Dear editor,

Thank you for taking the time to facilitate the reviews. Thanks to the reviewers for taking the time and effort to constructively comment on our manuscript. It helped us improve it considerably for this resubmission.

Based on the reviewer comments, we thoroughly revised the manuscript to improve the language. We also noted that the reviewers want to see a more thorough description of how "as-you-go" reporting would look like in practice. We extended this somewhat in the network structure to the extent possible, but note that the technical implementation details require a wholly separate manuscript. We also added a set of references to be more complete in our literature coverage.

We also made some additional changes upon revising the manuscript. These include:

- Replace 'metrics' with 'indicators' to be more in line with the Metric Tide report
- Replace references to 'pieces' and 'small units' with 'modular' to be consistent with the manuscript detailing the technical details of the infrastructure

We provide a point-by-point response to the reviewer's comments below. The tracked changes of the revisions are available in the history of our GitHub repository and can be viewed at

We look forward to your response.

Cheers,
Marino van Zelst and Chris Hartgerink

Reviewer 1

S1: The paper present interesting ideas that challenge current models of scholarly publications, the "as-you-go" approach.

Thank you for the kind words 💁

S2: However, it then mainly focuses on network analysis for research evaluation. Network analysis is already widely used for some of the purposes described in the 'use cases', especially when looking at academic recruitment to existing research groups. Just adding a complex measure, which would be very opaque to most, does not appear to be helpful to academic recruitment.

Current network analysis is indeed widely used and focuses on citation networks or author networks. In this manuscript we propose a network of direct descendancy

between the various pieces of information that are contained in traditional articles. As such, this would add an additional layer to the network analysis. We agree we did not differentiate this very well and therefore have tried to clarify this by adding the following to the network structure section

"Incorporating the direct descendancy of these knowledge modules builds a different kind of network than citation and authorship networks. As such, this network would be an addition to these already existing citation and authorship networks; it does not seek to replace them."

S3: We all agree that current measures are weak and, in some cases, deceitful, but these are normally tempered through the peer review system. In the UK, the largest national research assessment exercise (REF) is still a peer review system.

We do not contend systems like the REF (and similar), nor do we want to undermine them. We want to provide additional tools to assist in the evaluation process that surpass the weak measures currently applied.

S4: I think the focus, in the central part of the paper, on the evaluation side detracts from the interesting "as-you-go" approach which should be explored in more detail. The question that needs to be addressed is why this approach will result in 'better, higher quality research' and not 'why it will make research metrics and evaluation more rigorous'.

We agree with the reviewer that domains such as research quality are important and would like to share that these are part and parcel of the underlying ethos with which we wrote this manuscript. However, we also agree that this should get a more prominent place in the manuscript. We have introduced a section in the introduction relating the proposed communication infrastructure to the functions of scholarly communication:

"A scholarly communication system is supposed to serve five functions, but can do so in a narrow sense as it currently does, or in a wider sense. These functions of the scholarly communication system are (1) registration-, (2) certification-, (3) awareness-, and (4) archival [@roosendaal1997], and (5) incentives [@10.1045/september2004-vandesompel]. A narrow fulfillment of for example the registration function would mean that findings that are published are registered, but not all findings are registered [e.g., due to selective publication; @10.1126/science.1255484]. Similarly, certification is supposed to occur through peer review, but peer review can exacerbate human biases in the assessment of quality [e.g., statistical significance increasing the perceived quality of methods; @10.1007/bf01173636]."

We also added a section on how this network structure could provide a wider interpretation of these scholarly functions, at the end of the network structure section.

"A continuous and network based communication system could take a wider interpretation of the scholarly functions it is supposed to serve [@roosendaal1997; @10.1045/september2004-vandesompel]. Registration would become more complete, because selective publication based on results is preempted by embedding communication before any results are known. Certification is improved by embedding the chronology of a research cycle into the communication of research, ensuring that predictions precede results [@10.1073/pnas.1708274114]. Awareness is improved by using open by design principles, whereas awareness is now limited by financial means to access scholarly articles [@10.12688/f1000research.12037.3]. Archival would not only be simplified with peer-to-peer protocols, but also allows anyone to create a copy and could result in excessive redundancy under the Lots Of Copies Keeps Stuff Safe principle [LOCKSS; @10.1045/june2001-reich]. In the next sections, we extend on how incentives could be adjusted in such a network structure, to facilitate both the evaluation of research(ers) and the planning of research."

