

REPORT DOCUMENTATION PAGE

1. REPORT DATE 20220425		2. REPORT TYPE Final Performance Report		3. DATES COVERED	
				START DATE 20180312	END DATE 20211231
4. TITLE AND SUBTITLE Research Performance Report for "Provenance Analytics for Command & Control" project					
5a. CONTRACT NUMBER		5b. GRANT NUMBER N62909-18-1-2079		5c. PROGRAM ELEMENT NUMBER	
5d. PROJECT NUMBER		5e. TASK NUMBER		5f. WORK UNIT NUMBER	
6. AUTHOR(S) Huynh, Trung Dong and Moreau, Luc					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) King's College London The Strand London WC2R 2LS United Kingdom				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research Global ONR REG Office Boston, 495 Summer Street, Room 627 Boston MA 02210-2109			10. SPONSOR/MONITOR'S ACRONYM(S)		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for Public Release; Distribution is Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This is the final research performance report for the project entitled "Provenance Analytics for Command & Control". It summarizes the key findings of the project: Provenance Kernels (PK), Provenance Types Library, PROV-Template compiler, and Explanations for classification by PKs. The report also highlights the technology transfer of provenance modeling and the associated tools from King's College London to the research team at the Naval Information Warfare Center Pacific. Finally, it enumerates the project's outputs in terms of publications and dissemination.					
15. SUBJECT TERMS provenance, machine learning, command and control					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT		18. NUMBER OF PAGES
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU	UU		10
19a. NAME OF RESPONSIBLE PERSON Professor Luc Moreau				19b. PHONE NUMBER (Include area code) -44 207 848 6808	

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE.

Full publication date, including day, month, if available.
Must cite at least the year and be Year 2000 compliant, e.g.
30-06-1998; xx-06-1998; xx-xx-1998.

2. REPORT TYPE.

State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED.

Indicate the time during which the work was performed and the report was written.

4. TITLE.

Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

5a. CONTRACT NUMBER.

Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER.

Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER.

Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER.

Enter all project numbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER.

Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/monitoring agency, if available, e.g. BRL-TR-829; -215.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.



**Provenance Analytics for
Command and Control**

Final Research Performance Report (RPPR)

Award No. N62909-18-1-2079

Distribution Statement

- DISTRIBUTION A. Approved for public release: distribution unlimited.

Accomplishments

What were the major goals and objectives of the project?

The objectives of the project are the following:

01. To design provenance analytics techniques that combine domain-agnostic properties of provenance, such as network metrics and topology, with domain-specific information, to relate provenance logs about the creation of plans to quality of plans in terms of their outcomes.
02. To devise analytics techniques (both network-based and summarization-based) capable of mining provenance and extracting discrete activities performed by operators and unmanned vehicles, and compare the performance of these techniques.
03. In the context of unmanned vehicle management systems, to devise principles of accountability allowing the system and the analytics techniques to be accountable to their stakeholders.
04. To study the scalability of the techniques investigated in the project.

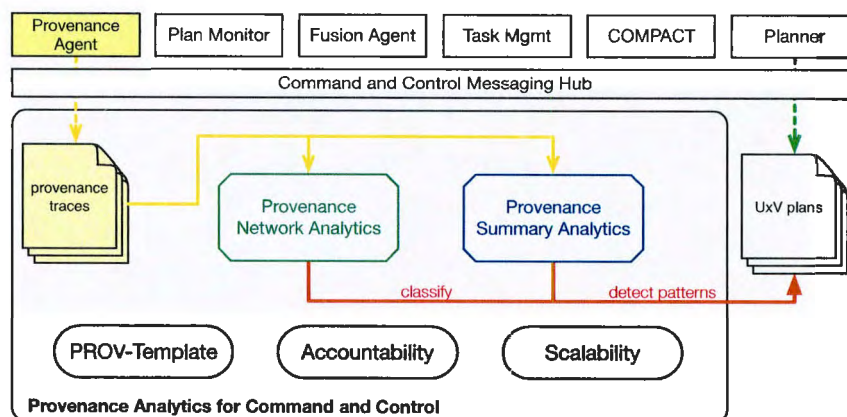


Figure 1. An overview of the proposed work in the PA4C2 project.

What was accomplished towards achieving these goals?

