



Published in final edited form as:

J Transl Med Epidemiol. 2014 ; 2(2): 1026–.

Research Networking Systems: The State of Adoption at Institutions Aiming to Augment Translational Research Infrastructure

Jihad S Obeid^{1,*}, Layne M Johnson², Sarah Stallings³, and David Eichmann⁴

¹South Carolina Clinical and Translational Research Institute, Medical University of South Carolina, USA

²Institute for Health Informatics, University of Minnesota, Minneapolis, USA

³Vanderbilt Institution for Clinical and Translational Research, Vanderbilt University Medical Center, USA

⁴School of Library and Information Science, University of Iowa, USA

Abstract

Fostering collaborations across multiple disciplines within and across institutional boundaries is becoming increasingly important with the growing emphasis on translational research. As a result, Research Networking Systems that facilitate discovery of potential collaborators have received significant attention by institutions aiming to augment their research infrastructure. We have conducted a survey to assess the state of adoption of these new tools at the Clinical and Translational Science Award (CTSA) funded institutions. Survey results demonstrate that most CTSA funded institutions have either already adopted or were planning to adopt one of several available research networking systems. Moreover a good number of these institutions have exposed or plan to expose the data on research expertise using linked open data, an established approach to semantic web services. Preliminary exploration of these publically-available data shows promising utility in assessing cross-institutional collaborations. Further adoption of these technologies and analysis of the data are needed, however, before their impact on cross-institutional collaboration in research can be appreciated and measured.

Keywords

Collaboration science; Informatics; Translational science; Linked data

INTRODUCTION

One of the key tenets of translational science is the breaking down of barriers between disciplines and institutions to create new collaborations that expedite the translation of new

*Corresponding author Jihad Obeid, South Carolina Clinical and Translational Research Institute, Medical University of South Carolina, 55 Bee St, MSC 200, Charleston, SC 29425, USA, Tel: 843-792-0272; Fax: jobeid@musc.edu.

Cite this article: Obeid JS, Johnson LM, Stallings S, Eichmann D (2014) Research Networking Systems: The State of Adoption at Institutions Aiming to Augment Translational Research Infrastructure. *J Transl Med Epidemiol* 2(2): 1026.

discoveries from inception in the laboratory to clinical practice at the bedside [1]. The National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) program was launched in 2006 to create academic homes for translational research, and provide enhanced infrastructure and training for translational science [1,2]. Inherent to that mission is building better bridges between pre-clinical, clinical, and population health sciences, providing a foundation of shared resources and expertise, and developing partnerships to facilitate integration of research across sites and disciplines. The NIH granted awards to an average of 12 institutions per year over a period of 5 years. In November 2012 there were a total of 61 consortium member institutions. The consortium was organized across several key function areas, such as community engagement, education, informatics, and biostatistics. The CTSA Consortium Steering Committee, which was comprised of the principal investigators of the CTSA sites, key function committee chairs and NIH representatives, provided leadership and oversight to the consortium, and played a key role in setting strategic direction and goals of the consortium. In November 2008 several strategic goal committees were assembled to achieve those goals [3]. One of those committees, Strategic Goal Committee 3 (SGC3), sought to enhance collaborations across the consortium by providing the framework and governance to ensure that researchers can readily locate experts, resources and methodologies, and find networks of investigators to effect collaborative research [3]. The Research Networking Affinity Group, part of the Informatics Key Function Committee, self-assembled in early 2009 to facilitate the implementation of research networking platforms across the consortium. Membership in this group included 60 individuals from various institutions including the NIH, with expertise and/or interest in informatics, research networking and team science. Informed by the Research Networking Affinity Group, SGC3 provided several recommendations regarding research networking to the CTSA Consortium:

- That institutions with CTSA adopt research networking tools supporting Linked Open Data;
- That these data be shared in a form compatible with the VIVO ontology[explained below]; and
- That the Research Networking Affinity Group assists institutions where feasible in achieving these goals.

These recommendations were adopted by the CTSA Executive and Steering Committee in October of 2011. This paper reports the results of two preliminary assessments of Research Networking System (RNS) adoption by CTSA funded institutions that began in the summer of 2012, including: 1) a survey of the degree of adoption of research networking technologies by members of the CTSA Consortium and their plans for research networking over the two year period following the survey; and 2) a preliminary analysis of concept coverage in the VIVO ontology from publically available data at some of the institutions that were surveyed. Although there is perception in the research community, based on anecdotal evidence, that research networking enhances cross-institutional and multi-disciplinary collaboration, it should be noted that the assessment of the impact of these tools on the formation of research teams and the quality of resulting science is outside the scope

of this report. Given the relatively recent implementation and ongoing evolution of these tools at this time, it is too early to assess the long-term impact.