The goal of this manuscript subsequently extends more on the incentive function, considering that this is our main goal for this manuscript (see also our response to S8).

S5: I suggest: In the introduction, focus on the weaknesses of the current publication cycle and how a new approach will improve the quality of the production and dissemination of research

See response above (S4).

S6: What is the role of peer review in the new approach?

This is a good question. Peer review, or peer discussion, could be facilitated by more timely communication, where discussion can still improve theory, design, hypotheses, etc. We believe that peer review could improve in a modular system, as research can also improve 'on-the-go' and not only after the research has been fully conducted. If review only occurs "after-the-fact", insightful comments cannot be incorporated because the research was already conducted. How to technically implement this discussion, we consider efforts such as dokie.li to be very promising..

S7: Asses if the new approach will apply across all disciplines, I am not sure it will especially in the humanities. Which research will accrue most benefits from this approach

This proposal is initially aimed at empirical research, but has relevance for any area of research where provenance of information is relevant (e.g., anthropological research, archival study). We also added a short section in the manuscript on this "Additionally, "as-you-go" modules could be used in scholarly fields where the provenance of information is important to findings."

S8: Treat metrics and evaluation as an important by-product but maybe not the focus of all benefits

The is a Catch-22 for us: Our focus is what indicators could be fruitful in an "as-you-go" based scholarly communication system and the reviewer proposes us to shift our scope. We agree there are other benefits and tried to bring this to the forefront more (see also our response to S4).

S9: Use Cases have too many sweeping statements. It would be more interesting to work through a research topic. For example, there are quite a few references to data and codes, do you have a real case where this approach was used? What benefits did it produce compared to traditional publication models?

We don't have real data, and therefore sketch a situation where we would. We tried to mitigate the number of sweeping statements we recognized.
This is preliminary work for providing evaluation measures at the point in time that this infrastructure, which the first author is working on, is in place and gets used by individuals. We will report on real data separately once we have these.

S10: Quite a few points I have made above are touched upon in the 'discussion' section. These should be explored in the main section of the paper

We shifted around some of these points per the reviewer's suggestion, as can also be seen in our response on S4.

S11: A 'conclusion' section should be added.

Added:

"The current scholarly communication system based on research articles is "after-the-fact" and can be supplemented by an "as-you-go" based communication system. By doing so, the functions of a scholarly communication system can be interpreted more widely, making registration complete, certification part of the process

instead of just the judgment of peers, access to everything for everyone based on peer-to-peer protocols, simplify archival, and facilitate incentive structures that could align researcher's interests with that of scholarly research."

S12: The paper is well written and easy to read.

Thanks!

S13: I hope the comments above will be helpful. As I said at the start, this is an interesting and topical area that should be explored further.

We appreciate the time that went into writing this review and the constructive manner it is written. We hope the reviewer agrees with our scope limitation, because this work is connected to a larger vision that we can hardly describe in one manuscript.

Reviewer 2

S1: This is a very interesting concept. I thought that the point about using network analysis to compare a research group at University X with similar groups at other universities was excellent.

Thank you! 💁‍♀️

S2: However, the presentation lacks some of the rigor of a scholarly journal article. For example, the use of the first person tense (I/we) is generally frowned upon.

We are used to writing in the active tense from our respective fields, hence, did so here as well. We understand that this can be field dependent and make the stylistic decision to stick with the active tense.

S3: The use of "severe" to describe the extent / nature of threats (l.13) and changes (l. 41) is not appropriate; you should pick adjectives which are more qualitative.

We removed all mentions of "severe".

S4: I assume that with the use of "things" (l. 60), you are actually referring to research outputs? Obviously my point is that you should ideally aim for a more scholarly use of language.

We replaced "things" with research outputs and took the reviewer's comment into account when revising the manuscript.

S5: Broad, unsubstantiated assertions: At the beginning (l. 36), you mention "high rejection rates". According to whom?

Added a reference to the APA summary of 2016 that shows an average rejection rate of 71%.

S6: You mention that "the communicative design of papers embeds hindsight and its biases" (l. 74). Who says so? I would suggest adding some appropriate references.

We agree with the reviewer and refer to our response to Reviewer 1, S4.

S7: Literature review: You do mention major references. However, I believe that you should have a brief literature review section at the beginning.

We've added appropriate references where needed. We also added some crucial references that we indeed missed.

S8: Definition of concepts: You have based your concept on the fact that the current scholarly communication is "paper" based. You should define "paper" in this context since it obviously can also refer to a physical format, which I assume is not your intention.

