Scientometrics (2015) 103:1003–1022
DOI 10.1007/s11192-015-1576-8

Collaborative interdisciplinary astrobiology research: a bibliometric study of the NASA Astrobiology Institute

Zehra Taşkın[1] · Arsev U. Aydinoglu[2]

Received: 9 October 2014 / Published online: 26 March 2015

Abstract This study aims to undertake a bibliometric investigation of the NASA Astrobiology Institute (NAI) funded research that was published between 2008 and 2012 (by teams of Cooperative Agreement Notice Four and Five). For this purpose, the study creates an inventory of publications co-authored through NAI funding and investigates journal preferences, international and institutional collaboration, and citation behaviors of researchers to reach a better understanding of interdisciplinary and collaborative astrobiology research funded by the NAI. Using the NAI annual reports, 1210 peer-reviewed publications are analyzed. The following conclusions are drawn: (1) NAI researchers prefer publishing in high-impact multidisciplinary journals. (2) Astronomy and astrophysics are the most preferred categories to publish based on Web of Science subject categories. (3) NAI is indeed a virtual institution; researchers collaborate with other researchers outside their organization and in some cases outside the U.S. (4) There are prominent scholars in the NAI co-author network but none of them dominates astrobiology.

Keywords Bibliometrics · Astrobiology · NAI · Social network analysis · CiteSpace · VosViewer

✉ Zehra Taşkın
ztaskin@hacettepe.edu.tr

Arsev U. Aydinoglu
arsevu@gmail.com

1 Department of Information Management, Hacettepe University, 06800 Ankara, Turkey

2 Center for Science and Technology Policies, Middle East Technical University, 06800 Ankara, Turkey

Introduction

According to the NASA Astrobiology Institute's 2012 Annual Report, there are 772 active researchers affiliated with 148 institutions in 14 active teams[1] (nodes) that are studying astrobiology (the origins, evolution, distribution, and future of life) related questions (NAI 2013a). Figure 1 shows the distribution of researchers in the U.S. based on their affiliations. Researchers have different levels of expertise ranging from senior researchers to undergraduate students. Since 1998, 42 teams have received funding from NAI to conduct interdisciplinary astrobiology research. NAI has promoted interdisciplinary research, stimulated scientific achievements, and contributed to the establishment of new astrobiology programs.

However, according to the NRC report (2008) measures of interdisciplinarity and collaboration among its members were lacking at NAI. The report (2008, p. 31) recommended that "The NAI should improve the tracking and critical assessment of its publications." The report also suggested that NAI should take some actions, such as establishing a database of publications resulting from NAI funding, inclusion and exclusion of certain types of data and scientific output, and foci of analysis. The NAI has utilized the Report's recommendations in general and improved its tracking system. There has been a very detailed Annual Report process and it collects not only bibliometric data but also project information, updates on education and public outreach activities, team membership info and so on. This study addresses the NRC report's recommendation and provides a bibliometric analysis of NAI-funded research between 2008 and 2012.

Background

According to the National Academies of Science definition:

> Interdisciplinary research (IDR) is a mode of research by teams or individuals that integrates information, data, techniques, tools, perspective, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or field of research practice (NRC 2004, 26).

This definition suggests a very broad spectrum of interactions among researchers from engaging in an informal conversation at a conference to sending samples to a different lab and to having a formal collaboration to investigate a complex problem. Some of these interactions may not necessarily lead to a co-authored scholarly publication. However, it has also been widely acknowledged that a scientific effort is only complete when there is a publication reporting on the work (Wagner et al. 2011); thus, bibliometric analysis is a standard tool for evaluation and in this piece we limit our analysis to it. Both Stokols et al. (2008) and Wagner et al. (2011) emphasize the use of bibliometric tools and network analysis of collaborative efforts and many scholars conducted detailed analyses on

[1] Teams are named after the principal investigator's institution; however, this naming is misleading because these teams are in fact a consortium of researchers from different institutions which create distributed networks. For instance, the Pennsylvania State University Team has researchers affiliated with 40 other institutions in addition to the Pennsylvania State University (41 institutions in total) or the Virtual Planetary Laboratory at the University of Washington Team members are affiliated with 25 institutions all over the world. In addition, a researcher can contribute to more than one team.

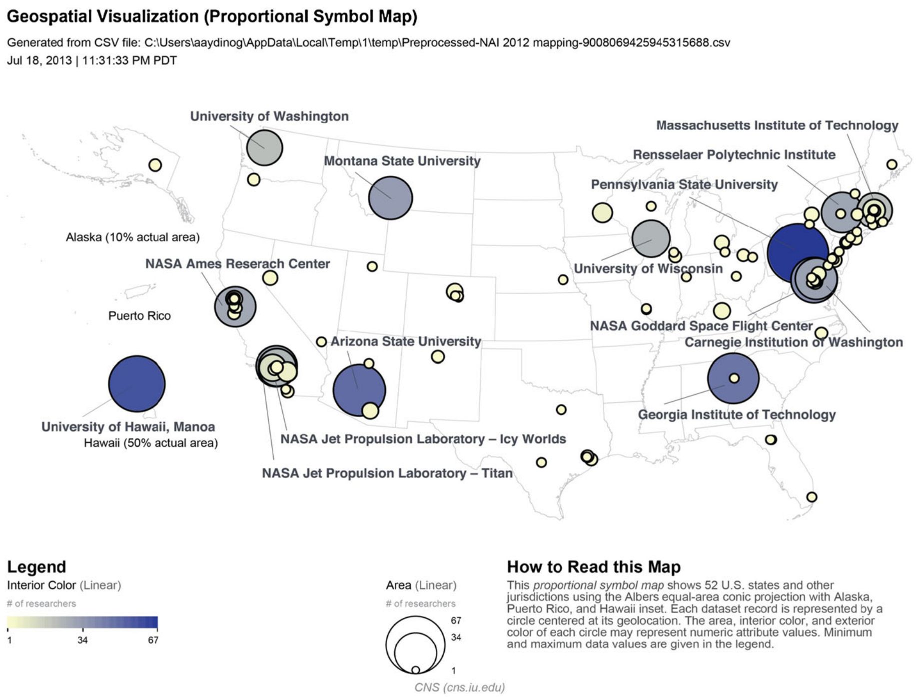

Fig. 1 2012 NAI Network in the U.S. showing the institutional distribution of the members of the 14 teams. Each *circle* represents an institution. The map is generated through Sci2Tool (Sci2 Team 2009)

publications to investigate interdisciplinary research. For instance, Katz and Hicks (1995) looked at ISI journal classifications to investigate cross-disciplinarity. Morillo et al. (2003) also studied ISI journal classifications. Some studies focus on the papers—not journals or journal categories. A highly cited study by Wuchty et al. (2007) that examined 19.9 million papers and 2.1 patents found that science is becoming more collaborative and cross-disciplinary. A follow-up study by Porter and Rafols (2009) also found that there are "major increases in the number of cited disciplines and references per article". Hall et al. (2012) examined the outputs of transdisciplinary team science initiatives and investigator-initiated grants between 1994 and 2014 and found that the former had higher publication rates and average number of coauthors per publication. Porter et al. (2007) increased the scope by analyzing the venue of the publication and the research domains citing it to track the interdisciplinary impact of a researcher or a publication.