RESEARCH NETWORKING SYSTEMS: BACKGROUND

A new breed of tools that enable users to discover research expertise across multiple disciplines, identify potential collaborators, mentors, or expert reviewers, and assemble teams of science, based on publication history, grants and/or biographical data has recently emerged [4-8]. These tools are referred to as Research Networking Systems in translation science circles and reference social networking with a focus on academic expertise and accomplishments. Schleyer *et al.* describes RNSs as “systems which support individual researchers’ efforts to form and maintain optimal collaborative relationships for conducting productive research within a specific context” [9]. These systems allow users to view researcher profiles derived from data on publications, research activity, and/or areas of expertise with the intention of catalyzing collaboration across disciplinary and institutional boundaries. Investigators can use RNSs to identify possible collaborators using keyword or name searches. Users can also identify networks of authors, explore fields of interests, and identify leaders in a field through the use of these networks. Mentees can identify mentors using the same information [10]. Research Networking Systems help address some barriers to collaboration development as identified through investigator interviews by Spallek, *et al.* [11]. Collaboration barriers include limited and underutilized professional networks and low quality of electronically-available information about potential colleagues. Research Networking Systems are anticipated to overcome these barriers because the networking is topic-driven rather than professional society- or organization-driven. Moreover, several systems are based on new semantic web technologies and data standards allowing interoperability across systems and organizations. By catalyzing the rapid formation of cohesive teams based on well-characterized expertise, RNSs could allow more competitive and effective targeting of funding opportunities.

Research Networking Systems contain information about researcher activities and author relationships within and across institutions. The culture of sharing this information is well-established in research through posting of curriculum vitae (CV) and publication lists on web sites, curation of publication metadata by federal agencies (e.g., NIH's US National Library of Medicine's PubMed), and commercial bibliographic database platforms (e.g., Elsevier's Scopus, and Thomson Reuters' Web of Science). A distinct advantage for RNSs over these approaches is the focus on structured data, as compared to a CV or publication list, and on the individual and the institution, as compared to a citation database. The information in RNSs is a valuable resource not only to research investigators, but also to administrators and institutional leaders, who can leverage RNS data to support institutional intelligence and strategic planning. For example, these data may provide insight into faculty productivity on an institution-wide scale or across institutions to characterize administrative units or departments, and allow the discovery of focus areas and trends that can enhance faculty support, recruitment, and retention programs.

Commercial platforms for social networking such as Facebook, LinkedIn and Google+ excel at linking people through shared interests and acquaintanceships, but these systems lack

focus in scientific activities, collaborations and relationships. More recent commercial systems such as Biomed Experts and Research Gate have attempted to fill that gap but provide limited support for user input and little to no institutional system integration. Moreover, the underlying data in commercial platforms are not open to the public or to the institutions from which they were collected. As a result these data are not available to library science, information science, team science or network researchers who may be interested in mining them for secondary uses.

RESEARCH NETWORKING SYSTEMS AND LINKED OPEN DATA

Research Networking Systems take an alternative approach to these data that emphasizes openness and integration. The information contained in an institution's RNS can be represented as a mesh of Linked Open Data (LOD) [12]. Linked data, the underpinning of the semantic web, are inherently open to the public and provide facile cross-boundary or cross-institutional links. There are clearly identifiable conceptual entities in the persons, publications, grants, and other data found in an individual's profile, and this information is typically curated by both the individual and the institution in the form of one's curriculum vita. The corresponding relationships and roles are also both informally understood by the community and may formally be characterized and organized through the specification of ontological classes, relationships and properties. VIVO is one of the key platforms used in research networking. It was created to represent information about research and researchers, their scholarly work, research interests, and organizational relationships [4]. VIVO was designed around an expressive ontology (the VIVO ontology) and tools for creating and managing LOD for representing scholarship and facilitating discovery of expertise.

Berners-Lee's tiered characterization of LOD [12,13] offers a framework to interpret the existing culture of web-based unstructured information, the existing capabilities of RNSs, and the potential for maturation of the technology (see Table 1).

INFORMATION REPRESENTED IN RESEARCH NETWORKING SYSTEMS

The historical roots of RNS information content lie at two ends of a continuum. At one end lies traditional self-description of an individual through their curriculum vitae, a self-managed document containing unstructured data to characterize one's scholarly activities. The accuracy and timeliness of these data are the sole responsibility of that individual. At the other end are large curation sites aggregating publication or grant metadata in a given domain (e.g., the National Library of Medicine's PubMed for biomedical publications and the NIHR ePORTER for grants). Here accuracy and timeliness are driven by the curating organization (the National Library of Medicine for PubMed) and/or their sources (e.g., journal publishers).

Early versions of RNSs established a functional midpoint on this continuum at the level of the institution. Creating an institutional RNS instance consumed publications and grants metadata, and depended on a combination of attributes such as author affiliation and local knowledge to disambiguate publications, in order to describe individual researchers. As RNSs have evolved, the scope of data not only describes individuals, but has been broadened to include a spectrum of information including courses taught, organizational

structure (e.g., colleges, centers, etc.) and the appointment status of individuals within that structure. With the awarding by NIH of a cooperative agreement grant (U24 RR029822, VIVO: Enabling National Networking Of Scientists) [4], and the subsequent adoption of recommendations involving the VIVO ontology by the CTSA Consortium, the potential representational space describing an individual has become quite rich. However, the concept coverage, i.e. the completeness of the data within this ontology framework, has not been studied before.