We adjusted our first mention of the research paper to highlight the scope of what we indicate as a research paper

"The current scholarly communication system is primarily organized around researchers who publish static (digital) research papers in journals."

S9: l.53: "work that is not aptly captured by papers receives less recognition ...": Not only a reference required but also a slightly more detailed explanation, please.

Adjusted this sentence to
"Additionally, work that is not aptly captured by the authorship of papers is likely to receive less recognition (e.g., writing software code) due to reward systems counting publications instead of contributions (see also https://perma.cc/MUH7-VCA9)"

S10: Figures: While I appreciated the simplicity of Figure 1 (connected research stages), I was hoping for a slightly more in-depth visualization of what a network might look like. The mention of the Dat Project (protocol) was helpful for more context but did not further enlighten me as to what a scenario might look like. (Some readers are "visual people").

We include a reference to another manuscript by the first author that is more detailed on the technical details. We added another panel to Figure 1 to provide a more extensive Directed Acyclic Graph example.

Some random questions which sprang to mind:

S11: **l.136 (forward-looking evaluation): Would / could researchers "game the system", i.e. crunch the numbers and then manipulate their (yet to be produced) outputs to create a (positive but) artificial result?

Absolutely and we expect people to do so. The first author is planning to have explicit testing rounds to try and game the system and learn more about how that would work (a form of penetration testing). We hope that the diversity in questions and operationalizations regarding research evaluation mitigate the scope of gaming as we currently see with the Journal Impact Factor and citations (see also this paper).

S12: **l. 208: (computed density): The focus tends to be on the number of outputs. Where does the quality of the publishing outlet come into play? For example, how does the head of group know that applicant X, whose percentage points exceed the minimum, has not published in so-called "predatory" outlets so as to conflate their numbers?

Journals are not part of the proposed infrastructure and therefore are not discussed. In other words, there is no outlet in this infrastructure. As a result, the quality of the outputs is intended to be based on the quality of the outputs instead of the reputation of the outlet.

S13: **l.236-241 (central finding): Surely this approach could also be used to spark new ideas?

Yes :) We propose one use case but there are many more. We added a note in the discussion on this:
"The use cases presented for the various target groups (e.g., universities) can readily be transferred to suit other target groups (e.g., individuals)"

S14: **l.285 (rehash the same theory): In terms of replication, for example, does not the system rely on the reader following ALL relevant nodes? Is that realistic?

The reader could follow all relevant nodes to N-degrees of separation. The larger the distance, the less immediately relevant. Following the nodes through to the root node might become unrealistic to a human, yes. The idea is that machines can assist in this process because all content would be open by design. Furthermore, we expect that organizations would be able to offer services to academics and administrators that can help them sieve through the vast web of research output. Additionally, the current system of citations technically implies that readers could go down the rabbit hole, which is even less realistic because a lot of the research papers are inaccessible.

S15: **l.279 (communication of data): How would you combat the reluctance of researchers to publish first their data for fear of being gazumped / scooped by another researcher before they have the opportunity to publish their results?

This is a difficult issue we have thought about to great length and consider this fear of sharing similar to “security by obscurity”. After all, it seems to us more likely to get scooped if nobody hears your ideas (we informally call this the scooping paradox). From our own anecdotal experience, this is a mindset that can quickly be ameliorated by shortening feedback loops for sharing. That is one of the design principles of this modular infrastructure.

S16: **l.288 (papers): What evidence do you have that publishers (especially the Big Five) would treat the "as-you-go" units as preprints and therefore publish an "after-the-fact" version?

Absolutely no evidence whatsoever. We only know that preprints are increasingly accepted by journals for publication and that this is promising for these units (10.1126/science.aaq0747), which are broken up preprints in essence. We look forward to finding this out in practice.

S17: ** (question): Is this model you propose suitable for ALL disciplines? Do you foresee any challenges in some disciplines?

Please see our comments on Reviewer 1, remark S7, which addresses a similar issue.

We thank the reviewer for the insightful questions, which we appreciate greatly and will help us in our further efforts.

Reviewer 3

S1: This paper, “As-you-go” instead of “after-the-fact,” presents a proposal for breaking apart the unitary journal publishing model into chains of entities, connected via links, to enable network-based analysis and metrics of scholarly work. The paper is clearly written and easy to read. The limitations of current journal publishing models are used to motivate the proposed new approach to disaggregating scholarly work into discrete entities that can be represented via a network model.