These analyses combined with the recent developments in visualization software led an important area of research, science mapping, which is used to visually identify scientific domains, interconnectedness among them, size, etc. For instance Small (1999) developed one of most multidisciplinary high-level science map (at the time) using co-citation data. Boyack et al. (2005) analyzed over a million articles in 7121 journals in order to achieve structural accuracy (accuracy in global and local scale) and found that Biochemistry is the most interdisciplinary discipline in science, followed by General Medicine, Ecology/ Zoology, Social Psychology, Clinical Psychology. Klavans and Boyak (2006) utilized journal citation interactions to develop science maps.

A recent contribution was made by Rafols et al. (2010) to generate science overlay maps which were helpful in investigating less traditional disciplinary categories and their connections to other disciplines. A toolkit was prepared and opened for the community as well by these researchers. Bibliometrics was also used to support the facilitation of cross-disciplinary communication (Williams et al. 2013).

Despite the benefits of bibliometric analysis in understanding interdisciplinary research, there have been only a handful of studies for investigating the interdisciplinarity of astrobiology through them. The first study conducted in the field using publications compared the emergence of Geology to Astrobiology in order to determine whether the latter had become an isolated (or specialized) discipline (Brazelton and Sullivan 2009). A citation analysis of publications in the journal *Astrobiology* and *the International Journal of Astrobiology* revealed that Astrobiology was still interdisciplinary. More recently, Astrobiology Integrative Research Framework (AIRFrame) Group headed by Rich Gazan (University of Hawaii), have been investigating the interdisciplinarity of the NAI to foster understanding across domains, and thereby catalyze interdisciplinary collaboration (AIRFrame). The team used the information bottleneck algorithm developed by Slonim et al. (2002) to assess interdisciplinary research within the University of Hawaii Astrobiology Node using abstracts of the publications that the Node produced (Gowanlock and Gazan 2013; Miller et al. 2014). Afterwards, granted with a Director's Discretionary Fund, the team applied the algorithm to a bigger dataset (publications by NAI-funded research between 2008 and 2011) and identified topically related documents that are not necessarily in the same discipline and where collaborations take place in the greater NAI community.

Aim of study

The aim of this study was to undertake a bibliometric investigation of the NASA Astrobiology Institute (NAI) funded research conducted by CAN 4 and 5 teams that were published between 2008 and 2012. For this purpose the study created an inventory of publications co-authored through NAI funded research and investigated journal preferences, international and institutional collaboration, and citation behaviors of researchers to reach a better understanding of interdisciplinary and collaborative astrobiology research funded by the NAI.

To achieve this aim, following research questions were addressed;

- In which journals did the authors choose to publish their publications?
- What was the distribution of collaboration types (institutional and international) of publications produced by NAI-funded researchers?
- Which NAI-funded researchers were major knowledge producers?
- What were the citation preferences of NAI researchers?

In this section, NAI researchers or NAI authors refer to the co-authors of the publications that resulted from NAI funding to the NAI teams; source refers to journal; institution, organization, university refer to the affiliation(s) of the co-author; and most importantly, it has to be kept in mind that NAI team refers to more than one institution as each team is a conglomerate of researchers from different organizations. The team name is identified the affiliation of the principal investigator (PI). Moreover, this study considers NAI as a network and employs bibliometric and network tools accordingly to provide an assessment of NAI—not the individual teams. The analysis of individual teams is irrelevant to this study for two reasons. Firstly, this study is a response to the recommendations of the NRC Study, which treats the Institute as a whole. Secondly, although all of the NAI teams are multi/interdisciplinary, their research activities fall under different scientific disciplines, each of which has different values, workflows, and publication habits; therefore, an analysis based on the breakdown of teams will not only result in incomparable results between teams, which will be not only useless/irrelevant but also misleading (if one tries to compare one team to another).

Methodology and data

The main aim of bibliometric studies is to evaluate scientific publications and their references deeply. Revealing the impacts of scientific works becomes possible by the help of bibliometric techniques which depend on quantitative and qualitative analyses. The results of bibliometric studies are used by decision-makers and managers to identify effective knowledge producers (authors, institutions, countries etc.), to visualize scientific impact, and to distribute tenures and incentives. The main data tool for bibliometric analyses is citation databases, such as Web of Science and Scopus. Although traditional bibliometric studies are based on counting publications and citations; social network analyses (SNA) are used as the contemporary research method for studies. These analyses comprise the social structures of some actors, such as authors, countries, institutions and so on, and knots of the relationships between these actor pairs (Al et al. 2012, p. 42). Some software is designed to visualize social networks in the literature. Well known SNA tools are Pajek, CitaSpace, Sci2, and VosViewer.

To conduct a bibliometric study on NAI publications, we evaluated only peer-reviewed publications that were included in the NAI annual reports between 2008 and 2012 because at the time of this study, active NAI teams had started their work in 2008 and the latest data available was from 2012. The Thomson Reuters' Web of Science was used as a data tool. As a result, 1210 peer-reviewed publications produced by NAI-funded teams were gathered. A deep data cleaning process was carried out to access reliable and accurate results. The cut-off date for the citation datasets was June 2013. All information about author, institution and country names were unified into standardized format. Web of Sciences' subject categories were used to classify publications. SPSS and MS Excel were used for statistical calculations about frequencies and standard deviations.

One of the most important part of this study was social network analysis of NAI teams and their publications. We used the tools VosViewer (VosViewer 2013), developed by Centre for Science and Technology Studies, Leiden University (http://www.vosviewer.com/), and Citespace (Chen 2014a) created by Chaomi Chen from Drexel University (http://cluster.cis.drexel.edu/~cchen/citespace/) to produce networks and to visualize connections. Two of the tools are Java Applications, therefore they require Java-installed computers to create networks. We converted our dataset into two different.txt formats; field-delimited text was for CiteSpace and tab-delimited text for VosViewer. Then, the software processed data and produced networks automatically. Detailed information about how to create maps by using CiteSpace and VosViewer are in the manuals of the software (Chen 2014b; van Eck and Waltman 2013). Some terms that used in social network analyses terminology were explained in the relevant part of the study.