METHODS

Research networking survey

The Research Networking Affinity Group solicited survey responses from RNS experts at the 61 CTSA sites during the period of July to October of 2012 to assess the state of RNS adoption in the consortium and the impact of SGC3 recommendations on collaborative tools and LOD. The survey team, formed from members of the Research Networking Affinity Group, developed questions to assess the state of research networking across the consortium along the following themes: How many CTSA investigators are discoverable through a research networking system? How extensively are the research networking recommendations adopted by the CTSA Consortium Executive and Steering Committees being followed? What RNS platforms are in use across the consortium? How mature are the implementations?

Survey development also took into account issues that confound direct comparison across sites, including variation in institutional commitment to research networking and in implementation details (e.g., data sources used to populate profiles).

Questions were developed by the survey team, reviewed by the entire Research Networking Affinity Group, and edited by project managers at the CTSA Consortium Coordinating Center. The survey team queried informatics leaders across the consortium to identify local RNS experts and to establish authoritative contacts for completing the survey. The survey team identified contacts for every member of the consortium (n=61). The finalized survey (made available in supplementary material) was implemented in REDCap [15] and distributed to the site contacts. Respondents were given a total of three months to respond allowing for consultation with colleagues and planners at their institutions. Individuals were asked to submit one response for each unique RNS implementation in planning or in production at their institution. The survey was then closed and the data were analyzed using REDCap tools, MS Access and MS Excel.

COVERAGE ANALYSIS

As a second phase of the work reported here, we conducted an analysis of the ontology coverage of a small number of VIVO ontology-compliant RNSs. This was motivated by discussions within the VIVO Community regarding the institutional variations in the nature of data being added to these systems. Three implementations were selected: one mature VIVO site; one newer VIVO site; and one mature site that do not use the VIVO software but provides VIVO ontology compatibility using an ontology compliant query interface. These

sites provided variation both in how long an RNS had been in operation and in the underlying software that had been implemented. A Java program was written that used the VIVO OWL specification (version 1.4) to construct Simple Protocol and RDF Query Language (SPARQL) queries to count the number of instances of classes defined in the ontology. While our primary interest was in the classes representing the inheritance chain of Agent – Person – Faculty Member, frequency counts for all classes were captured to assess the extent each site was populating the 205 classes comprising the VIVO ontology.

RESULTS

Survey results

There were respondents from 48 individual institutions out of a total of 61 institutions, resulting in a response rate of 79%. Three institutions implemented two RNSs each, and one institution had implemented three RNS platforms, resulting in a total of 53 survey responses from 48 respondents. The majority of respondents (47 of 48) indicated that their institutions had either identified or implemented an RNS, 24 of which had fully implemented systems that were in production. The remaining respondents indicated their institutions were either in planning stages or had pilot implementations (Figure 1).

There was a significant diversity of planned or implemented RNS platforms used across the consortium, which ranged from non-proprietary open source systems to proprietary commercial or homegrown systems. At the time of data collection, 24% of the 53 responses indicated the use of SciVal® Experts [7], 23% used Harvard Profiles [5,10,16], 21% used VIVO [4,17], and 27% used other systems including; Loki [6], Faculty Information System (formerly Digital Vita) [18], Knode (by Knodeinc), Pivot™ (by ProQuest) or other custom implementations (Figure 2). Six percent of the respondents were either in planning phases or unable to specify an RNS platform.

Out of the 48 institutions, 10 (or 21%) were using LOD in the underlying systems implemented, 16 (33%) were planning to use LOD within the next two years and 12 (25%) were considering the adoption of LOD, bringing the total to 79% adoption of LOD if realized or at least 54% if those considering did not adopt (Figure 3).

The majority (72%) of the 53 responses indicated that they included, planned to include or were considering the inclusion of individuals with expertise from outside the life sciences disciplines (figure 4).

Roles of individuals included in these systems also varied significantly between different institutions. Forty-nine out of 53 respondents indicated that they included faculty or planned to include them over the next two years: 4 respondents did not provide an answer to this question. A smaller proportion of systems included post-docs (33) and research fellows (22), and an even smaller proportion included non-faculty staff (Figure 5).

Most implementations (53%) had mandatory inclusion of individuals represented in their RNS. Thirty three percent provided an opt-out policy, 10% had opt-in policies and 4% had no response regarding inclusion policy.

The survey included questions about the types of data included in the RNS. Out of the 53 responses, 51 included or were planning the inclusion of publication data. Forty one included or planned to include funding and grants data. Other RNS data types included patents, courses taught, and recommended funding opportunities (Figure 6). A few respondents indicated that they were unable to determine data types at the time of the survey.

COVERAGE ANALYSIS RESULTS

Table 2 presents an overview of data contained in three RNSs, each operating a publically available VIVO-compliant with SPARQL query interface. For the commonly populated class hierarchy of Agent – Person – Faculty Member, distinctly different protocols were used to populate these classes. University A's RNS had substantial numbers of persons who were not faculty members, and agents who were not persons (counted as instances within the class Agent, Person, or Faculty Member). University B had substantial numbers of persons listed in their RNS who were not faculty members, but included no agents who were not persons. University C did not include agents or persons who were not faculty members.