Thank you for the kind words! 💁

S2: The proposed “as-you-go” model is thought-provoking and if widely adopted could present some of the opportunities described in the paper for new kinds of analysis and scholarly metrics. The paper as written, however, is very light on practical implementation details and does not reflect enough on the reasons why the current state of journal-centric scholarship has maintained its dominance. I provide a few suggestions and pointers that could strengthen the overall paper.

We consciously chose to exclude practical implementations in order to focus the paper’s scope on the benefits of when such an infrastructure is already in place. We realize that this can be unconvincing to readers and the first author has since submitting this manuscript drafted a more

extensive technical description that we now reference in the paper for interested readers (see also https://dat-com-chris.hashbase.io).

S3: Why journals are dominant? In the introduction, the paper presents a discussion of problems with the current scholarly publishing model. These problems are significant, and do limit the kinds of projects and products that scholars are free to produce. The paper does not, however, reflect in any length on the flip side of this issue, namely, with all of these problems, why are traditionally structured journals still dominant? Proposals to radically (or modestly) reinvent journal publishing models have been promulgated since the 1990s, along with predictions of the death of traditional journals (see Odlyzko 1995).

The reviewer has a valid point. We don't intend to undermine the journal by proposing this system (although we do admit we'd be very happy if that were the outcome). Please see also our response to Reviewer 3 S19. Hence, we do not include a discussion of why journals remain dominant because the scope of the paper is to propose a new, additional model instead of a zero-sum model to replace journals. Plus, this is such an intricate issue we don't think we can do it justice by adding a paragraph or two (there are social, political, historical, legal, financial reasons that go beyond what is best for science).

S4: The idea of breaking journal publications into components or modules has likewise been around for at least a couple of decades (Kircz 1998). Reviews of this early literature can be found in a few places (Kling & Callahan 2003; Borgman 2000, chap 4).

We agree and certainly don't intend to claim novelty on modular communication and also admit that we omitted crucial references like Kircz, which we were aware of but simply forgot to mention. We thank the reviewer for the suggestions and have added these into the manuscript where fitting and added the suggested books to our reading list.

S5: Journals have switched largely to electronic form, but otherwise, despite these many studies and proposals, traditional journals still remain the primary mode for the publication of scholarly work. It is important to reflect on why this is.

See our response to Reviewer 3 S3.

S6: Scholarly communication research discusses how the journal publication model provides significant benefits, and serves to fulfill several key functions necessary for scholarship to thrive, e.g. legitimization, dissemination, access, preservation, and curation (Borgman 2007, chap 4; see also Van de Sompel et al 2004).

We agree. Based on comments from the other reviewers we added a section on this (see also our response to Reviewer 1 S4).

S7: For example, the last few paragraphs of the present paper provide a bit of detail on Dat's features around versioning and replication. But replication is not preservation or curation. Just because bits are safe and duplicated does not mean that the bits are useful or will have value in the future.

We concur with the reviewer on preservation. We added
"However, preservation is a social process that requires commitment. Hence, a peer-to-peer infrastructure would require committed and persistent peers (e.g., libraries) to make sure content is preserved."

S8: I think the paper's proposal would be much stronger if it directly addressed these core functions of scholarly communication. It would also be worth considering these functions in relation to the question of what might be lost (for the better or worse) in a switch to the "as you go" model.

We added these, please see also our response to Reviewer 1 S4.

S9: 2. How would this new model work? The paper is very short on practical details of how the "as you go" model might be implemented in practice. I realize that this is not the goal of the paper, but without more detail on how this would work in practice, it is hard to take the proposed model too seriously.

We now refer the reader to a more technical description that has been drafted since the submission of this manuscript.

S10: Significant practical impediments stand in the way of some of the characteristics of this model, e.g. it is currently challenging to identify relationships between scholarly entities. It is also difficult to characterize such relationships, validate them, and preserve them.

We look forward to tackling these barriers in various testing rounds with researchers. The model of descendancy is primarily intended to flow logically from one module directly to another and allows researchers to make these direct links (instead of backwards identifications), with citations still occurring within the modules. Validation and preservation of these relations is technically feasible and is described in the more technical piece on the decentralized scholarly communication infrastructure.

S11: Various initiatives focused on "data citation," "software citation", etc, are working on these issues now, but progress is slow and piecemeal at best (for the core issues at play, see the 2013 report by the Task Group on Data Citation Standards and Practices).