Results

Number of publications

The researchers of NAI-funded teams (Cooperative Agreement Notice Four and Five, 14 teams in total) have co-authored 1210 peer-reviewed publications in 221 different journals. Only eleven out of 221 journals have approximately 52 % of the publications, which indicates that NAI researchers target certain journals to disseminate their research. The table below presents the name of the eleven journals, number of publications in that journal (N), the impact factor of the journal (a measure reflecting the average number of citations to recent articles published in the journal), and the ranking of these journals in Journal Citation Reports (JCR),[2] which is a database containing journal quality indicators by using citations. Some indicators about journals such as impact factors, total cites, cited half-life, subject categories and rankings are calculated by JCR (Thomson Reuters 2014). JCR uses Web of Sciences' subject categories to define the most effective journals of certain areas. In our dataset, only one journal (Astrobiology) is present in three different categories according to Web of Science categories. NAI funded researchers prefer relatively high impact journals (see Table 1). Publishing in the first, second and fourth ranking of multidisciplinary journals indicate that astrobiology researchers do prefer multi/interdisciplinary journals to publish.

The 80-20 rule, which is also named as Pareto Principle, can be identified in the library and information science literature as "approximately 80 % of the circulations in a library are accounted for by about 20 % of the holdings" (Lancaster and Lee 1985, p. 390). This rule is used in many areas from economics to bibliometrics studies. We found that 80 % of the NAI publications were published in 45 journals (20 % of journals) (see Fig. 2), which meant that the publication pattern fit into 80-20 rule since 80 % of NAI publications (969) were in 45 journals. Since 221 is the total number of journals, the 80-20 rule was verified.

Journal categories and network topology

NAI researchers published their research in Astronomy and Astrophysics[3] journals the most (see Table 2). Out of 1210 publications, 464 were tagged in Astronomy and Astrophysics (among the 464, 84 % had only Astronomy and Astrophysics tag, and the rest had multiple tags including "Geosciences, Multidisciplinary"; "Biology"; and "Meteorology and Atmospheric Sciences"). Geochemistry and Geophysics was the second most popular journal category with 225 publications (18 were tagged in Mineralogy and two with Marine and Freshwater Biology and Oceanography); Geology was the third journal category with 169 publications. Only 25 % of the publications were published in a journal with more than one tag. When Category 56—multidisciplinary is added to that (publications that were tagged with more than one category −25 %), a little over one third (34.5 %) of all publications had multidisciplinarity based on journal categories.

Based on the publications funded through NAI between 2008 and 2012 and using Web of Science Journal Categories, a betweenness centrality analysis identified 44 nodes and 59

[2] The reason for using journal category is the assumption that certain journals have certain audiences based on their category. Publishing in a different category means reaching out to a different audience, hence a proxy for multidisciplinary interaction.

[3] In this section "Astronomy & Astrophysics" is the Web of Science Subject Category—not the journal title.

Table 1 Most common journals to publish in

Journal name	N	%	Impact factor	Journal rank
Astrophysical Journal	161	13.3	6.733	6 of 56[a]
Geochimica Et Cosmochimica Acta	87	7.2	8.884	6 of 76[b]
Icarus	68	5.6	3.161	18 of 56[a]
Science	55	4.5	31.027	2 of 56[c]
Astrobiology	48	4.0	2.803	21 of 56[a] 17 of 83[d] 35 of 170[e]
Earth and Planetary Science Letters	47	3.9	4.349	4 of 76[b]
Meteoritics & Planetary Science	43	3.5	2.800	19 of 76[b]
Astrophysical Journal Letters	39	3.2	6.341	7 of 56[a]
PNAS	36	3.0	9.737	4 of 56[c]
Astronomical Journal	26	2.1	4.965	12 of 56[a]
Nature	25	2.0	38.597	1 of 56[c]

[a] Category, Astronomy & Astrophysics

[b] Category, Geochemistry & Geophysics

[c] Category, Multidisciplinary

[d] Category, Biology

[e] Category, Geosciences, Multidisciplinary

(Data Source: Journal Citation Reports 2012 Edition)

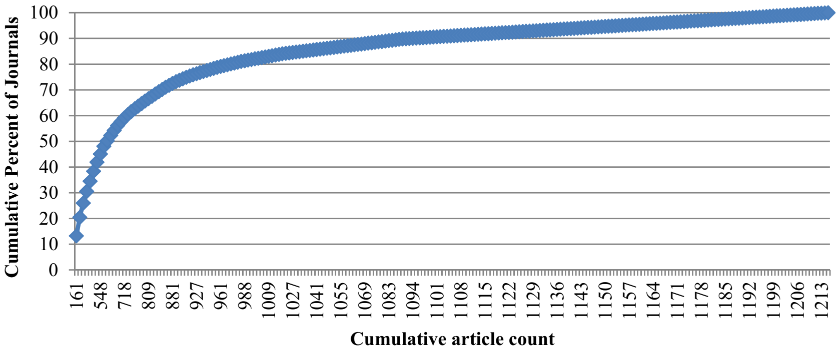

Fig. 2 80/20 rule for NAI publications

edges. Betweenness Centrality—a measure of a node's centrality in a network—is equal to the number of shortest paths from all vertices to all others that pass through that node.[4] We used "SC" (subject category) column of Web of Science to calculate and visualize

[4] *Betweenness centrality* is a more useful measure (than just connectivity) both the load and importance of a node. The former is more global to the network, whereas the latter is only a local effect. The thickness of the lines (edge) shows the degree of connection between the two nodes. The size of the node is the frequency of publications in that domain. The color of the line is the year of publication. Pink Circle means that that node is pivot node—that is the node that makes the interdisciplinary connection. These nodes are strategically important in pulling other nodes together; they have the highest betweenness centrality which is an indicator of a node's ability to make connections to other nodes in a network (Chen et al. 2008, p. 238).

Table 2 Frequency of publications based on journal categories

Category	No. of articles
Astronomy & Astrophysics	464
Geochemistry & Geophysics	225
Geology	169
Science & Technology—Other Topics	127
Life Sciences & Biomedicine—Other Topics	102
Chemistry	53
Environment Sciences & Ecology	52
Microbiology	51
Biochemistry & Molecular Biology	49
Physics	29

category data. According to Web of Science, a single publication can be indexed in two or more different categories. Therefore, CiteSpace creates connections between categories by using these publications. As it is evident from the Fig. 3, the most common categories were not well connected to the rest of the network. This might suggest less interdisciplinarity based on journal categories if we assume that certain journals have certain audiences. Geology was actually the only field that connects Astronomy & Astrophysics to the rest of the NAI network and prevented it from floating alone such as Geochemistry & Geophysics or Science & Technology—Other Topics. Geology, Life Sciences & Biomedicine—Other Topics, Chemistry, Physics, Evolutionary Biology, and Environmental Scientists are important for the overall connectivity of the network.

International collaboration

NAI is a network. As mentioned earlier, 770 researchers in 140 institutions all over the world collaborated to conduct astrobiology research under 14 teams between 2009 and 2012. NAI has a formal partnership with thirteen astrobiology networks outside the U.S. (NAI 2014) in every continent except Africa. Joining forces with the international science communities is an important function of NAI, which has been reflected in the co-authorship behaviors of its researchers (see Table 3).