Key findings

1. **Provenance Kernels (PK)** [ANALYTICS] are novel graph kernels [1] developed specifically to characterize provenance graphs by exploiting the types of nodes and edges contained therein, or their *provenance types*. The kernels provide us with a generic tool to compare and classify provenance graphs and, by extension, data from several applications with high accuracy without the need for manual selection of graph features as required by our existing Provenance Network Analytics method [2] while being 100s faster (see Figure 3). A summary of provenance kernels and their applications are provided below, with the full results are reported in [3].
2. **Provenance Types Library** [ANALYTICS & SCALABILITY] is an efficient data structure to record provenance types, designed particularly to support the analysis of provenance data generated from

continuously running applications [4]. Not only that it compactly stores provenance types, but it also supports the incremental (re)inference of the provenance types of existing provenance graphs when new provenance records are added to them without the need for expensive re-computation. Therefore, a provenance type library makes provenance types required by provenance kernels readily available for the analysis, classification, and summarization of provenance data whenever the need arises.

3. **PROV-Template compiler** [SCALABILITY] is a software development tool to facilitate provenance instrumentation in software and applications via the PROV-Template approach [5]. The tool compiles provenance templates into (a) an application-side library to record the right application data for the specified templates in a conforming CSV representation and (b) a provenance-side library to expand such CSV data into the corresponding provenance graphs. In addition to expediting the provenance instrumentation process, the approach enables applications that generate large amounts of provenance data to record them efficiently in a compact CSV data format. Moreover, the provenance-side library can compute provenance types directly from the CSV data without the need to expand them into provenance graphs before so doing. The tool is made available in the open-source ProvToolbox¹ package.
4. **Explanations for PK classifications** [ANALYTICS & ACCOUNTABILITY] can be produced by combining a machine learning (ML) model built on PKs with the LIME method [6] to identify which provenance type(s) significantly contribute to a specific classification decision. A narrative in a natural language describing the graph patterns associated with such a provenance type can then be produced to help users of the ML model understand why certain classification is picked by the model. An illustrating example of this method is presented in [3].

Provenance Kernels

The proposal's **O1** and **O2** objectives planned to study provenance graph metrics [2] and provenance summarization [7] as the analytics techniques for IMPACT provenance data. However, graph metrics have to be arbitrarily selected without any established principle; furthermore, domain-specific information does not naturally integrate with graph metrics. Thus, principles of graph kernels [1] were used in combination with the concept of provenance types introduced in [7], allowing semantic information to be integrated into provenance analytics in a systematic manner. Graph kernels are tools used to capture similarities of networks according to different aspects of their structure (or graph patterns), expressed in terms of type information associated with graph nodes and edges. Such aspects can be identified, or "discovered", by graph kernels without relying on human selection of graph features like the PNA method. Those can be associated with the number of nodes in a graph or their connectivity, or to the set of possible paths or subtrees encountered in them, as well as more abstract notions such as the Lovász number, central in graph theory and combinatorial optimization. Once a table of similarities between graphs is calculated with the aid of a graph kernel method, traditional machine learning algorithms can be applied to those graphs, for example, to classify them according to a particular characteristic.

¹ Available at <https://lucmoreau.github.io/ProvToolbox/>.

As part of this project, to analyze provenance graphs, we have developed a notion of provenance types that captures (1) provenance semantics, (2) local graph structures, and (3) application-domain information from provenance graphs. Provenance types are, in loose terms, an encoding of the neighborhood of nodes in a graph in such a way that each node is assigned a *type* that summarizes the history of that provenance expression. Based on such provenance types, we built the Provenance Kernel method, which involves comparing provenance graphs according to how many of such neighborhood structures (i.e., provenance types) they have in common. Provenance Kernels can then be used similarly to traditional graph kernels for the classification or clustering of provenance graphs (to gain insights from them as shown in the Applications below). As part of this line of work, we *designed* an algorithm to infer provenance types in *linear time* with respect to the size of input graphs. Thanks to this, the computing time of Provenance Kernels is 100x faster than that of the network metrics employed by PNA on the same graphs (O4). We showed that Provenance Kernels not only give us high accuracy in prediction tasks (see the Applications section) but also can be done incrementally as new data becomes available, as discussed in more detail in [4].