Additionally, out of the 205 classes defined in version 1.4 of the VIVO ontology, only 159 classes or 77.6% at best, were populated, and that was at University A, with the most mature VIVO installation. When limited to classes with at least 100 instances, population rates drop even lower, to 43.4% at best (Table 2).

DISCUSSION

The survey data show a diverse spectrum of implementations across several dimensions, including levels of adoption or maturity of implementation, types of systems implemented, expression of the data as LOD, types of individuals included and types of data about the individuals included. The more mature implementations tended to include more data types and more types of individuals. Respondents reporting less mature implementations seemed to be planning to follow suit within two years. In an effort to keep the survey short and manageable, attitudes toward the systems were not assessed. Costs associated with implementing RNSs were not assessed in this survey, so return on investment and other comparative financial information cannot be discussed at this time.

The results also indicate that several of the SGC3 research networking recommendations have been adopted and have most likely impacted the evolution of RNS implementations. The recommendations provided a common means of sharing research productivity metadata and the basis for building a metrics-driven framework to allow for determining the extent of utilization and impact of RNSs on translational research. Developers and vendors alike have indicated that they were adopting LOD and the VIVO ontology to make their tools interoperable with other RNS platforms and the semantic web.

The adoption of LOD along with a single ontology standard paves the way for a national federated research network (4, 14, 16), permits refined and aggregated searches to facilitate and enhance collaboration across institutions, and permits interoperability or integration with other LOD-based research resources such as the CTSA connect project [19] and eagle-I

[20,21]. Utilization parameters continue to be reviewed by the Research Networking Affinity Group to inform the development of strategies to measure and evaluate the impact of using RNS tools on the research process and the formation of research teams. In 2013, Weber described a novel method of leveraging data from an RNS for measuring the impact of scientific initiatives, funding agencies and individual scientists on translational medicine [22]. Fazel-Zarandi, *et al.* reported on a recommender system for facilitating the formation of teams of expert scientists that leverages the semantic web along with network analysis [23]. Such systems could also be used for identifying job candidates, reviewers, or mentors for students.

The utility of a common representation, the VIVO ontology, has been clear – particularly in the enabling of secondary federated search systems such as CTSA search [14] and VIVO Search [24]. Studies by Bhavnani, *et al.* [25] and Schleyer *et al.* [26] have evaluated user needs for research networking systems and found that: 1) researchers need authoritative information federation in informatics systems supporting interdisciplinary research; and 2) the use of information beyond publications, such as research interests, grant topics, and patents, provides necessary information for reaching across disciplines to establish collaborations. Linked open data and a common representation like the VIVO ontology facilitate both of these requirements. However, even when a single common representation is used, there is still remarkable variation in its implementation, as shown in the results of the coverage analysis in Table 2. The low population rates in classes with more than 100 instances are not a failing of the ontology or of the implementations by the respective sites, but rather an indication of the early stage of maturation for RNS technology. Currently a gap clearly exists between the idealized notion of an RNS as modeled by the VIVO ontology and the pragmatism of fielding an implementation heavily focused on faculty, publications, and grants, all with readily accessible, public information sources. We anticipate coverage to expand as the field matures

Our survey instrument did not include questions relating to ontology coverage issues, and hence we have no data on the nature of the policy decisions resulting in this variation. Given the variability in the coverage data collected, researchers seeking to assess institutional characteristics using this data should consider the lack of data associated with ontology classes, and will need to independently establish whether it is due to lack of related activity in the areas represented by those classes or just omissions in populating these particular data.

Analyzing this snapshot of the RNS landscape using the framework proposed by Berners-Lee for linked data provides an outlook on the possibilities for cross-institutional collaboration tools as institutions move toward the 5-star criteria for research networking. Linked open data seem to be the future of data sharing for both expertise and resources with the ultimate objective of faster formation of successful collaborations and reduced costs due to resource sharing. Additionally, LOD provides an open standard that can be leveraged for any number of institutional needs and applications – including faculty reporting, monitoring research activities, and gaining insight into departmental or college research programs. Regardless of the listed potential benefits, further study is needed before an assertion of the impact of research networking and LOD on the quality of collaboration and resulting science can be made.

CONCLUSION

The survey and coverage analysis results represent an early assessment of RNS adoption across a group of institutions similar in that they are CTSA awarded sites and, as such, have committed to advancing clinical and translational research. The data show that most of the institutions surveyed were either in advanced stages of implementation of RNSs or well on their way to do so within two years. The results also illustrate the effect that multi-institution adoption of common data representation standards can have on the design and implementation of RNSs.

Although preliminary exploration of this publically available data through the coverage analysis shows promising utility in assessing cross-institutional and cross-disciplinary collaborations, further adoption of these technologies and analysis of the data are needed before the impact on collaborative research can be addressed.