Institutional collaboration

NAI was established as a virtual institute envisioned as "a distributed network of scientists from different disciplines spread across many sites nationally and internationally to work on projects in which they are mutually interested…" (Blumberg 2003). Each NAI "team" is a conglomerate of organizations, average 17.36 institutions per NAI team (SD = 9.54). The data below represents the NAI as a whole, as an institute—not research teams. The data comes from the author affiliations from the peer-reviewed publications. If a co-author had more than one affiliation, they were represented as well. The institutional collaboration network had 247 nodes (institutions) and 189 edges (connections) with a density of .0064 (see Fig. 4). Although the density of the network was low for evaluation, it showed the main clusters and nodes for institutional collaboration. In the top left of Fig. 4, the Arizona State University refers to NAI-funded papers that was coauthored by researchers who had Arizona State University as their affiliation. The lines to the University of California

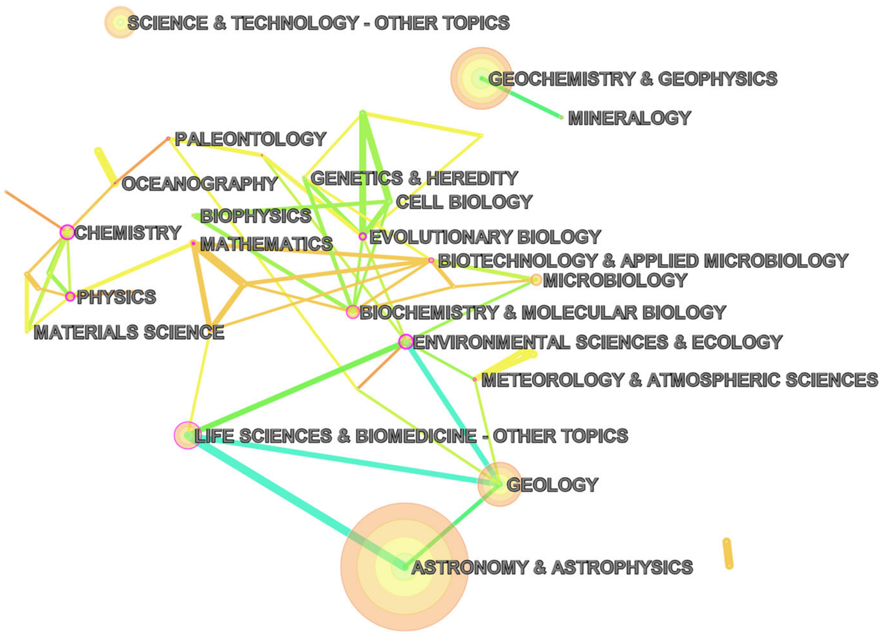

Fig. 3 Network topology of journal categories based on NAI-funded publications [*colors* represent the year of publication or connection (*orange* 2012, *yellow* 2011, *green* 2010, *light blue* 2009, *dark blue* 2008). The *thickness* refers to the number of publications in that year.]. (Color figure online)

Table 3 Top 10 international collaboration based on co-authorship

Collaborator's location	No. of co-authored papers
England	48
France	45
Germany	40
Netherlands	34
Canada	28
Australia	27
Spain	20
Mexico	15
Denmark	13
Italy	12

Riverside and Johns Hopkins University meant that some of these papers had coauthors in these institutions. Density is the sum of the ties divided by the number of possible ties. The density of a network can give us insights on how information diffuses among the nodes, and which actors have higher influence in the network (Hanneman and Riddle 2005). If every node in the network have connection with each other, the maximum density can be calculated as 1.0. A good number of density may be between .40 and .60 (Carpenter et al. 2009, p. 455)".

The visualization of the network analysis is provided in Fig. 4 below. NASA, Carnegie Institute of Washington (CIW), California Institute of Technology (CalTech),

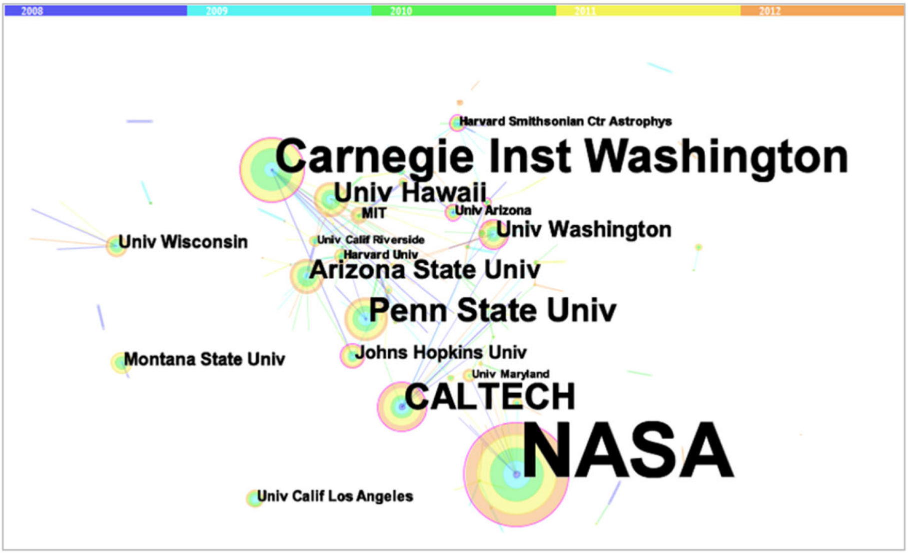

Fig. 4 Network topology based on the institutional affiliations of co-authorship (in order not to clutter the visual with texts of every institutions, a cut-off is applied and only significant nodes and connections are shown) [*colors* represent the year of publication or connection (*orange* 2012, *yellow* 2011, *green* 2010, *light blue* 2009, *dark blue* 2008). The *thickness* refers to the number of publications in that year.]. (Color figure online)

Table 4 Most productive NAI-funded authors are; (by freq)

Author	Freq.	Cent.	Interdisciplinarity of the author[b]	Ph.D. degree of the author
SC Solomon	57	0.05	10 different category	Geophysics
RP Butler	40	0.02	5 dif. cat.	Astronomy
P Ehrenfreund	33	0.31	19 dif. cat.	Astrophysics
GW Marcy[a]	29	0.03	3 diff. cat.	Astronomy & Astrophysics
JW Head[a]	27	0.01	6 dif. cat.	Geological Sciences
LR Nittler	27	0.42	4 diff. cat.	Physics
TW Lyons	26	0.05	16 dif. cat.	Geology—Geochemistry
A Steele	25	0.11	15 dif. cat.	Biotechnology
GD Cody	24	0.05	16 dif. cat.	Geosciences
JW Peters	23	0.00	24 dif. cat.	Biochemistry

[a] Not a member of CAN 4 or 5 teams. What is interesting here is that both GW Marcy and JW Head collaborated and co-authored with NAI-funded researchers. Even though they did not receive funding from NAI and they were not identified as a team member by the Principal Investigators, they contributed (in terms of co-authorship) so much that they are in the top-10 list. The reason for this is unknown as they fall beyond the reach of bibliometric tools

[b] Here interdisciplinarity is defined as in how many different Web of Science Journal Categories an author has a publication. For instance, P Ehrenfreund's NAI relevant publications were published in 19 different categories

Pennsylvania State University, Arizona State University (ASU), and University of Washington hold the most productive and connected researchers. On the periphery, the University of Wisconsin, University of California—Los Angeles, Montana State University, and Rensselaer Polytechnic Institute are productive. It is not surprising that in most of the cases, these institutions overlap with the PIship of the teams. CalTech (California Institute of Technology) is the official employer for the researchers at the NASA Jet Propulsion Laboratory (JPL) which hosts two NAI teams.