Objective O3 was addressed by defining a set of templates for provenance for unmanned vehicle management systems [8]. Combined with the methodology “Explainability by Design” conceived in a sister project², they allow for explanations of the systems to be generated to address the user’s needs. In addition, the provenance kernel approach was shown to be able to provide explanation of the most significant features of a graph, in term of semantic types and the PROV provenance vocabulary [3].

Applications

We have applied Provenance Graph Kernels in several suitable provenance data sets available to us to build predictive models for data in those applications. Specifically, we have successfully applied the method to predict (see Figure 2):

1. whether a patient’s admission to a hospital resulting in a fatality (Figure 2) based on the provenance graph recording their journeys through the hospital – the provenance graphs were built from logs of critical care data in the MIMIC-III dataset [9];
2. the data quality label of buildings, routes, and route sets crowdsourced in the CollabMap application [10] from their provenance; and
3. the team to which a player belongs based on their actions, as recorded in provenance graphs, in Pokémon game simulations, where each team employs a different strategy.

² The PLEAD project – Provenance-driven & Legally-grounded Explanations for Automated Decisions. See <https://plead-project.org>.

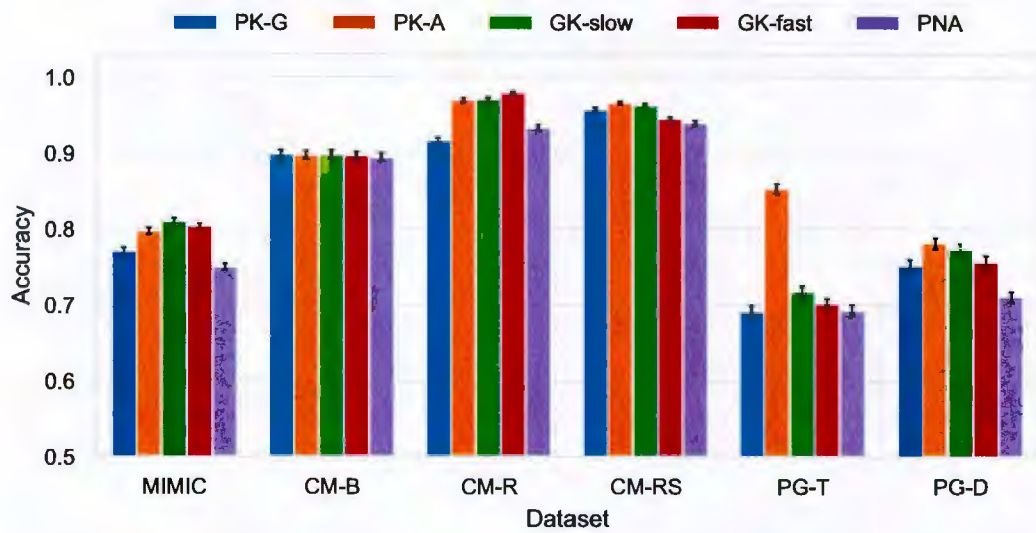


Figure 2. The accuracy (0.0 – 1.0) of two variants of provenance kernels (PK-G and PK-A), the best graph kernels, and the PNA method in six classification tasks from the MIMIC-III dataset [9], the CollabMap (CM) datasets [10], and the provenance of Pokémon game simulations (PG). The error bars show the 95% confidence intervals.

