observatory, what they are trying to do with the data, and the extent to which they use or are aware of the resources provided to them. By contacting investigators listed in the dozens of papers published using AIBS data, they found that PIs were more

concerned with the quality of the data, while graduate students were more likely to seek out the SDK. Further, very few users found out about the Brain Observatory through the community and were much more likely to have had personal contact with a scientist at AIBS who directed them to the website. For help, users read documentation tutorials or asked their contact at AIBS directly, and were less likely to consult source code, GitHub, or the website. This presents a challenge to AIBS, as more work is needed to increase community awareness of the Brain Observatory and streamline online tutorials. Dr. de Vries suggested that, in building the IBI Inventory, it is important to create tools in line with the true needs of the scientific community, rather than what we, as organizers, *think* the community needs.

Dr. Gandhi responded to this information blitz with his experience of using the Brain Observatory in teaching. He mentioned that his undergraduates benefited from the user-friendliness of the website and emphasized the immeasurable role that inspiration plays in cultivating science. The AIBS website gave his students a visible example of a community and the activities therein.

Dr. Bernard emphasized the lesson that AIBS learned: one must understand why they are building a resource. The vision for the Brain Observatory was parallel to the vision that many scientists have for their own work: to learn, to publish, and to create visually appealing figures that do justice to their findings. This is what AIBS assumed the community wanted. But the field is constantly changing, and it's important to continuously ask oneself, "who is your community?" and step outside of one's own circle. It is important to understand the accessibility and relevance of your data. AIBS also learned that, despite not offering educational outreach or customer support, the community expects that level of support. Lacking that, a disconnect with users is created. She stressed that in some cases, being a successful nonprofit aggregator of data is not enough.

PRESENTER: CINDY DANIELSON (EXPERIENCES COORDINATING NIH DATABASE TOOLS THAT INTEGRATE MULTIPLE TYPES OF INFORMATION FROM MULTIPLE FUNDERS (NIH REPORTER, FEDERAL REPORTER, AND WORLD REPORT))

Dr. Danielson walked the workshop participants through the three NIH database tools. NIH RePORTER pulls together NIH expenditures and results of funding activity. Federal RePORTER expands upon the NIH RePORTER by adding data from other federal funders of science. There are 18 agencies consulted in Federal RePORTER, capturing around 980,000 projects. For example, USDA, DOD, ED, EPA, HHS, NASA, NSF, and VA are all considered in these searches. In World RePORT, 12 worldwide funders and 295,000 projects (both public and private) are considered. Consequentially, World RePORT is the most heterogeneous of the three tools. The biggest challenges to developing these tools has been the manual effort involved with integrating these data, especially when naming conventions are inconsistent.

Dr. Danielson suggests considering whether your resource automatically retrieves information from other databases and to automate as much as possible. For international projects, make sure the IBI inventory can accommodate information in multiple languages. She also clarified that Federal RePORTER has an API and one for NIH RePORTER is in development. World RePORTER can be downloaded as a spreadsheet, but lacks an API.

PRESENTER: STEPHEN LARSON (EXPERIENCES IN SYSTEMS IMPLEMENTATION IN NEUROSCIENCE INFORMATION (NIF, INCF, OPENWORM, METACELL))

As founder of OpenWorm, Dr. Larson has experience with building organizations from the ground-up. Much of the process, for him, has been in leveraging free resources (such as YouTube, Slack, and Google Sheets) to create conversation without spending resources. Larson broke down his recommendations to the workshop into three major areas: sociological, technical, and operational. He placed the heaviest importance on the sociological component, due to the tendency of data aggregation to be a tragedy of the commons. While many in the scientific community would benefit from open data sharing, the diversity in the ways that people have been wanting to solve this deficiency in cross-talk has been unproductive at best or counterproductive at worst. To solve the sociological challenges, Dr. Larson suggested that the community needs to recognize the point at which a true, engineering mindset is needed and focus singularly on maintaining the clarity of the problem to be solved. The technical needs of the project are in establishing widespread standardization. Finally, operationally, existing resources should be utilized. He left us by posing a literary metaphor: does the IBI want to create a "newspaper" that is issued on a regular basis, in the hands of people quickly, and contains less important, but dynamic, information? Does the IBI want to create a bookstore, that houses substantial and meticulously edited resources, that only leave the store once? Or does the IBI want to create a library, that focuses on archiving and creating comprehensive, historical weight behind the information it contains?