In addition to the institutions above, the richness at the center of the network demonstrates that the NAI is well distributed into the national and international research network through collaborators at University of Arizona, Johns Hopkins University, University of California (Berkeley, Riverside, Santa Cruz, San Diego), University of Maryland, University of Colorado, Harvard University, and Brown University to name a few among nationals and the Centre National de la Recherche Scientifique, Universiteit Leiden, University of New South Wales, and The Universidad Nacional Autónoma de México among internationals. As for the strength of collaborations between institutions CIW and Johns Hopkins University, Massachusetts Institute of Technology (MIT) and Pennsylvania State University, MIT and UC Berkeley, ASU and UC Riverside are the most prominent (thickness of edges).

Most productive NAI co-authors

The top-10 most scholarly productive NAI researchers (the number of NAI-funded co-authored publications between 2008 and 2012) are listed in Table 4. The first six researchers in the table were from astronomical sciences; and the remaining are Lyons TW (biogeochemistry), Steele A (microbiology/astrobiology, Cody GD (geosciences), and Peters JW (biochemistry). We also looked at the Web of Science Category and Subject Category areas to see how multidisciplinary their publications are.

We also looked at the betweenness centrality of co-authors in order to understand how vital they were to the rest of the co-authorship network. These authors were the ones who created the network, who connected different co-authorship networks in the greater NAI network; therefore, vital to the collaborative nature of the NAI-funded research. However, the centrality scores were quite low (the highest .42, and declined rapidly) which suggested that even the most productive authors were not prominent in the network. Centrality metrics provide a computational method for finding pivotal points between different specialties or tipping points in an evolving network (Chen 2006, p. 362). It is estimated that average centrality rate may be between .40 and .60. In our dataset, we found that there were no actors who dominated the network. This might be due to the multidisciplinary nature of the field—a researcher's influence does not go beyond his/her immediate domain.

Citation analysis

The authors of NAI used 70,752 references for their papers. The average number of references for each publication is identified as 58.47. In addition, publications that resulted from NAI funding were cited 22,056 times between 2008 and 2012. The publications in recent years need more time to be cited as there is a temporal dependence on citations.

Among the NAI funded publications, the ones that were cited the most, the number of their citations, their publication year, and the journals where they are published are shown on Table 5 (mean of the citations is 18.23 and the median is 8.00.)

Table 5 Top-10 mostly cited NAI-funded publications (citations from Web of Science system)

Title	Journal	Publication year	No. of citations
The HITRAN 2008 molecular spectroscopic database	Journal of Quantitative Spectroscopy & Radiative Transfer	2009	1081
Structure of the 70S ribosome complexed with mrna and trna	Science	2006	597
Kepler planet-detection mission: Introduction and first results	Science	2010	364
Characteristics of planetary candidates observed by kepler. Ii. Analysis of the first four months of data	Astrophysical Journal	2011	278
Kepler mission design, realized photometric performance, and early science	Astrophysical Journal Letters	2010	258
Application of Fe isotopes to tracing the geochemical and biological cycling of Fe	Chemical Geology	2003	235
The Keck Planet Search: Detectability and the minimum mass and orbital period distribution of extrasolar planets	Publications of the Astronomical Society of the Pacific	2008	213
Chemistry and mineralogy of outcrops at Meridiani Planum	Earth and Planetary Science Letters	2005	186
Multiple sulfur isotopes and the evolution of the atmosphere	Earth and Planetary Science Letters	2003	183
The anaerobic oxidation of methane and sulfate reduction in sediments from Gulf of Mexico cold seeps	Chemical Geology	2004	177

Mostly cited sources by NAI co-authors

VosViewer visualization tool identified 5 clusters for mostly cited journals by NAI authors (see Fig. 5). Co-occurrence matrix of data is important for VosViewer. It creates maps in

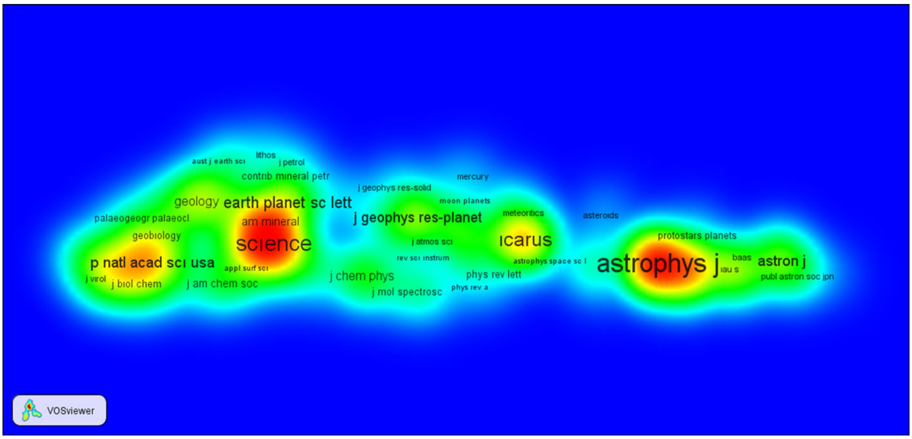

Fig. 5 Mostly cited journals

Table 6 The list of mostly cited journals by frequency and by centrality

By frequency			By centrality		
Freq	Cent.	Journal name	Freq	Cent.	Journal name
899	0.07	Science	115	0.33	American Mineralogist
833	0.11	Nature	56	0.24	Physical Review B
472	0.01	The Astrophysical Journal	237	0.22	Astrobiology
466	0.11	Icarus	111	0.22	Philosophical Transactions of the Royal Society B
418	0.17	PNAS	174	0.19	Applied and Environmental Microbiology
415	0.06	Geochimica et Cosmochimica Acta	90	0.19	Molecular Biology and Evolution
405	0.01	Astronomy & Astrophysics	174	0.18	Journal of Geophysical Research
395	0.02	Earth and Planetary Science Letters	181	0.17	Annual Review of Earth and Planetary Sciences
297	0.00	Monthly Notices of the Royal Astronomical Society	418	0.17	PNAS
280	0.05	The Astronomical Journal	159	0.16	Geobiology

three steps; calculating similarity metrics, mapping and translation-rotation-reflection (van Eck and Waltman 2010, pp. 530–531). It calculates clusters of related items by using similarity metrics. The determined clusters in our dataset were: cluster 1, PNAS is at the center; cluster 2, Science at the center (nature is at the same spot but not visible in this visual); cluster 3, Journal of Geophysical Research at the center; cluster 4, Icarus at the center; and Cluster 5, the Astrophysical Journal at center. A surprising finding here was that none of the top-10 mostly cited sources, except for PNAS, are in the top-10 centrality list. This meant that the most cited journals were not the most vital ones in terms of connectivity in a network. The centrality and frequency scores of mostly cited journals are shown on the Table 6.