Thanks for that reference! A very interesting report we'll definitely take into account moving forward.

S12: I suggest that the authors provide more detail on how the Dat protocol works, perhaps giving an example, or at least fuller explication of the ways that this scheme would work in practice.

See our previous reply to S2 (for reviewer 3).

S13: Research is not linear.

We wholeheartedly agree.

S14: The paper repeatedly uses hypothesis-driven research as the exemplar for how the "as you go" model would work. But what proportion of scholarly research is actually hypothesis driven?

We do not know the exact proportion. The hypothesis driven model is the starting point, but we note that all provenance based research could be communicated this way (see also Reviewer 1 S7).

S15: Likewise, preregistration is given as an example of something that could be included in the as-you-go networks. I don't know of any field outside of biomedicine where preregistration takes place.

From our personal experience, we know that registered reports and preregistration are also increasingly applied in the social sciences. Furthermore, 96 journals (as of March 29 2018) offer Registered Reports across various fields (https://cos.io/rr/).

S16: Maybe the authors think it should be done in more fields. If so, the practicalities of doing this should be discussed. Many research projects are exploratory, without having a predefined hypothesis, and many start with a hypothesis but end up far afield for a variety of reasons.

The preregistration debate is covered extensively in other papers that have this discussion within their scope. We don't include that discussion here but include a note on this "[exploratory research could be communicated without hypotheses; for a more extensive discussion of the benefits and limits of preregistration see @10.1073/pnas.1708274114]"

S17: Research is also rarely chronological. It is usually highly iterative, oscillating between data collection, analysis, conceptualization, presentation, and back again in every conceivable permutation (and not always with clear logic or clear links/boundaries between the stages).

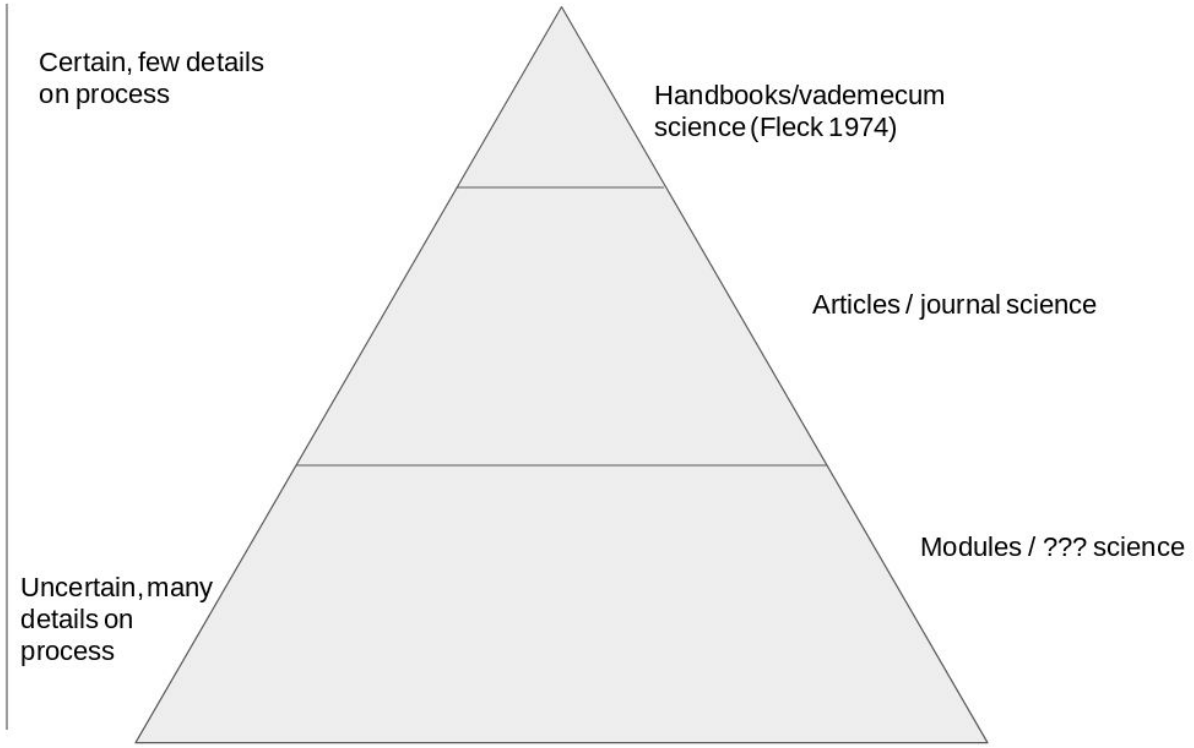
We agree.