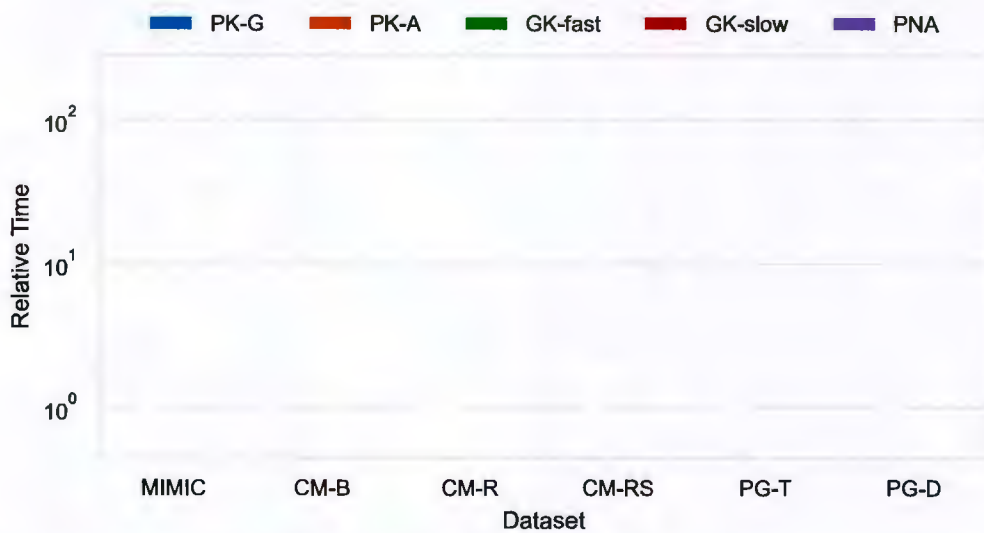


Figure 3. The relative time costs of the best-performing methods shown in Figure 2 plotted on the log scale. It shows that, in terms of computing cost, provenance kernels are competitive compared to the faster, existing graph kernels while significantly out-perform the PNA method in analyzing provenance data.

See [3] for a more detailed report on the provenance kernels and their performance.

What opportunities for training and professional development did the project provide?

On November 20th, 2019, Luc Moreau (LM) and Trung Dong Huynh (TDH) ran the Provenance Tutorial session at the [Open Data Science Conference 2019 in London](#) for about 30 conference participants. The recording is available at <https://www.youtube.com/watch?v=v8nojFOYFEk>. We explained the motivations for provenance, introduced the PROV standard, presented some applications of provenance, and reviewed some tools for provenance. We concluded the tutorial with a practical session on provenance modeling. We acknowledge funding from ONR-Global in our presentation (<https://openprovenance.org/tutorial/odsc2019.html>).

How were the results disseminated to communities of interest?

In addition to our fortnightly teleconferences with the SPAWAR team, we undertook the following engagements:

1. IMPACT Provenance Modelling workshop at KCL (March 14–15, 2018) between KCL and SPAWAR teams, resulting in an initial set of provenance templates for IMPACT. The provenance template approach is introduced to the SPAWAR team.
2. Visit SPAWAR in San Diego (Feb 2019): LM and TDH visited SPAWAR between Feb 19th and 22nd to report our research and to plan the collaboration between the two teams. We also worked with the SPAWAR team on provenance modeling and recording during the visit.
3. MAPLE 4 Trial at QinetiQ Portsmouth Technology Park (May 18th and 22nd, 2019). KCL team was introduced to the MAPLE project and the role of IMPACT in the whole MAPLE exercise. We reviewed additional provenance templates created by the SPAWAR team and helped the SPAWAR team integrate the ProvToolbox library into their provenance service within IMPACT, enabling it to generate provenance directly without the need for external utilities.
4. Global Maritime Forum in Liverpool, April 30th, 2019. LM presented ongoing work in the project and took part in a panel on Big Data Analytics
https://nmio.ise.gov/Portals/16/GMF%20Reports/Docs/GMF_2019_Final.pdf.
5. The SPAWAR team (Crisrael Lucero, Doug Lange) visited KCL in London on May 21st and 22nd, 2019, and then September 31st and Oct 3rd, 2019. We discussed the research agenda and worked on the provenance recording for IMPACT.
6. ONR Science of Autonomy Meeting in Washington, DC, August 5th–8th, 2019: LM presented the PA4C2 project to the Science of Autonomy program.
7. Open Data Science Conference Europe, London, Sept 2020, LM presented “Provenance: a fundamental data governance tool - a case study for data science pipelines and their explanations” <https://odsc.com/speakers/provenance-a-fundamental-data-governance-tool-%E2%8E%AF-a-case-study-for-data-science-pipelines-and-their-explanations/>
8. Plenary presentation “Explanations on the Web: A Provenance-based Approach”, by LM, WEBIST 2020, 05 November, <https://vimeo.com/channels/webist/478043709>
9. “Explanations on the Web: a Provenance-based Approach” by LM at Warwick Cybersecurity Expert Seminar series, <https://warwick.ac.uk/research/priorities/cyber-security/seminarseries/>

10. AI Strategic Challenge Assurance Workshop organized by The Technical Cooperation Program (TTCP) (September 30th and October 1st, 2021). LM and TDH presented “Provenance-based Explanations” to the workshop. We also discussed how provenance could benefit assurance cases in command-and-control applications.
11. Visit to NIWC Pacific (December 6–10, 2021): TDH (and LM remotely) reported the findings and results of the project to NIWC team. We discussed future research collaboration and opportunities. TDH was introduced to NIWC collaborators at the UK Maritime Warfare Centre (MWC), leading to a funding award to apply provenance technology to wargaming exercises with MWC (see below).