NAI-funded authors utilized publications from a great variety of sources to cite in their research. They cited six thousand seven hundred and seventy four (6774) unique sources between 2008 and 2012. The most popular journals for citation were; Astrophysical Journal, Science, Nature, Astronomy & Astrophysics and Icarus. It is obvious that the authors generally cited journals which they publish their publications. The journal preferences for citations and publications were important to show core journals in the field.

Mostly cited authors by NAI co-authors

The mostly cited authors by NAI researchers were: Canfield, DE; Kasting, and JF; Butler, RP (see Table 7). The co-citation map of mostly cited authors is shown on Fig. 6. The mostly cited researchers for the NAI network seemed to be the researchers who were already funded by NAI, except for Mayor, M. However, centrality scores revealed different names, such as Schopf, JW; Kaltenegger, L.; Charbonneau, D. (Only Knoll, A. was on both lists.) Although, in sheer numbers the latter group was cited less, they had higher centrality scores, meaning that they were bridging different co-authorship networks. However, they were not being cited a lot, their influence in other domains was limited.

Table 7 The list of mostly cited authors by frequency and by centrality

By frequency			By centrality		
Author	Freq.	Cent.	Authors	Freq.	Cent.
DE Canfield	105	0.02	JW Schopf	32	0.35
JF Kasting	85	0.15	L Kaltenegger	27	0.28
RP Butler	67	0.15	D Charbonneau[a]	41	0.23
AD Anbar	61	0.08	AH Knoll	57	0.22
Mayor M[a]	61	0.05	GD Cody	53	0.22
P Ehrenfreund	58	0.10	A Boss	41	0.21
AH Knoll	57	0.22	SA Sandford	52	0.19
J Farquhar	55	0.05	HF Levison	47	0.19
EB Ford	53	0.08	SJ Kenyon[a]	26	0.19
MJ Mumma	53	0.04	JW Head	43	0.17

[a] Not a member of CAN 4 or 5 teams

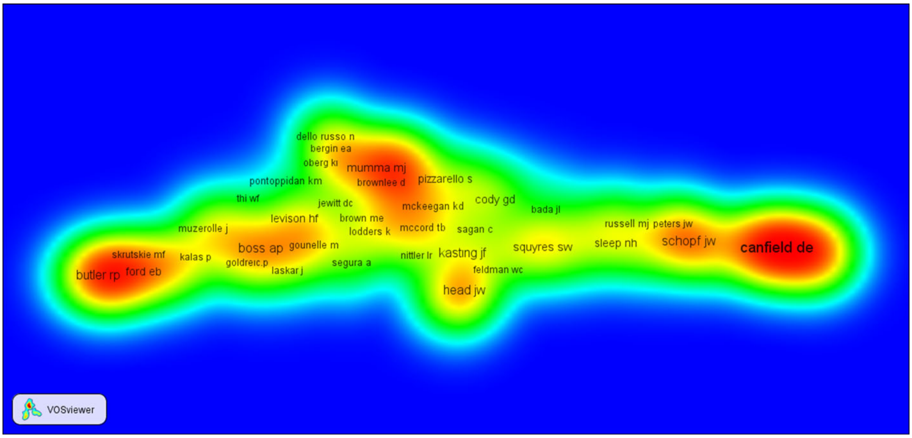

Fig. 6 Co-citation map of mostly cited authors

Figure 6 below is the heat map for co-citations of mostly cited authors (the bigger the font, the more publications from that author. The color groupings means co-authorship) (Table 8).

The publications that were cited the most by the NAI authors and the publications among the cited that had the highest centrality score are different (except for two publications). This suggests that, except for the two, the publications that were cited by the most were only cited by a certain co-authorship networks and they were not diffused to the rest of the greater NAI network. The tables for the most cited (Table 6) and the highest centrality scores (Table 9) are as follows:

Analysis of gathered citations for NAI-funded publications

When we applied the network analysis to the citations that NAI publications received, we found that the authors that cite NAI publications were quite distributed, that the density of the network was low (inbetweenness centrality density = 0.0113), and that there were not

Table 8 The list of top-10 most cited articles (by frequency)

Freq.	Cent.	Title	Journal
39	0.12	Attaining Doppler precision of 3 m s^{-1}	Publications of the Astronomical Society of the Pacific
39	0.01	Spectroscopic properties of cool stars (SPOCS). I. 1040 F, G, and K dwarfs from Keck, Lick, and AAT planet search programs	Astrophysical Journal Supplement Series
35	0.15	A hybrid symplectic integrator that permits close encounters between massive bodies	Monthly Notices of the Royal Astronomical Society
35	0.06	Tracing the stepwise oxygenation of the Proterozoic ocean	Nature
34	0.07	Research article—Comet 81P/Wild 2 under a microscope	Science
33	0.07	Catalog of nearby exoplanets	Astrophysical Journal
32	0.28	Habitable Zones Around Main-Sequence Stars	Icarus
31	0.02	A new model for Proterozoic ocean chemistry	Nature
29	0.21	The origin and evolution of chondrites recorded in the elemental and isotopic compositions of their macromolecular organic matter	Geochimica et Cosmochimica Acta
29	0.19	Interstellar ice: the infrared space observatory legacy	Astrophysical Journal Supplement Series

Table 9 The list of cited articles by the top-10 highest centrality scores

Freq.	Cent.	Title	Journal
26	0.41	A whiff of oxygen before the great oxidation event?	Science
28	0.37	Endogenous production, exogenous delivery and impact-shock synthesis of organic molecules: an inventory for the origins of life	Nature
19	0.30	In situ evidence for an ancient aqueous environment at Meridiani Planum, Mars	Science
32	0.28	Habitable zones around main-sequence stars	Icarus
15	0.27	A revised, hazy methane greenhouse for the Archean Earth	Astrobiology
18	0.23	Quantitative organic and light-element analysis of comet 81P/Wild 2 particles using C-, N-, and O-mu-XANES	Meteoritics & Planetary Science
16	0.23	The loss of mass-independent fractionation in sulfur due to a Palaeoproterozoic collapse of atmospheric methane	Geobiology
16	0.22	Organic haze, glaciations and multiple sulfur isotopes in the Mid-Archean Era	Earth and Planetary Science Letters
29	0.21	The origin and evolution of chondrites recorded in the elemental and isotopic compositions of their macromolecular organic matter	Geochimica et Cosmochimica Acta
14	0.20	Detection of thermal emission from an extrasolar planet	Astrophysical Journal

any dominant co-authors in the network. This might be because of the diverse nature of astrobiology in the sense that certain co-authors are followed by researchers in certain fields. Yet, in both of the maps prepared by VosViewer and CiteSpace, there were some prominent authors. CiteSpace identified five main clusters. Sean Solomon, Geoff Marcy,