PRESENTER: SEAN HILL (TOWARDS A GLOBAL BRAIN NEXUS)

Dr. Hill is co-director of Blue Brain, a project that seeks to digitally reconstruct and simulate the rodent brain with the ultimate goal of doing the same for humans. He challenged the workshop participants to imagine what a collaborative and searchable brain atlas of this nature would look like. This is an extensive undertaking, not least of all because data coming from different sources would lead to heterogeneous data formats and a lack of standardized ontologies. Organizing this information in a relational database, like Excel, works in some circumstances, but managing versions is unsustainable, as it requires the developer to recreate it to accommodate a new data source. An alternative is to represent all of the

information in a knowledge graph, like that of Google's Data Set Search, that indexes individual data sources rather than storing it. It enables sophisticated queries and navigation of heterogenous data. Google has also created a website for publishing their schemas for the metadata of their datasets (schema.org). Embedding these schemas into a website will index the website in Google's knowledge graph. Dr. Hill designed Nexus to be a neuroscience-centered software similar to Google's that ensures greater accessibility through an API, provides visualizations, and is searchable.

David Keator added that the INCF metadata working group has been meeting to discuss efforts in Canada and the US to define federated data elements. In general, the community has been looking for more top-down incentivisation by the NIH and the NSF to make data available, but those institutions are currently unable to enforce it.

Satrajit Ghosh offered that the reason that Google Schemas works as well as it does is because it is simple, but also very difficult to extend by adding new terms. Neuroscience, as an expanding domain with new tools being introduced constantly, cannot be captured by a schema that remains static as Google's does.

PRESENTER: MICHAEL GRAUER (MANAGEMENT AND HOSTING OF HETEROGENEOUS RESEARCH DATA)

As a software developer working on Girder, an open-source service for web-based management of heterogeneous data, Michael Grauer has experienced the difficulties of data curation. Girder also shares a mission similar to that of the IBI inventory, which is to enable scientific reproducibility. However, rather indexing data from disparate sources, Girder contributes an environment to host both files and metadata, allowing users to download large files from their browser. Having files and metadata in the same place also allows for linking visualizations of data and integrate with data analytics. Grauer has been working with Neurodata Without Borders and the Allen Instutute to build more of such visualization and analytics tools. Through this, Grauer has learned to focus on flexibility and sustainability of the archive, but has encountered similar difficulties to those discussed; namely, the lack of consensus for terms in metadata.

PRESENTER: DAVID KEATOR (SEMANTIC ANNOTATION OF EXPERIMENT DESCRIPTIONS WITH THE NEUROIMAGING DATA MODEL (NIDM))

NIDM has been a grassroots effort between European, Canadian, and US labs to capture the entire lifecycle of neuroimaging datasets with a single core ontology. The technique underlying this is that of a semantic web that makes use of NIDM core vocabulary and various specifications found in experiment descriptions and metadata (such as what study groups

exist, what protocols are used, what data are being collected, and what analyses were performed on the data to arrive at some results). Again, one challenge is that terminology is far from being consistent. The information is organized such that locally stored information, in the form of files in a local, native format) can be accessed via a tool that can extract variables in the metadata. NIDM is currently working on a series of API queries to, for example, extract all of the ages of human subjects in a sample. A future direction is to look at how terms represent data in different countries, such as between the US and Australia. Ultimately, a comprehensive list of neuroscience terms may be required. Dr. Keator suggested dynamic lexicons such as Neurolex or its successor, Interlex.

PRESENTER: ELIZABETH WU (LESSONS LEARNED IN BUILDING A DISEASE-FOCUSED INFORMATION AND KNOWLEDGE RESOURCE FOR THE GLOBAL RESEARCH COMMUNITY)

As Director of Innovation Development for Alzforum, Elizabeth Wu has experience with building an online community from scratch for the purposes of accelerating scientific discovery. Built in 1996, Alzforum has a clear mission with an identified target audience: researchers. The staff currently is comprised of 10 employees, all of whom work remotely. Besides Wu, who manages the programming and interface of the site, Alzforum employs one programmer, one curator, an executive editor, a managing editor, four newswriters, and one part-time copy editor. By employing newswriters, Alzforum can elaborate on recent publications by putting new research into context and offering explanations of its impact. Through incorporating a searchable database atop numerous newsletter-like writeups, Alzforum serves as a continuously-expanding and comprehensive archive of information related to Alzheimer's Disease. They have also created extensive meta-analyses of various genes, biomarkers, and antibodies by collecting information from hundreds of papers, and designed accompanying visualizations. In the process, Elizabeth Wu and the team at Alzforum have learned the value of color to make the figures more approachable, easy to interpret, and enjoyable to interact with.