Finally, we believe this report will inform future implementations of research networking at other institutions by providing a snapshot of the current landscape and by highlighting the need for making these data publically available in a consumable and computable format.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENT

This work was funded by the National Institutes of Health, Clinical and Translational Science Award program grant numbers 8UL1TR000114-02, UL1TR000123 & UL1TR000445, UL1RR024979 & UL1TR000442, and UL1RR029882 & UL1TR000062. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

ABBREVIATIONS

NIH	National Institutes of Health
CTSA	Clinical and Translational Science Award
SGC3	Strategic Goal Committee 3
RNS	Research Networking System
LOD	Linked Open Data
URI	Uniform Resource Identifier

REFERENCES

1. Zerhouni EA. Translational and clinical science--time for a new vision. *N Engl J Med.* 2005; 353:1621–1623. [PubMed: 16221788]
2. Woolf SH. The meaning of translational research and why it matters. *JAMA.* 2008; 299:211–213. [PubMed: 18182604]
3. Reis SE, Berglund L, Bernard GR, Califf RM, Fitzgerald GA, Johnson PC, et al. Reengineering the national clinical and translational research enterprise: the strategic plan of the National Clinical and Translational Science Awards Consortium. *Academic med.* 2010; 85:463–469.

4. Börner, K.; Conlon, M.; Corson-Rikert, J.; Ding, Y. VIVO: A Semantic Approach to Scholarly Networking and Discovery. Morgan & Claypool Publishers; 2012.
5. Profiles RNS. Profiles Research Networking Software; 2013.
6. Iowa, Uo. Loki. 2014
7. Vardell E, Feddern-Bekcan T, Moore M. SciVal Experts: a collaborative tool. Med Ref Serv Q. 2011; 30:283–294. [PubMed: 21800985]
8. Weber GM, Barnett W, Conlon M, Eichmann D, Kibbe W, Falk-Krzesinski H, et al. Direct2Experts: a pilot national network to demonstrate interoperability among research-networking platforms. J Am Med Inform Assoc. 2011; 18(Suppl 1):i157–160. [PubMed: 22037890]
9. Schleyer T, Butler BS, Song M, Spallek H. Conceptualizing and Advancing Research Networking Systems. ACM Trans Comput Hum Interact. 2012; 19
10. Kahlon M, Yuan L, Daigre J, Meeks E, Nelson K, Piontkowski C, et al. The use and significance of a research networking system. J Med Internet Res. 2014; 16:e46. [PubMed: 24509520]
11. Spallek, H.; Schleyer, T.; Butler, BS. Good partners are hard to find: the search for and selection of collaborators in the health sciences.. Fourth IEEE International Conference on eScience; Indianapolis, IN. 2008; IEEE Computer Society;
12. W3.org. Linked Data. 2009
13. Berners-Lee T, Hendler J, Lassila O. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American. 2001
14. CTSAssearch 2013.
15. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009; 42:377–381. [PubMed: 18929686]
16. Weber, GM.; Barnett, W.; Conlon, M.; Eichmann, D.; Kibbe, W.; Falk-Krzesinski, H., et al. J Am Med Inform Assoc. Vol. 18. Suppl 1: 2011. Direct2Experts: a pilot national network to demonstrate interoperability among research-networking platforms.; p. i157-160.
17. <http://www.vivoweb.org/>
18. Pittsburgh, Uo. The Faculty Information System (originally called Digital Vita). 2014
19. Torniai C, Essaid S, Barnes C, Conlon M, Williams S, Hajagos JG, et al. From EHRs to Linked Data: representing and mining encounter data for clinical expertise evaluation. AMIA Jt Summits Transl Sci Proc. 2013; 2013:165. [PubMed: 24303330]
20. Vasilevsky N, Johnson T, Corday K, Torniai C, Brush M, Segerdell E, et al. Research resources: curating the new eagle-i discovery system. Database (Oxford). 2012; 2012:bar067. [PubMed: 22434835]
21. Eagle-i. 2012.
22. Weber GM. Identifying translational science within the triangle of biomedicine. J Transl Med. 2013; 11:126. [PubMed: 23705970]
23. Maryam Fazel-Zarandi, HJD.; Huang, Yun; Contractor, Noshir. Expert recommendation based on social drivers, social network analysis, and semantic data representation.. 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems; 2011;
24. VIVO Search. 2013.
25. Bhavnani SK, Warden M, Zheng K, Hill M, Athey BD. Researchers' needs for resource discovery and collaboration tools: a qualitative investigation of translational scientists. J Med Internet Res. 2012; 14:e75. [PubMed: 22668750]
26. Schleyer T, Spallek H, Butler BS, Subramanian S, Weiss D, Poythress ML, et al. Facebook for scientists: requirements and services for optimizing how scientific collaborations are established. J Med Internet Res. 2008; 10:e24. [PubMed: 18701421]