Further Project and Funding

Provenance of Robotic Planning: We initiated a new line of work with colleagues from the [Planning Group at KCL](#) in which we record the provenance of plans and plan execution from ROSPlan [11], a planning system for robotic systems in ROS³. Once this is completed, we plan to simulate various planning and plan execution scenarios akin to those in dynamic command and control environments with the aim to apply provenance analytics techniques over their provenance.

Provenance Analytics for War Gaming: The KCL team applied and awarded innovation funding to work with the UK Maritime Warfare Centre, Deep Blue C Technology Ltd⁴, and NIWC Pacific to apply provenance analytics techniques to wargaming exercises in the Serge environment⁵. We will exploit results from this project to demonstrate provenance-based analytics capability over user interactions within war games.

Honors: What honors or awards were received under this project in this reporting period?

In a closely related but distinct project, under the theme of accountability (see Figure 1), from March to June 2019, and subsequently in 2020–2021, we investigated and developed a demonstrator⁶ in which textual explanations are generated from the provenance of automated decisions to answer various questions with respect to their data protection concerns. This work was carried out in collaboration with legal experts at the University of Southampton and with inputs from the UK Information Commissioner’s Office (ICO), the data protection regulator in the UK.

The ICO published its guidance on *Explaining decisions made with Artificial Intelligence* in May 2020. The guidance refers to our above work and introduces our provenance-based explanation approach as an

³ The Robotic Operating System: <https://www.ros.org>.

⁴ <http://deepbluec.com>.

⁵ <https://sites.google.com/deepbluec.com/serge/>.

⁶ The demonstrator is available online at <https://explain.openprovenance.org/loan/>.

example technique⁷ “to help provide information for the data explanation ... [and to] provide details for the responsibility, and safety and performance explanations”.

The provenance kernel approach has been shown to contribute to this explainability capability and was mentioned in talks at WEBIST and Warwick.

Technology Transfer

Please list any technology transfer activities including patent applications (include patent number, title, authors, and application date and status), inventions, licenses (include license title, application date and status). Please describe interactions with Navy laboratories or other DoD Agencies. Describe any commercialization efforts.

Provenance Modelling and technical support

We had regular fortnightly teleconferences and face-to-face meetings with the SPAWAR team, both in San Diego and in London (see above). During those engagements, we provided the following:

1. Training on the PROV data model and provenance modeling via workshops and tutorials.
2. Training on the PROV-Template approach [5] and on using the ProvToolbox library to instrument provenance recording for IMPACT.

The SPAWAR team managed to integrate the above provenance technology into their software, leading to the following publications reporting results from their provenance-related work:

1. C. Lucero, P.H. Nguyen, “A Modern Approach to Big Provenance,” *2019 IEEE International Conference on Big Data (Big Data)*, 2019, pp. 6142–6144.
2. C. Lucero, T.D. Huynh, D. Lange, L. Moreau, “Information Provenance for Evaluating Autonomy,” *IJCAI-19 Workshop on Evaluation of Adaptive Systems for Human-Autonomy Teaming (EASyHAT)*, 2019.
3. C. Lucero, T.D. Huynh, D.S. Lange, L. Moreau, “Modeling Provenance of Decisions within the Human Autonomy Team,” *NATO STO-MP-HFM-300 Symposium on Human-Autonomy Teaming*, 2019.
4. S. Wark, M. Nowina-Krowicki, C. Lucero, D. Lange, “But Why? Generating Narratives Using Provenance,” *OzCHI 2018 Workshop on Interaction Design for Explainable AI*, 2018, pp. 5–8.
5. C. Lucero, B. Coronado, E. Gustafson, D.S. Lange, “Tracking provenance in decision making between the human and autonomy,” *International Conference on Engineering Psychology and Cognitive Ergonomics*, 2018, pp. 171–180, doi: 10.1007/978-3-319-91122-9_15.
6. C. Lucero, B. Coronado, O. Hui, D.S. Lange, “Exploring Explainable Artificial Intelligence and Autonomy through Provenance,” *Proceedings of the IJCAI-ECAI 2018 2nd Workshop on Explainable Artificial Intelligence (XAI)*, 2018.
7. M. Nowina-Krowicki, P. Argent, C. Lucero, D. Lange, S. Wark, “Now This Changes Everything: Managing Provenance using Multimedia Narrative,” *ICRA Workshop on Robot Teammates Operating in Dynamic, Unstructured Environments (RT-DUNE)*, 2018.