PRESENTER: KELSEY ROSELL (HOW COLLABORATION AND INTELLIGENT SYSTEMS CAN DELIVER!)

Dimensions is a research information system that is committed to building a collaborative and sustainable scientific community. As a platform that links grant information, clinical trials, researcher profiles, publications, and patents, Dimensions facilitates exploration of resources, collaborations, and funding opportunities. Among numerous other functions, Dimensions allows users to see the citation ratios of institutions or individuals as a function of a particular search term or custom category. This allows the user to extract the amount of interest a particular investigator has in specific fields and how often they get cited within

that community. It also provides visualizations of researchers' networks to facilitate the identification of collaborators, by including the journals in which they are published, the timelines of their research, and who their funders are. A more global database filters information by active grants and their associated timeframes while separating by country. Rosell's presentation positioned Dimensions as a possible contending software resource that could suit the needs of the Inventory working group.

PRESENTER: MATHEW ABRAMS (THE INTERNATIONAL NEUROINFORMATICS COORDINATING FACILITY (INCF))

Nearing the end of the information blitz session, Dr. Abrams kept his talking points brief. He boiled down the INCF's expertise in information network solutions into four parts. First, the INCF has had a role in the IBI itself, which positions them well in understanding the community needs for the inventory. Secondly, as an organization centered around facilitating neuroinformatics coordination, INCF has improved and arrived at standards and best practices that will prove valuable in the development of the IBI inventory. Third, much thought has been invested in improving the accessibility of the information that is put online, which falls under the umbrella of training and education. To this end, INCF has set up an online training portal to promote career pathways in neuroinformatics, brain and medicine, and computational neuroscience. Finally, and importantly, those involved in the INCF have discovered through experience what *not* to do when coordinating and disseminating information. Thus, Dr. Abrams offered his expertise and support of his organization to advise in the creation of the IBI inventory.

BREAKOUT DISCUSSIONS

Day 1 of the IBI Inventory Workshop concluded with breakouts into discussion groups:

- Database infrastructure
- Interoperability/integration with existing efforts
- Resource usability and international user experience
- Mechanisms for sustainability

The results of these small-group discussions were discussed during Day 2.

DAY 2

SUMMARY OF BREAKOUT DISCUSSIONS

The Inventory Working Group convened during the second day to discuss the results of the previous day's breakout sessions.

DATABASE INFRASTRUCTURE WAS REPORTED ON BY DAVID KEATOR:

This breakout group primarily discussed the strengths and weaknesses of using Dimensions as the software solution for the inventory. The group established that the inventory should begin at the micro-scale in building a global database of funded projects. Then, the next course of action would be to reach out to other international teams to further understand accessibility and improve upon the system. An important consideration is when to "build versus buy." In the latter case, Dimensions represents a good opportunity. They next discussed the extent to which Dimensions meets the requirement for the inventory. In terms of scalability, it can be hosted on Amazon. Whether it can operate in all participating countries and the number of instances that can be supported is to be determined. In the ways of privacy protections, it is possible to restrict access at all levels of the administrative hierarchy. The intended user base of Dimensions is currently oriented towards scientists, giving investigators access to the same information used by funders. A global Dimensions database of efforts made by IBI sites therefore caters to both funders and investigators. Finally, the database infrastructure discussion group proposed that the next steps in this process might be to populate a Dimensions database with awards, publications, and patents and build in additional features over time and as needed.

INTEROPERABILITY AND INTEGRATION WITH EXISTING EFFORTS WAS REPORTED ON BY AGNES MCMAHON:

Some existing efforts to integrate research metadata include doi handles, ORCIDs, PubMed, NIH RePORTER, World RePORTER, and preprint archives such as Bioarchive for new techniques that have yet to be published. For some, preprint archives are more useful than information on awarded grants because tools may be in development far before they appear in a published article.

The discussion was diverted in the direction of community engagement in the chosen platform, with McMahon expressing the importance that the inventory becomes a platform to integrate community conversations and networks. She asked the group: how do we start

the process of community engagement? **Mathew Abrams** proposed starting with program officers and letting involvement trickle down to PIs and graduate students. **Michael Yassa** suggested building a dynamic, constantly updating, news layer. This would be labor-intensive, but could be a secondary priority and future goal.

