Making it count: a computational approach to attribution

Kristi Holmes* - Northwestern University Feinberg School of Medicine

Karen Gutzman - Northwestern University Feinberg School of Medicine

Patty Smith - Northwestern University Feinberg School of Medicine

David Eichmann - University of Iowa

Melissa Haendel - Oregon Health & Science University

National Center for Data to Health (CD2H)

Abstract

People from a diverse array of backgrounds play important roles in research, often in ways that cannot be quantified through traditional metrics of scholarly impact. This demands better approaches to evaluate and ultimately communicate scholarly outcomes beyond the narrow criteria of publications and grants. It is imperative to develop a structure to track a much wider diversity of contributor roles and research objects - and do so in a manner that is easily populated with real data. This presentation will share details about the project, our team and approach, upcoming opportunities to collaborate, metrics for success, and integration and application of this work to date.

Introduction

Reproducible science depends in part on knowing what has been specifically performed and by whom. However, even as data scientists and software engineers play a greater role, they still lack a standard mechanism for attribution and subsequent citation of these individuals' creative contributions. Not only does this hamper reproducibility and evidence for scientific conclusions, it also disincentivizes these types of contributions, slowing research by making it less reliable - and encouraging some of our brightest contributors into more lucrative, and potentially less open and acknowledged fields.

Tracking contributor roles and research objects is an urgent need to support today's team-based, interdisciplinary work. In order for this to be successful, it is essential to do so in a manner that can easily leverage and populated with real data from real systems. This can be accomplished by computational models by which contributions and research objects can be recorded and the data made openly available to third-party sites such as scholar profiles, ORCID, publishers, funders, etc. Relationships between people and their products/activities can be used to track research trends, to understand and leverage influences or projects, to promote collaboration and team formation, to support research recommender systems, and to present a complete record of research. ***Fundamentally, the data about the contributions that scholars make should be as open as the data and resources themselves if we really aim to incentivize sharing and open science.***

Work to date

This project builds on strong preliminary workshops[1-2] and existing evaluation frameworks[3-8] which identified many activities and outputs for which people want credit. Here, we aim to define computational models for collecting and disseminating contributor attribution data for a wide range of scholarly object types. A first draft of a contribution ontology was completed in April 2016 and debuted at the FORCE11 2016 conference in the OpenVIVO project (http://openvivo.org/), where it has been used to annotate the contributions users have made to their various research outputs.[9] This first version of the ontology leveraged the CRediT Taxonomy[10] in combination with the outputs recorded in the previous workshops, providing a unique perspective on the work people do and want credit. We aim to enable our efforts to

integrate with multiple vocabulary frameworks such as the schema.org vocabulary. We note that all the above have been primarily volunteer efforts through communities such as the Force11 Attribution working group, the NISO Altmetrics working group[5], and the VIVO open source community.

An open invitation for collaboration

Standards development is largely community-driven, with stakeholders and systems working together with clear understanding that everyone needs to make progress toward the shared goal to achieve a successful outcome. We actively welcome broad input and collaboration on the data models, pilot implementations, metrics for success, and next steps from diverse stakeholders such as researchers, developers, funders, publishers, scholarly organizations, and agencies. The stakeholders and systems that must be engaged as part of an attribution effort are the same partners that have long demonstrated their commitment to improving the scholarly ecosystem through other community-driven works and we are grateful for ongoing and new partnerships. This presentation will introduce the project and our collaborators, our team and approach, share details about upcoming opportunities to collaborate, metrics for success, and integration and application of this work to date.

Acknowledgement

This work is being developed under the auspices of the National Center for Data to Health (CD2H), grant U24TR002306, from the National Center for Advancing Translational Sciences at NIH. The CD2H is charged with coalescing and coordinating informatics activities to provide collaborative research infrastructure and create an ecosystem for discovery and sharing of software, data, and other research resources.

---

[1] Measuring Success through Improved Attribution. Workshop at the 2015 VIVO Conference, Cambridge, MA. Slides available at http://goo.gl/64pvjn.

[2] Contribution and Attribution in the Context of the Scholar. Workshop at 2015 Force15 Conference, Oxford, UK.

[3] CRediT (Contributor Roles Taxonomy). (2016). Retrieved from http://docs.casrai.org/CRediT.

[4] CC Sarli, EK Dubinsky, and KL Holmes. Beyond Citation Analysis: A Model for Assessment of Research Impact. J Med Libr Assoc 2010 Jan;98(1): 17-23. PMCID: PMC2801963

[5] National Information Standards Organization. (2016). Outputs of the NISO Alternative Assessment Metrics Project. Recommended Practice RP-25-2016. Retrieved from https://goo.gl/n7JV2z.

[6] Panel on Return on Investment in Health Research, 2009. Making an Impact: A Preferred Framework and Indicators to Measure Returns on Investment in Health Research, Canadian Academy of Health Sciences, Ottawa, ON, Canada

[7] Academic Careers Understood through Measurement and Norms (ACUMEN). (2018). Retrieved from http://research-acumen.eu/.

[8] Researchfish, research impact tracking and reporting. (2018) Retrieved from http://www.researchfish.net/.

[9] Ilik V, Conlon M, Triggs G, White M, Javed M, Brush M, Gutzman K Essaid S, Friedman P, Porter S, Szomszor M, Haendel MA, Eichmann D and Holmes KL (2018) OpenVIVO: Transparency in Scholarship. Front. Res. Metr. Anal. 2:12. doi: 10.3389/frma.2017.00012

[10] Allen L, Scott J, Brand A, Hlava M, Altman M. Publishing: Credit where credit is due. Nature. 2014 Apr 17;508(7496):312-3. PubMed PMID: 24745070.