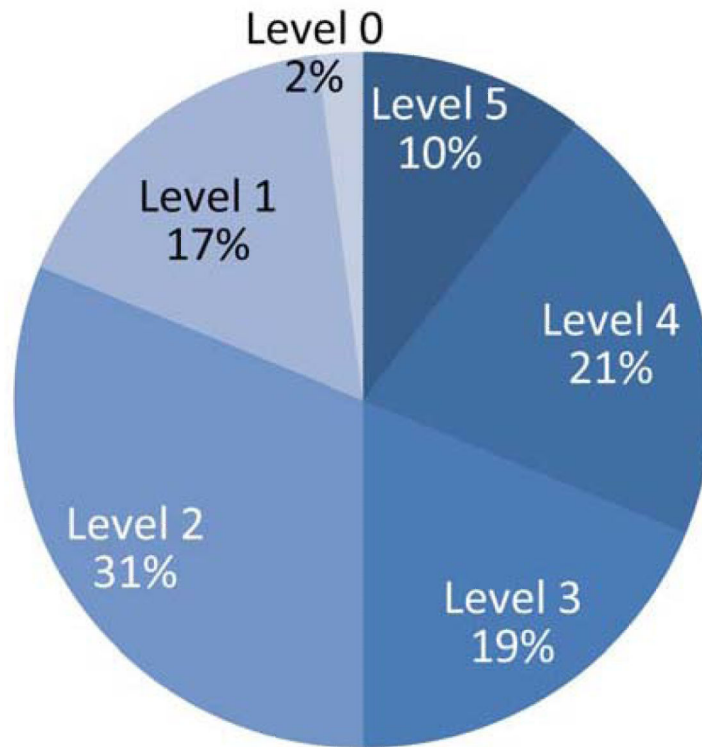


Figure 1. Levels of adoption, n=48 institutions. Level 0 - no concrete plans or steps have been taken towards research networking; Level 1 – identification of tools and planning well underway; Level 2 - initial implementation of RNS performed and data loaded; Level 3 - managed implementation, with regular updates of the data; Level 4 – same as above with automated feeds, and/or greater functionality for administrative and research support processes such as automatic curriculum vitae or biographical sketch generation; Level 5 - optimized implementation which includes the above plus integrating network analyses of team science activities, predictive analytics and prospective grant opportunity assessment.

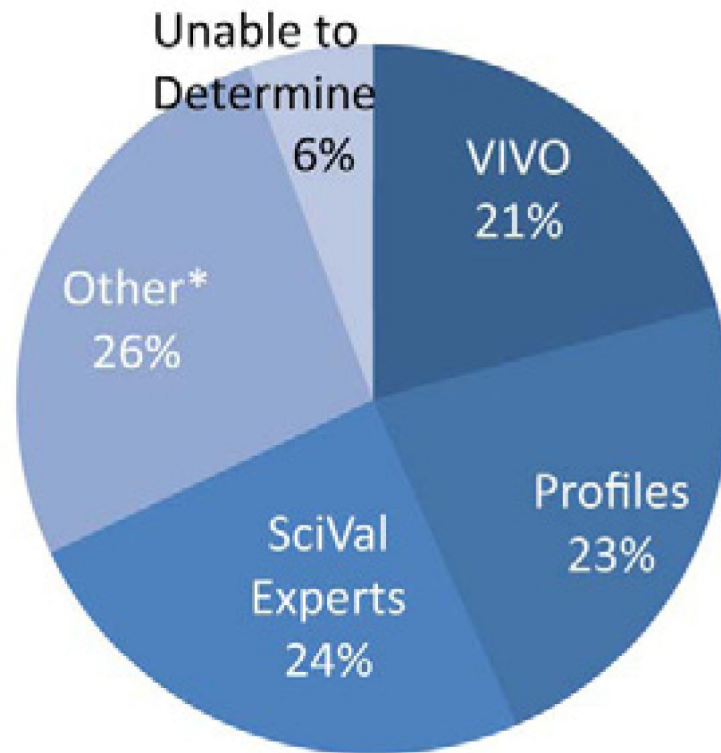


Figure 2. Diversity of systems planned or implemented installations (n=53 responses at 48 institutions). *Other includes: home grown systems or other commercially available systems.

