WORKSHOP AGENDA: Workshop on Supporting Scientific Discovery through Norms and Practices for Software and Data Citation and Attribution
Jan 29-30, 2015
Gallery 2 Room, Hilton Arlington

Goals: During the morning presentations, identify and capture challenges in software and data citation and begin to formulate how these can be made actionable. In the afternoon breakout sessions, select specific software and data citation use cases of interest. Formulate actionable outcomes as "critical asks" around these use cases for guiding infrastructure and methodology development and widespread adoption of norms and practices for software and data citation and attribution.

Day 1:	7:15 a.m. - 8:00 a.m.	Continental Breakfast
	8:00 a.m. - 8:10 a.m.	**Stan Ahalt** (RENCI, PI): Welcome: Overview, Goals and Desired Workshop "Actionable" Outcomes
	8:10 a.m. - 8:15 a.m.	**Knowledge Innovation Facilitation Team**: Workshop facilitation
	8:15 a.m. - 8:45 a.m.	**Luis Ibanez** (Featured Guest Speaker): Supporting Reproducible Scientific Research with Open Source Practices for Software and Data Citation and Attribution, a 15-Years Perspective and Vision for the Future
	8:45 a.m. - 9:00 a.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
	9:00 a.m. - 9:25 a.m.	**Dan Katz** (**NSF** Presentation): Metrics & Citation for Software (and Data)
	9:25 a.m. - 9:45 a.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
	9:45 a.m. – 10:10 a.m.	**Thomas Carsey and Jonathan Crabtree** (UNC Odum Institute for Research in Social Science): Automated Data Citation in the Social Sciences using the Dataverse Network Open-Source Software

10:10 a.m. - 10:25 a.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
10:25 a.m. – 10:40 a.m.	Break
10:40 a.m. – 11:10 a.m.	**Sweitze Roffel and Mike Taylor** (Elsevier): Linking data in and outside a scientific publishing house – a perspective from a publisher.
11:10 a.m. - 11:25 a.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
11:20 a.m. – 11:45 a.m.	**Jennifer Lin** (PLOS) **and Matt Jones** (UCSB, NCEAS): "Make Data Count" Open source software collecting metrics for data and software use.
11:45 a.m. - 12:00 p.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
12:00 p.m. – 12:15 p.m.	**James Howison** (Univ. of Texas Austin): How software is mentioned/cited in the biology literature.
12:15 p.m. - 12:30 p.m.	ALL: While you are listening to the presentations above, capture challenges that you see. Next, briefly discuss these as groups for 10 min and choose and put forward 1-3 actionable ideas.
12:30 p.m. - 1:30 p.m.	Lunch
1:30 p.m. - 2:00 p.m.	Organize and decide upon some number of use cases and challenges that will be used as motivators for rest of day. The decided-upon use cases will be allocated to 4-5 breakout groups in the afternoon and another 4-5 breakout sessions the next morning. A running list of currently contributed use cases is located on the workshop site https://softwaredatacitation.renci.org/Pages/Use-Cases.aspx.

2:00 p.m. - 4:00 p.m.		ROUND-1

Breakout Sessions: 4-5 breakout sessions will be held concurrently with one or more use cases decided upon in the morning session allocated to each breakout session. Attendees will self-select which breakout sessions they attend as a function of which use cases they are interested in actionably acting upon after the workshop. Each breakout session group should discuss the use cases assigned to their group through the lens of NSF topics proposed in their Dear Colleague Letter listed at the end of this agenda.

4:00 p.m. - 4:10 p.m.	Break
4:10 p.m. - 5:15 p.m.	Report Out from Breakout Groups
5:30 p.m.	Adjourn

Attendees are encouraged self-organize into multiple groups to go to dinner Thursday evening. Restaurant recommendations are provided on the workshop website: https://softwaredatacitation.org/Pages/Locations-and-Directions.aspx

Day 2:	8:00 a.m.	Continental Breakfast
	8:15 a.m. - 8:45 a.m.	Plenary Discussion
	8:45 a.m. - 10:30 a.m.	ROUND-2

Breakout Sessions: 4-5 breakout sessions will be held concurrently with one or more use cases decided upon in the morning session allocated to each breakout session. Attendees will self-select which breakout sessions they attend as a function of which use cases they are interested in actionably acting upon after the workshop. Each breakout session group should discuss the use cases assigned to their group through the lens of NSF topics proposed in their Dear Colleague Letter listed at the end of this agenda.

10:30 a.m. - 10:50 a.m.	Break

10:50 a.m. - 12:30 p.m.	Reports from Day 2 breakout sessions
12:30 p.m. - 12:40 p.m.	Workshop PI and NSF closing comments
12:40 p.m. - 1:30 p.m.	Lunch
1:30 p.m.	Adjourn
1:00 p.m. - 2:00 p.m.	Organizing committee generates report writing assignments, timeline, and due dates

NSF topics to address in breakout sessions from Dear Colleague Letter
(http://www.nsf.gov/pubs/2014/nsf14059/nsf14059.jsp):
Citation and attribution:
- Novel mechanisms for citation of software and datasets as distinct products of scholarship, promoting standards of academic credit and rigor for these cyberinfrastructure components
- Novel citation methods for new forms of publication and scientific expression so that researchers are able to ensure their work is citable, and others are able to discover and access it
- Citation patterns that include a role for citations (e.g. to value activities such as "data provider/curator" and/or "software tool provider" alongside "data analyzer" or "computational modeler"), which can help create a credit market for data and software sharing

Metrics of impact:
- Introduction of appropriate metrics that match the effort necessary for successful development and maintenance of scientific software and data frameworks
- Development and use of metrics that measure software and data framework usage and impact on science, engineering and education
- Establishment of metrics that recognize open access policies and sharing
- Comparison of impact for publication of software and data in a citable form before paper publication, as advocated in initiatives such as SageCite (http://blogs.ukoln.ac.uk/sagecite/), versus current practices
- Creation of specific project metrics that assess and monitor effective availability and accessibility of software and data
- Identification of sources of information about researchers' productivity and impact
- Development of ways in which researchers' scientific activity can be automatically captured and validated