⁷ The provenance-based explanation approach is described in the guidance's Part 2, Task 2, available at <https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/part-2-explaining-ai-in-practice/task-2-collect/>.

Products

You have the option of selecting “nothing to report” in this section.

There are no limitations to the number of entries you submit and you can also pull information directly using the publication DOI.

Below is the information detailed for each product submission:

1. Journal Article
 - a. Title: **Provenance Graph Kernels**
 - b. Authors: David Kohan Marzagão, Trung Dong Huynh, Ayah Helal, Luc Moreau.
 - c. Publication Status: pre-print.
 - d. Publication Date: 20 Oct 2020
 - e. Publication Identifier Type: ArXiv
 - f. Publication Identifier: <https://arxiv.org/abs/2010.10343>
 - g. Acknowledgement of Federal Support? Yes
2. Conference Paper
 - a. Title: **Incremental Inference of Provenance Types**
 - b. Authors: David Kohan Marzagão, Trung Dong Huynh, and Luc Moreau.
 - c. Conference Name: [International Provenance and Annotation Workshop \(IPAW\)](#)
 - d. Conference Date: July 19th – 22nd, 2021
 - e. Conference Location: Due to Covid-19, it will now be held virtually and online.
 - f. Publication Status: Awaiting publication.
 - g. Publication Date: July 9th, 2021.
 - h. Publication Identifier Type: DOI
 - i. Publication Identifier: [10.1007/978-3-030-80960-7_9](https://doi.org/10.1007/978-3-030-80960-7_9)
 - j. Acknowledgement of Federal Support? Yes
3. Conference Paper
 - a. Title: **The PROV-JSONLD Serialization. A JSON-LD Representation for the PROV Data Model**
 - b. Authors: Luc Moreau and Trung Dong Huynh.
 - c. Conference Name: [International Provenance and Annotation Workshop \(IPAW\)](#)
 - d. Conference Date: July 19th – 22nd, 2021
 - e. Conference Location: Due to Covid-19, it is held virtually and online.
 - f. Publication Status: Published.
 - g. Publication Date: July 9th, 2021.
 - h. Publication Identifier Type: DOI
 - i. Publication Identifier: [10.1007/978-3-030-80960-7_4](https://doi.org/10.1007/978-3-030-80960-7_4)
 - j. Acknowledgement of Federal Support? Yes
4. Conference Paper
 - a. Title: **ProvViz: An Intuitive Prov Editor and Visualiser**