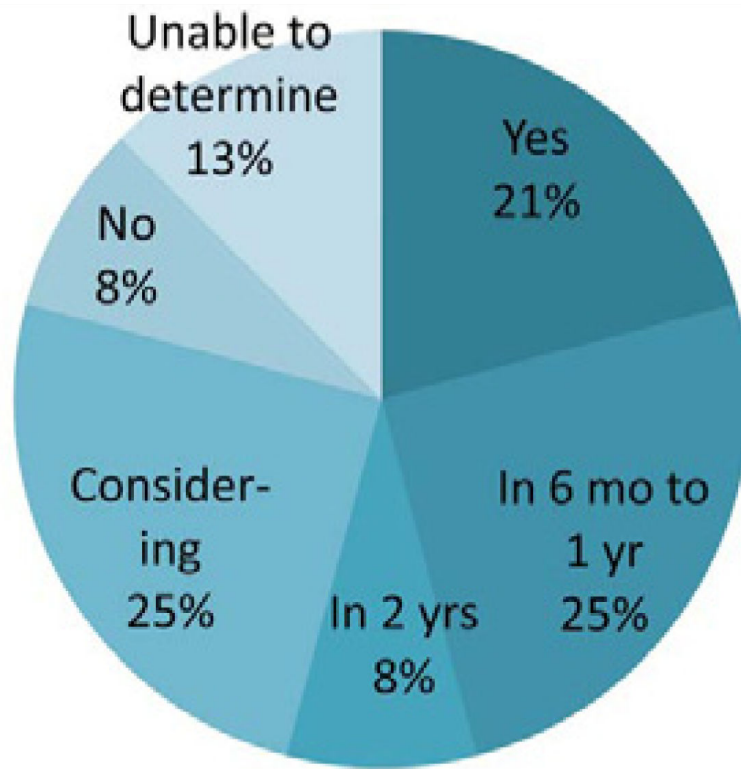


Figure 3.
Adoption of linked open data (n=48 institutions).

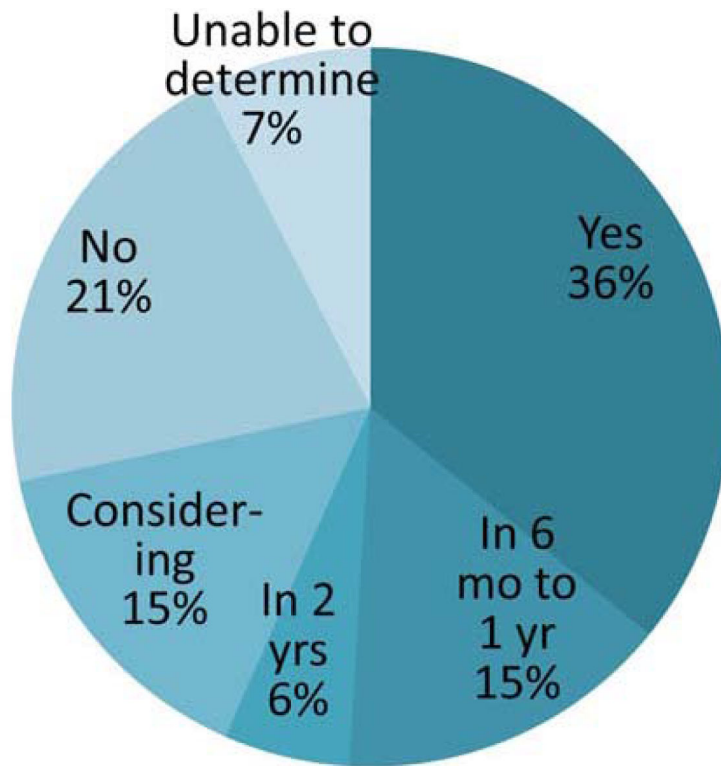


Figure 4.
Inclusion of personnel or faculty outside of the biomedical sciences.

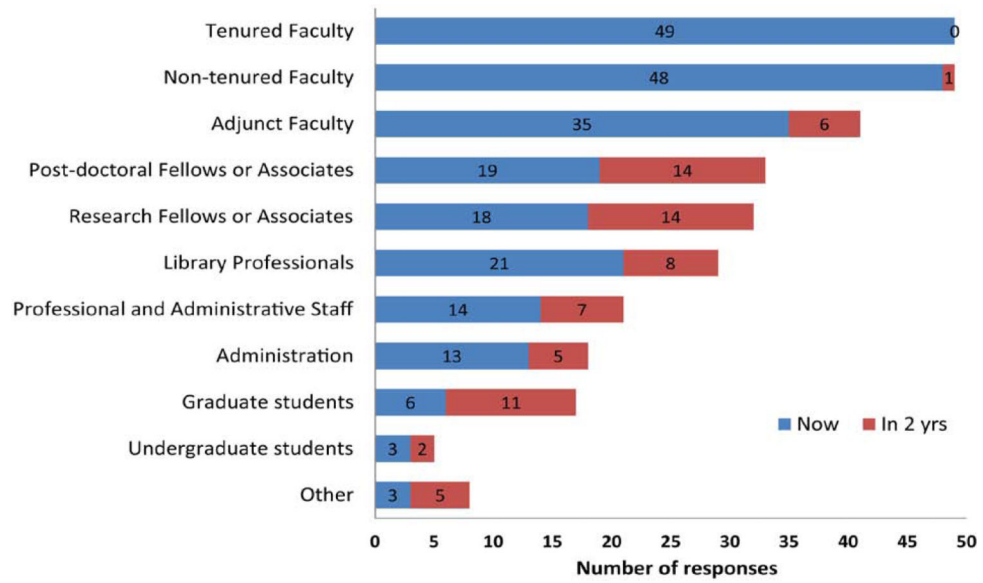


Figure 5.

Roles of individuals whose data were included in the RNS at the time of the survey and those that were planned to be included over the next 2 years. The y-axis represents the roles of individuals, and the x-axis is the number of responses out of a total of 53. Four respondents did not respond to this question.