S18: Does this "as you go" model work outside of chronological hypothesis-driven research? If so, would the resulting network maps be understandable or useful?

Yes. This is covered more in the technical details of the paper, but put simply: Each link in the network refers to a versioned folder. The versions are saved, but the folder can still be updated. As such, iterating between theory V1 => hypotheses V1 => theory V2 => hypotheses V2 => design V1 => theory V3 => ... is possible. These iterations are useful because they also show the process of knowledge generation to an extent that now is not captured by the research article. Moreover, if a researcher does not want to see these iterations they could be filtered away and just see the unique modules' latest registered versions.

S19: Journal articles are highly useful as a genre of communication because of their standardized layouts, linguistic conventions, and narrative forms. Reiterating my point above, what would be lost if those genre conventions would be lost?

Journals can co-exist if people feel they have a need for these (but see also Reviewer 3 S3). We see modules as a way of increasing uncertainty in claims but that provide more detail about the process. Conversely, we have handbooks that have more certainty but provide less insight into the process.



S20: How would this proposed "as you go" model replace those genre conventions, which streamline the cognitive load in reading articles.

See our response above. Additionally, because all content is open by design, reducing cognitive load is facilitated by unrestricted Text- and Data Mining (TDM) options on the registered content

(circumventing the copyright issues that currently restrict reducing cognitive load in journal articles).

WORKS CITED

Borgman, C.L. (2000). From Gutenberg to the Global Information Infrastructure. Cambridge MA: MIT Press.

Borgman, C.L. (2007). Scholarship in the Digital Age. Cambridge MA: MIT Press.

Kircz, J.G. (1998). Modularity: the next form of scientific information representation? Journal of Documentation, 54(2): 210-235. https://doi.org/10.1108/EUM0000000007185

Kling, R. and E. Callahan. (2003). Electronic journals, the internet, and scholarly communication. Annual Review of Information Science and Technology. 37: 127-177. https://doi.org/10.1002/aris.1440370105

Odlyzko, A. M. (1995). Tragic loss or good riddance? The impending demise of traditional scholarly journals. International Journal of Human-Computer Studies, 42(1), 71–122. https://doi.org/10.1006/ijhc.1995.1004

Task Group on Data Citation Standards and Practices. (2013). Out of Cite, Out of Mind: The Current State of Practice, Policy, and Technology for the Citation of Data. Data Science Journal. 12, pp.CIDCR1–CIDCR7. DOI: http://doi.org/10.2481/dsj.OSOM13-043

Van de Sompel, H., Payette, S., Erickson, J., Lagoze, C., & Warner, S. (2004). Rethinking scholarly communication. D-Lib Magazine, 10(9). https://doi.org/10.1045/september2004-vandesompel

Reviewer 4

S1: An interesting premise, this paper lacks a few nuances of the subject, but overall is thought provoking and novel.

Thank you!

S2: Some close reading of the paper shoudl be done before publication: e.g.,
The paragraph that runs from lines 230-235 jumps in the middle. "Katz centrality" should be defined.

We've added

“Katz centrality extends this and computes the centrality over N generations of child nodes [pp. 206-210; @Wasserman1994] whereas traditional in-degree centrality calculates centrality for $N=1$ generations.”
as a definition of Katz centrality and explained why this is another useful measure of impact aside from in-degree
“If multi-generation indicators are relevant, Katz centrality measures could provide operationalizations of such measures.”

We’ve also extended our argument of these tools as proposals
“The indicators we highlighted here are simple proposals. Indicators from network analysis and graph theory could be applied to the study of knowledge development when a network structure is available and we hope to see suggestions to answer questions about the network. These kinds of analyses are already done within citation networks [e.g., @10.1126/science.aao0185] and authorship networks [e.g., @10.1371/journal.pntd.0000501], but we cannot do so with the provenance or planning of knowledge generation in the current scholarly communication system.”

We hope on a proliferation of ways to evaluate research qualitatively and quantitatively. We are unsure what the reviewer refers to as a jump in the middle of the sentence.

S3: More discussion of pre-registration, how peer-review might fit into the system, and other safeguards for academic conduct would be interesting to fill out this idea as viable. Additional information about how the infrastructure would be supported or who would manage the nodes and analysis would also be of interest. Overall a very interesting topic.

See reviewer 1 S4 for additions we made on these topics.