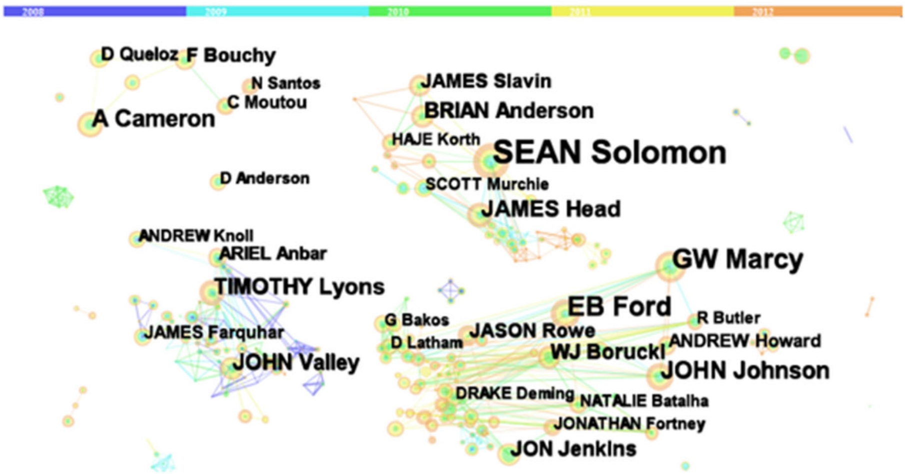

Fig. 7 Co-authorship clusters

Table 10 The list of cited articles category by the top-10 highest centrality scores

Journal category	Freq.	%
Astronomy & Astrophysics	8753	38.1
Science & Technology—Other Topics	2204	9.6
Geochemistry & Geophysics	2198	9.6
Geology	1128	4.9
Zoology	906	3.9
Life Sciences & Biomedicine—Other Topics	860	3.7

John Johnson, John Valley, Timothy Lyons, and Andrew Collier Cameron are pivot nodes among the five clusters (see Fig. 7). While four of the five clusters are somehow connected, Cameron's cluster is not connected to the rest of network.

The subject categories of the citing publications

Unfortunately the dataset is inconclusive for this type of analysis since we only have $\sim$65 % of the publications associated with a category in our data set. However, the order is similar to the NAI-funded publications list. Astronomy & Astrophysics dominates the list, followed by Geo-sciences and Science & Technology—Other Topics. Life Sciences (and Zoology) is a very small category (see Table 10).

Discussion and limitations

1. Table 1 demonstrates that NAI researchers prefer publishing in high-impact multidisciplinary journals (for publications between 2008 and 2012). As "conducting, supporting, and catalyzing collaborative interdisciplinary research" (NAI 2013a) is the very first goal of NAI, having more than half of their publications published in high

impact multidisciplinary journals for a newly established field (Blumberg 2003) is clearly an achievement. As for the domains, astro- and geo- sciences[5] dominate the field. Bio- science researchers might be publishing in geo- science journals and reaching out to a broader audience but there is not enough data to follow this thread.

2. Publications in bio- sciences tagged journals are more interdisciplinary; however, this is probably a result of Web of Science journal categories. These preliminary findings were presented to the PIs and NAI Central staff. The follow-up discussions revealed the limitations of journal categories. For instance, two very different topics such as cosmochemistry and planetary science are tagged under Astronomy & Astrophysics (A&A) but they are as different as microbiology and evolutionary biology, which are tagged separately in Web of Science. This WoS tagging makes A&A seem less “interdisciplinary”. However, Web of Science categories are the standard categories used in bibliometric studies; thus, they are not that off the target. The results can be compared to studies examining other funding programs and agencies. Furthermore, a disconnect between A&A and bio- sciences is obvious. In this regard, Miller et al. (2014) study is an important contribution since the information bottleneck algorithm does not have any assumptions (meaning no Web of Science Subject Categories). These categories can be a hurdle in the analysis especially for an emerging multidisciplinary field where disciplinary boundaries are fuzzy. However, the more traditional bibliometric approach that we employed in this study helped us to look other dimensions, such as international & institutional collaboration. In a nutshell, Miller et al. (2014) looks at interdisciplinary collaboration deeper, whereas our study covers more grounds from publication inventory to citation behaviors to different collaboration patterns.
3. NAI-funded researchers collaborate with researchers outside the U.S. Although, there is no monetary funding for researchers outside the U.S., strong co-authorship relations exist between the researchers in and outside the U.S. The Institute initiated the development of an international partnership network (thirteen members as of September 2014). A great majority of co-authors are from the partner countries.
4. NAI-funded researchers collaborate with researchers outside their organizations—an indicator of being a virtual institute.
5. There are prominent co-authors in the NAI-network but none of them dominates the field. The reason for that might be the multidisciplinary and emerging nature of the field. Astrobiology is like an umbrella term to define a quite diverse domain. Researchers might be known out of their immediate domain but that is not enough for their publications to be cited by researchers outside their domain. Over time, hopefully, there will be more integration between domains as envisioned in the Astrobiology Roadmap (Des Marais et al. 2008).
6. This type of research depends on bibliometric datasets. In order to have a better understanding for astrobiology relevant research, better and bigger datasets are needed. In this study, we focused only on NAI funded publications between 2008 and 2012 that were mentioned in the NAI annual reports, which are publicly available on the Institute’s website (NAI 2013b). Moreover, researchers have been publishing with NAI funds since 2000. A longitudinal study can tell us more about what the trends are in astrobiology or whether researchers are focusing only on certain areas or not. In addition, there are other funding streams for astrobiology research—such as the Exobiology Program at NASA. Publications datasets from such programs need to be

[5] Here astro-, geo-, and bio- sciences are used in the broadest, most general sense.

integrated to the dataset and analyzed. Therefore, it would be better if the NASA Astrobiology Program establishes a database that covers all funding streams (e.g. NAI, exobiology, etc.) and longer periods (say since the start of funding streams). Another future work direction might be performing similar analysis so as to see if other multi- & interdisciplinary fields have similar publishing practices.

7. It has to be mentioned that this study is a quantitative study, it only looks at the frequencies and connections of publications and citations. However, from the information science literature we know that it is important to know whether a citation is positive, neutral, or negative (Moravcsik and ve Murugesan 1975; Oppenheim 1996). For instance, in the case of the arsenic-based life article (Wolfe-Simon et al. 2010), the citations are not always positive (Benner et al. 2013). Another point is the place of the citation—whether it is in the literature review or methodology or discussion setting. Qualitative analysis by experts on selected publications can provide new insights as well.