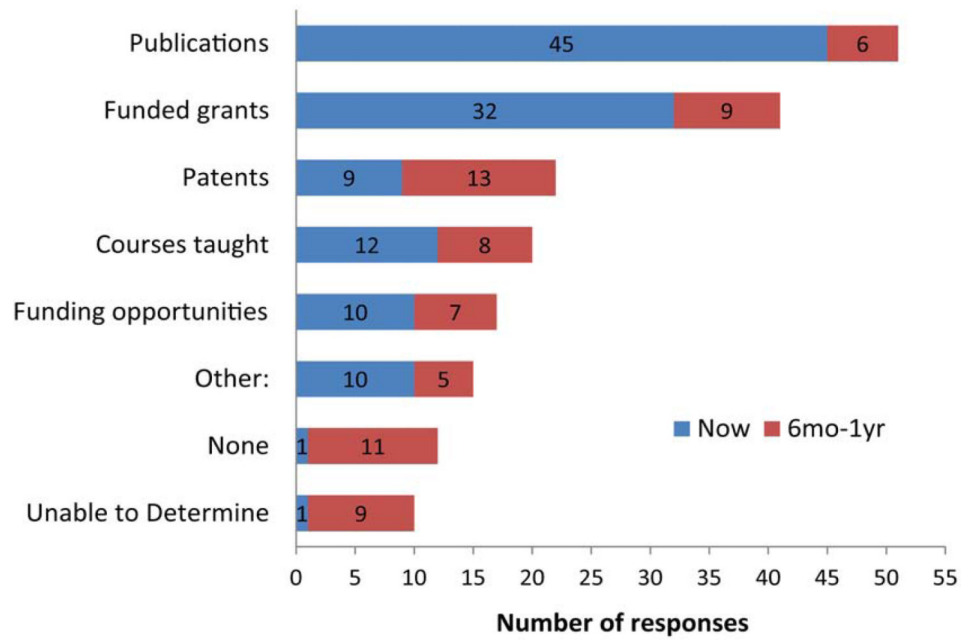


Figure 6. Types of data included in the RNS at the time of the survey and those that were planned to be included over the next 1 year. The y-axis represents the types of data, and the x-axis is the number of responses out of a total of 53.

Table 1

Linked Open Data 5 Star Scheme and Implications for Research Networking Systems [12].

# Stars	Criteria	RNS Implications
1	<i>Available on the web in any format, with an open license</i>	HTML and PDF versions of CVs and publication lists meet this requirement, except in those cases where the author of the artifact asserts copyright. The non-VIVO-compliant (4) RNSs also meet (only) this first criterion, as they appear to the outside world as a content-rich web page, irrespective of the structured data underlying the interface.
2	<i>Available as machine readable structured data (e.g., a spreadsheet)</i>	Providing a link to a library of citations (e.g., EndNote, Mendeley, or RefWorks libraries) from a web page meets this criterion, but with interesting constraints. The most obvious of these packages employ proprietary, binary formats, leading to a proliferation of custom software to access the data. Another constraint is the need to establish the relationship from the linking document and its owner to the included data.
3	<i>as (2) in a non-proprietary format</i>	This criterion removes the first constraint from (2), as the format alternatives are likely to be far fewer in number, with clear documentation as to their nature. The second constraint still holds, however.
4	<i>All of the above, with entities identified through standards (e.g., URIs)</i>	This criterion marks the transition to the establishment of distinct identities for persons, publications, etc. that is independent of the context of an instance of that data. The power of such identifiers has been established by the virtual ubiquity of PubMed identifiers (PMIDs) as references to papers in the biomedical literature. This level still exhibits problems, however. The same publication appearing in multiple RNS instances typically has a distinct Uniform Resource Identifier (URI) in each instance in which it appears. At the time of the writing of this paper, only one RNS instance (University of California, San Francisco [UCSF] Profiles)(10) had begun a pilot to provide extra-institutional link data, and then only for a limited number of other institutions.
5	<i>All of the above, with links to non-local data</i>	This criterion has as its ideal the unique, shared specification of a single URI for any given entity across all systems referring to that entity, in whatever role. This is an admirable goal, but from a practical perspective, much of this inter-RNS linkage will likely be provided by SAME-AS assertions provided through third party solutions (e.g., CTSAssearch)(14) defining equivalences between pairs of RNS instance specific URIs that refer to the same entity (e.g., a person). Even then, the result is a single multi-institutional mesh of data linkages supporting navigation and inference across the RNS content domain.

Table 2

Variations in instance counts from three RNSs for a commonly populated VIVO ontology class hierarchy. The data was collected August 2012 from respective SPARQL endpoints.

	University A (VIVO)	University B (VIVO)	University C (VIVO-compliant)
Class Instance Counts			
Class = Agent	109,572	49,256	3,318
Class = Person	85,582	49,256	3,318
Class = Faculty Member	5,386	6,907	3,318
Populated Classes (out of 205 classes [*])	159 (77.6%)	118 (57.6%)	11 (5.4%)
Populated Classes with > 100 occurrences (out of 205 classes [*])	89 (43.4%)	47 (22.9%)	11 (5.4%)

* Out of a total of 205 classes in the VIVO ontology version 1.4.