RESOURCE USABILITY AND INTERNATIONAL USER EXPERIENCE WAS REPORTED ON BY AMY BERNARD:

To begin optimizing the inventory for use in the United States and abroad, it is important to define specific user groups. Having a list of specific people and research groups that could provide feedback on the inventory would help the IBI pick out aspects of the inventory that best serve different users' interests. This breakout group also reminded themselves and the rest of the workshop that some researchers have access only to dial-up internet, so it is important to remember that elaborate visualizations of data won't be universally accessible. A mobile-friendly version would be an addition to consider. Multilingual support, or at least a statement of accessibility that is multilingual, would also improve international user experience. One way of incorporating research in language other than English, although insufficient in the ways of broadening the reach of English language resources, would be to simply have the platform point towards journals in other languages. International funding sources could also be listed, as well as written guides to help international investigators determine under which circumstances they might be eligible for US funding.

The group briefly considered the extent to which the NIH World RePORTER could be repurposed for the inventory, but remembered from Dr. Danielson's presentation that it is the most heterogeneous and least extensive of the NIH tools.

MECHANISMS FOR SUSTAINABILITY WERE REPORTED ON BY MICHAEL YASSA:

What is the best way to continuously fund a resource without having to repeatedly show milestones, as required by funding agencies, to stave off the risk of losing it? One possible solution is to use a hybrid subscription model, where light to moderate users can access the resource for free and heavier users must pay a subscription. Instituting a hybrid subscription model would supplement grant funding, which could be considered an indirect cost towards which sponsors put 0.5% or 1% of the budget. These sponsors, such as NSF and NIH, would likely want to know how many staff are required for curating and producing content for the inventory. The inventory working group would need to anticipate a number, perhaps falling between 5 and 7 people, slightly fewer than Alzforum.

Satrajit Ghosh added that including private companies in the inventory development may contribute to sustainability. Perhaps the inventory could work in more pieces that align with priorities of different stakeholder groups.

DISCUSSION OF BREAKOUT GROUPS SUBSEQUENTLY ENDED, AND THE FOCUS SHIFTED TO ACTION ITEMS AND FUTURE DIRECTIONS. THESE POINTS WERE AGGREGATED IN THE EXECUTIVE SUMMARY, COPIED BELOW FOR CONVENIENCE:

EXECUTIVE SUMMARY OF THE SYSTEMS IMPLEMENTATION WORKSHOP

IRVINE, CA | OCTOBER 22-23, 2018

RESTATEMENT OF THE GENERAL GOALS AND STRUCTURE OF IBI:

To unite 7 major brain initiatives across the world:

- Australian Brain Initiative (in development; est. 5-year duration)
- Canadian Brain Research Strategy (in development)
- China Brain Project (in development; est. 15-year duration)
- EU Human Brain Project (2013-2023)
- Japan Brain (2014-2013)
- Korea Brain Initiative (2016-2026)
- US BRAIN Initiative (2013-2025)

Each initiative to appoint representative(s).

Working groups:

- Data standards and sharing
- Neuroethics
- *Inventory of global brain projects*
- Education and training
- Communication and outreach
- Tool and technology dissemination

A central strategy committee oversees coordination, which includes a secretariat that serves with a one-year term. However, nobody in the governance structure currently dedicates 100% time to the IBI.

Across all working groups, there are four stated aspirational goals:

1. To promote coordination and leadership
2. To transcend borders
3. To share and disseminate knowledge
4. To shape the future

…and five use cases for the Inventory were touched upon by the first survey:

1. Identifying and maintaining collaboration
2. Resource awareness and training
3. Identifying potential research topics and areas of interest
4. Facilitating data use and reuse
5. Transparency in funding and funding opportunities

The major audiences considered during this workshop were Investigators and Funders, with private companies as a possible third party that may contribute to sustainability.

