



UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Collaborative provenance for workflow-driven science and engineering

Altıntaş, İ.

Publication date
2011

[Link to publication](#)

Citation for published version (APA):

Altıntaş, İ. (2011). *Collaborative provenance for workflow-driven science and engineering*.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

List of Tables

3.1	Summary of the characteristics of different scientific workflow provenance approaches.	39
5.1	Example queries across workflow executions and collaborations	60
6.1	Relation instances of the provenance schema corresponding to the example in Figure 6.2.	68
6.2	Basic QLP constructs and functions, where e_n is a node expression comprised of either a data artifact identifier, a run identifier, a data artifact type (denoting the set of artifacts having that type), or a workflow (denoting the set of runs of the workflow). We use p to denote a QLP path query, and r_i to denote a run. a is attribute for user collaboration graph, which can be either of these $\{\phi, N, W, S, NW, NS, WS, NWS\}$	75
6.3	Example queries expressed using the dependency functions defined for QLP. .	76
6.4	Mapping of basic Open Provenance Model nodes and edges to the collaborative entities and relationships.	78
7.1	Relation instances of the provenance schema corresponding to the example in Figure 7.4.	89
7.2	Example CAMERA queries 1 through 6 expressed in Datalog and QLP extensions.	93
7.3	Example CAMERA query 7 expressed in Datalog and QLP extensions. . . .	94
8.1	Collaborative provenance views for data dependency, run dependency and user collaboration expressed in PostgreSQL.	102
8.2	Example CAMERA queries Q1 through Q6 expressed in PostgreSQL.	103
8.3	Example CAMERA query Q7 expressed in PostgreSQL.	104
8.4	The size of database (in number of tuples) for different datasets.	105
8.5	The query execution time (in ms) for collaborative provenance views over datasets of 5, 10, 25 and 50 run dependencies.	105

- 8.6 The query execution time (in ms) for evaluation queries 1 through 7 over datasets of 5, 10, 25 and 50 run dependencies. 105
- 9.1 The publish and run observables in interoperable PC3 scenario. The contents of the table shall be read as follows: e.g., the second row refers to run_1 performed by u_2 using $wf_{preload}$ published by u_1 . In run_1 , u_2 used $d_{J062941}$ as an input and the run produced $d_{J062941-1}$ as its output. 112