We are hoping that this study and future studies (that investigates longer periods, the outputs of other astrobiology funding streams, and qualitative bibliometric studies) will shed more light on the field, researchers, and publications. Researchers can identify potential collaborators and understand how their own research fits in the broader astrobiology research. Funding agencies can see their impact in interdisciplinary, collaborative astrobiology research and better assess the research that is being done by seeing the collaborations among domains (or lack thereof) and identify neglected or over-supported domains and so on.

Acknowledgments This study was supported by the NASA Astrobiology Institute. In addition, we would like to thank to Thomson Reuters for making their relevant bibliometric datasets available to us.

References

AIRFrame. (2014). *Astrobiology integrative research framework project*. http://airframe.ics.hawaii.edu/. Accessed January 28, 2014.

Al, U., Taşkın, Z., & Düzyol, G. (2012). Use of social network analysis in bibliometric researches. BOBCATSSS 2012. In *Information on E-motion* (pp. 40–44). BOBCATSSS 2012, Amsterdam, January 23–25, 2012.

Benner, S. A., Bains, W., & Seager, S. (2013). Models and standards of proff in cross-disciplinary science: The case of arsenic DNA. *Astrobiology, 13*, 510–513.

Blumberg, B. (2003). The NASA Astrobiology Institute: Early history and organization. *Astrobiology, 3*, 463–470.

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*(3), 351–374.

Brazelton, W. J., & Sullivan, W. T. (2009). Understanding the nineteenth century origins of disciplines: Lessons for astrobiology today? *International Journal of Astrobiology, 8*, 257–266.

Carpenter, M., Bauer, T., & Erdogan, B. (2009). *Principles of management*. Washington: Flat World Knowledge.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology, 57*, 359–377.

Chen, C. (2014a). *CiteSpace: Visualizing patterns and trends in scientific literature*. http://cluster.cis.drexel.edu/~cchen/citespace/. Accessed December 22, 2014.

Chen, C. (2014b). *The CiteSpace manual*. http://cluster.ischool.drexel.edu/~cchen/citespace/CiteSpaceManual.pdf. Accessed December 22, 2014.

Chen, C., Song, I.-Y., Yuan, X., & Zhang, J. (2008). The thematic and citation landscape of data and knowledge engineering (1985–2007). *Data & Knowledge Engineering, 67*, 234–259.

Des Marais, D., Nuth, J. A., Allamandola, L. J., Boss, A. P., Farmer, J. D., Hoehler, T. M., et al. (2008). The NASA astrobiology roadmap. *Astrobiology, 8*(4), 715–730.

Gowanlock, M., & Gazan, R. (2013). Assessing researcher interdisciplinarity: A case study of the University of Hawaii NASA Astrobiology Institute. *Scientometrics, 94*, 133–161.

Hall, K. L., Stokols, D., Stipelman, B. A., Vogel, A., Feng, A., Masimore, B., et al. (2012). Does team science add value? A bibliometric study comparing the productivity of NIH-funded team science center grants with single investigator driven grants. *American Journal of Preventive Medicine, 42*(2), 157–163.

Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside, CA: University of California.

Katz, J. S., & Hicks, D. (1995). *The classification of interdisciplinary journals: A new approach*. In M. E. D. Koenig & A. Bookstein (Eds.), *Proceedings of the fifth biennial conference of the International Society for Scientometrics and Informetrics* (pp. 245–254). Medford: Learned Information.

Klavans, R., & Boyack, K. W. (2006). Quantitative evaluation of large maps of science. *Scientometrics, 68*(3), 475–499.

Lancaster, F. W., & Lee, J. L. (1985). Bibliometric techniques applied to issues management: A case study. *Journal of the American Society for Information Science, 36*, 389–397.

Miller, L. J., Gazan, R., & Still, S. (2014). Unsupervised document classification and visualization of unstructured text for the support of interdisciplinary collaboration. In *Proceedings of the 17th ACM conference on computer supported cooperative work & social computing (CSCW 2014)* (pp. 1033–1042). 15–19 February, Baltimore, MD.

Moravcsik, M. J., & ve Murugesan, P. (1975). Some results on the function and quality of citations. *Social Studies of Science, 5*, 86–92.

Morillo, F., Bordons, M., & Gómez, I. (2003). Interdisciplinarity in science: A tentative typology of disciplines and research areas. *Journal of the American Society for Information Science and Technology, 54*(13), 1237–1249.

NAI. (2014). International partners. https://astrobiology.nasa.gov/nai/international-partners/. Accessed 16 March 2015.

NASA Astrobiology Institute Website. (2013a). https://astrobiology.nasa.gov/nai/about/. Accessed February 1, 2014.

NASA Astrobiology Institute Website. (2013b). https://astrobiology.nasa.gov/nai/reports/annual-reports/. Accessed April 23, 2014.

National Research Council. (2004). *Facilitating interdisciplinary research*. Washington, DC: The National Academies Press.

National Research Council. (2008). *Assessment of the NASA Astrobiology Institute*. Washington, DC: The National Academies Press.

Oppenheim, C. (1996). Do citations count? Citation indexing and the research assessment exercise (RAE). *Serials, 9*(2), 155–161.

Porter, A. L., Cohen, A. S., Roessner, J. D., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics, 72*(1), 117–147.

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics, 81*(3), 719–745.

Rafols, I., Porter, A. L., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology, 61*, 1871–1887.

Slonim, N., Friedman, N., & Tishby, N. (2002). Unsupervised document classification using sequential information maximization. In *Proceedings of the 25th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 129–136). New York, USA.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science, 50*(9), 799–813.

Stokols, D., Hall, K. L., Taylor, B. K., & Moser, R. P. (2008). The science of team science: Overview of the field and introduction to the supplement. *American Journal of Preventive Medicine, 35*(2), 77–89.

Thomson Reuters. (2014). *Journal citation reports*. http://wokinfo.com/products_tools/analytical/jcr/. Accessed April 12, 2014.

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84*, 523–538.

van Eck, N. J., & Waltman, L. (2013). *VOSviewer manual*. http://www.vosviewer.com/documentation/Manual_VOSviewer_1.5.4.pdf. Accessed December 22, 2014.

VosViewer. (2013). http://www.vosviewer.com/. Accessed December 22, 2014.

Wagner, C. S., Roessner, J. D., Bobb, K., Klein, J. T., Boyack, J. W., Keyton, J., et al. (2011). Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature. *Journal of Informetrics, 165*, 14–26.

Williams, C. J., O'Rourke, M., Eigenbrode, S. D., O'Loughlin, I., & Crowley, S. (2013). Using bibliometrics to support the facilitation of cross-disciplinary communication. *Journal of the American Society for Information Science and Technology, 64*(9), 1768–1779.

Wolfe-Simon, F., Blum, J. S., Kulp, T. R., Gordon, G. W., Hoeft, S. H., Pett-Ridge, J., et al. (2010). A bacterium that can grow by using arsenic instead of phosphorus. *Science, 332*, 1163–1166.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science, 316*, 1036–1039.