PHASE 1 (NOW):

- Identify potential collaborators through data on currently funded projects
    - Build vs. buy: potential best option is to use Dimensions. To what extent does it meet our requirements?
        - Scalability: can be hosted on Amazon. TBD if this can operate in all participating countries and the number of instances that can be supported.
        - Protections: it is possible to restrict access at all levels of the admin hierarchy
        - Audience: Dimensions is currently a scientist-oriented service that gives investigators access to the same information used by funders. A global Dimensions database of efforts made by IBI sites therefore caters to both audiences (funders & investigators)
        - Begin with awards, publications, and patents; build in additional features over time and as needed
    - Action Items:
        - finding out which IBI data would already be in the system
        - (Eventually) enroll in the trial version of Dimensions (up to 90 days)
- Demonstrate research impact and productivity
    - This is more of an outcome of the platform, rather than something the platform can deliver
    - No clear conclusion reached on how to measure this

    - Some measure of success based on whether there are new, boundary-traversing collaborations where there formerly were none
    - When dynamic content is implemented, this could be measured by the number of website visits over time
- Find and fill gaps in brain research
    - Awards and publications are low-hanging fruit – more difficult is finding tools that have yet to be published
- Create a community discourse platform
    - Inventory cannot be static – phase 1 will incorporate a forum, while phase 2 may introduce a dynamic news layer
    - Fills the need to engage *networks* of individuals and their conversations
    - What communities already exist? Can we find out through the use of ORCIDs?
    - Does a discourse platform already exist? The activation energy for starting a brand new one could be high, as there would be no foundational online community:
        - Could Neurostars can be repurposed? Could we piggyback off of new IBI site when it rolls out -- IBI.neurostars.org ?
        - Other possible platforms: Reddit (mature, but presents curation challenges), Discourse (open source and skinnable; searchable; login element that could be coupled with ORCID; unclear if internationally accessible but possibly best option), and Discuss (not open source)

SOME RELATED TO-DOS:

- Refine the survey
    - Goal of the survey: treat the values that we have converged upon in this work group as hypotheses to test
    - Survey should only take 5 minutes, and should capture the intention of: "here is what we are intending to build; what would *you* do with it?"
    - Arrive at some concise sentences that summarize the mission:
        - The International Brain Initiative (IBI) working group on the Inventory of Brain Projects seeks to release a resource that will aggregate data for the funded projects, funding opportunities, and resources created by the Initiatives involved in the IBI. Scientists and funding bodies will use it to discover and record research projects, tools, and collaborators. In order to develop a resource that provides the most utility to our community, we would like end-user feedback for the Inventory. The following survey/session will attempt to capture the neuroscience community's priorities for an IBI Inventory.
    - Disseminate at SfN; scan badges to create beta-test email list

    - Disseminate at FENS?
    - To what extent can the communication and outreach working group promote the Inventory?
    - Here is the link to make edits to the current survey content: https://docs.google.com/document/d/1EA8UlclQ0RqwYaZPLIXHrjeLVrc5AGKwyzGXyHamIVM/edit?usp=sharing
- When survey results have been collected, the next steps need to be handled gracefully within the organizational structure of IBI
    - Group discussed where along the "asking for permission versus forgiveness" continuum our implementation will be, but the IBI intends to avoid any top-down model for organization and working groups will be encouraged to function semi-autonomously
    - This should resemble the path taken by the neuroethics committee, which did *not* go through the strategy committee before producing results
    - Best course of action is to treat the inventory as a proposal or a set of proposals that could be encouraged by the strategy committee
    - Permission will need to be obtained before reaching out to international program officers – will require socialization that runs high up the organizational hierarchy
- Expressed need to ensure that this group will be the next group of people to work on the Inventory, unless the product is easily handed off
    - Aggregate a list of *everyone* who has been involved with the conception of the Inventory
- As there is no centralized nonprofit for this project, some organization must own the governance structure so that we can actually obtain services such as Dimensions
    - It will otherwise fall on the funding agencies
    - This issue should be brought to the strategy committee
    - Once this is initially funded, sustainability can be met through a hybrid model: free access unless particular institutions are heavy users and must pay some amount to use. Can also be considered an indirect cost that sponsors put 0.5% or 1% of budget to this infrastructure

RESOURCES

The following resources were mentioned during the workshop:

- ALLEN BRAIN OBSERVATORY: "… presents the first standardized IN VIVO survey of physiological activity in the mouse visual cortex, featuring representations of visually evoked calcium responses from GCaMP6-expressing neurons in selected cortical layers, visual areas and Cre lines."

- ALZFORUM: "…expands the traditional mode of scientific communication by reporting the latest scientific findings and industry news with insightful analysis that puts breaking news into context. In addition, we advance research by developing open-access databases of curated, highly specific scientific content to visualize and facilitate the exploration of complex data."

- BLUE BRAIN: "The goal of the Blue Brain Project is to build biologically detailed digital reconstructions and simulations of the rodent, and ultimately the human brain. The supercomputer-based reconstructions and simulations built by the project offer a radically new approach for understanding the multilevel structure and function of the brain." The Blue Brain project includes the knowledge graph, NEXUS.

- DIMENSIONS: "a dynamic, easy to use, linked-research data platform that re-imagines the way research can be discovered, accessed and analyzed. Within Dimensions, users can explore the connections between grants, publications, clinical trials, patents and policy documents."

- DKNET: "…an information portal created to provide information to connect the NIDDK community of researchers with scientific resources. The portal includes access to large pools of data and research resources relevant to the mission of The National Institute of Diabetes Digestive and Kidney Disease (NIDDK)."

- GIRDER: "…a free and open source web-based data management platform developed by Kitware as part of the Resonant data and analytics ecosystem. What does that mean? Girder is both a standalone application and a platform for building new web services."

- INCF (INTERNATIONAL NEUROINFORMATICS COORDINATING FACILITY): "…an independent international facilitator catalyzing and coordinating the global development of neuroinformatics, and advancing training in the field." INCF develops standards and

best practices, training and educational resources, promotes open science and sharing of data, partners with international stakeholders, and engages funding partners in collaborative, community-driven projects.

- INTERLEX: "…a dynamic lexicon of biomedical terms. Unlike an encyclopedia, a lexicon provides the meaning of a term, and not all there is to know about it. InterLex is being constructed to help improve the way that biomedical scientists communicate about their data, so that information systems like NIF and dkNET can find data more easily and provide more powerful means of integrating that data across distributed resources."

- METACELL: "… a software company that puts neuroscience data online**,** creating compelling visual and collaborative experiences. Our software applications unlock the true value of neuroscience data and models, including microscopy and MRI images, EEG and electrophysiology data, computer simulations, and much more." MetaCell offers custom software, user interfaces and big data visualizations, and consulting.

- NEUROLEX: a dynamic lexicon of neuroscience concepts

- NEUROSTARS: an online forum primarily used in the neuroimaging community. "The StackOverflow of neuroscience by @INCForg"

- NIDM: "...a collection of specification documents that define extensions the W3C PROV standard for the domain of human brain mapping. NIDM uses provenance information as means to link components from different stages of the scientific research process from dataset descriptors and computational workflow, to derived data and publication."

- NIF: "…an initiative of the NIH Blueprint Consortium, which brings together 16 NIH Institutes, Centers and Offices that support neuroscience research into a collaborative framework to coordinate their ongoing efforts and to plan new cross-cutting initiatives" NIF includes the NIF Registry, a data sharing service, LinkOut Broker, and Ontology Engineering.

- NIH REPORTER: "an electronic tool that allows users to search a repository of both intramural and extramural NIH-funded research projects from the past 25 years and access publications (since 1985) and patents resulting from NIH funding. In addition to NIH-funded research, the system provides access to research supported by the

Centers for Disease Control and Prevention, the Agency for Healthcare Research and Quality, the Health Resources and Services Administration, the Substance Abuse and Mental Health Services Administration, and the U.S. Department of Veterans Affairs."

  - FEDERAL REPORTER: includes grant information from the DOD, ED, EPA, HHS, NASA, NSF, USDA, and VA
  - WORLD REPORT: "…an open-access, interactive mapping database project highlighting biomedical research investments and partnerships from some of the world's largest funding organizations. The site is designed to facilitate communication and coordination of biomedical research."

- NITRC: includes "the Resources Registry (NITRC-R), Image Repository (NITRC-IR), and Computational Environment (NITRC-CE), [with which] a researcher can obtain pilot or proof-of-concept data to validate a hypothesis for just a few dollars… NITRC's scientific focus includes: PET/SPECT, CT, EEG/MEG, optical imaging, clinical neuroinformatics, computational neuroscience, and imaging genomics software tools, data, and computational resources."

- OPEN HUMANS: "…a platform that allows you to upload, connect, and privately store your personal data – such as genetic, activity, or social media data… For researchers and citizen scientists, *Open Humans* offers a toolbox to easily create new projects which can efficiently ask an engaged audience of participants to join and contribute. or join research projects."

- OPENWORM: "…an open source project dedicated to creating the first virtual organism in a computer… because modeling a simple nervous system is a first step toward fully understanding complex systems like the human brain."