- b. Authors: Ben Werner and Luc Moreau.
 - c. Conference Name: [International Provenance and Annotation Workshop \(IPAW\)](#)
 - d. Conference Date: July 19th – 22nd, 2021
 - e. Conference Location: Due to Covid-19, it is held virtually and online.
 - f. Publication Status: Published.
 - g. Publication Date: July 9th, 2021.
 - h. Publication Identifier Type: DOI
 - i. Publication Identifier: [10.1007/978-3-030-80960-7_18](#)
 - j. Acknowledgement of Federal Support? Yes
5. Conference Paper
- a. Title: **Information Provenance for Evaluating Autonomy**
 - b. Authors: Crisrael Lucero, Trung Dong Huynh, Doug Lange, Luc Moreau.
 - c. Conference Name: [IJCAI-19 Workshop on Evaluation of Adaptive Systems for Human-Autonomy Teaming \(EASyHAT\)](#)
 - d. Publication Status: Published.
 - e. Publication Date: 10 Aug 2019
 - f. URL: [https://kclpure.kcl.ac.uk/portal/en/publications/information-provenance-for-evaluating-autonomy\(8a095afd-a683-4831-826f-1baaa7fe6801\).html](https://kclpure.kcl.ac.uk/portal/en/publications/information-provenance-for-evaluating-autonomy(8a095afd-a683-4831-826f-1baaa7fe6801).html)
6. Conference Paper
- a. Title: **Modeling Provenance of Decisions within the Human Autonomy Team**
 - b. Authors: Crisrael Lucero, Trung Dong Huynh, Doug Lange, Luc Moreau.
 - c. Conference Name: NATO STO-MP-HFM-300 Symposium on Human-Autonomy Teaming
 - d. Publication Status: Published.
 - e. Publication Date: Apr 2019.
 - f. URL: <https://www.sto.nato.int/publications/STO%20Meeting%20Proceedings/STO-MP-HFM-300/MP-HFM-300-11.pdf>.
7. Software Library (Other Product)
- ProvToolbox** is a Java library to create and convert W3C PROV data model representations. It incorporates the PROV-Template compiler as part of its functionality. The library is available online at <https://lucmoreau.github.io/ProvToolbox/>.
8. Code & Data (Other Product)
- The code and data used for the **Evaluation of Provenance Kernel** is published at <https://github.com/trungdong/provenance-kernel-evaluation>.

References

- [1] S. V. N. Vishwanathan, N. N. Schraudolph, R. Kondor, and K. M. Borgwardt, "Graph Kernels," *J. Mach. Learn. Res.*, vol. 11, no. Apr, pp. 1201–1242, 2010, [Online]. Available: <http://www.jmlr.org/papers/v11/vishwanathan10a.html>.
- [2] T. D. Huynh, M. Ebden, J. Fischer, S. Roberts, and L. Moreau, "Provenance Network Analytics," *Data Min. Knowl. Discov.*, vol. 32, no. 3, pp. 708–735, Feb. 2018, doi: 10.1007/s10618-017-0549-3.
- [3] D. K. Marzagão, T. D. Huynh, A. Helal, S. Baccas, and L. Moreau, "Provenance Graph Kernel," Oct. 2020. doi: 10.48550/arxiv.2010.10343.
- [4] D. Kohan Marzagão, T. D. Huynh, and L. Moreau, "Incremental Inference of Provenance Types," in *Lecture Notes in Computer Science*, vol. 12839, B. Glavic, V. Braganholo, and C. Koop, Eds. Springer, Cham, 2021, pp. 145–162.
- [5] L. Moreau, B. V. Batlajery, T. D. Huynh, D. Michaelides, and H. Packer, "A Templating System to Generate Provenance," *IEEE Trans. Softw. Eng.*, vol. 44, no. 2, pp. 103–121, Feb. 2018, doi: 10.1109/TSE.2017.2659745.
- [6] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.
- [7] L. Moreau, "Aggregation by Provenance Types: A Technique for Summarising Provenance Graphs," in *Graphs as Models 2015 (An ETAPS'15 workshop), Electronic Proceedings in Theoretical Computer Science*, Apr. 2015, pp. 129–144, doi: 10.4204/EPTCS.181.9.
- [8] C. Lucero, T. D. Huynh, D. S. Lange, and L. Moreau, "Modeling Provenance of Decisions within the Human-Autonomy Team," in *NATO STO-MP-HFM-300 Symposium on Human-Autonomy Teaming*, 2019, pp. 11-1-11–10, doi: 10.14339/STO-MP-HFM-300-11-PDF.
- [9] A. E. W. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Sci. Data*, vol. 3, no. 1, p. 160035, Dec. 2016, doi: 10.1038/sdata.2016.35.
- [10] S. D. Ramchurn, T. D. Huynh, M. Venanzi, and B. Shi, "Collabmap: Crowdsourcing Maps for Emergency Planning," in *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13*, 2013, pp. 326–335, doi: 10.1145/2464464.2464508.
- [11] M. Cashmore *et al.*, "ROSPlan: Planning in the robot operating system," in *Proceedings International Conference on Automated Planning and Scheduling, ICAPS*, 2015, pp. 333–341, [Online]. Available: <https://www.aaai.org/ocs/index.php/ICAPS/ICAPS15/paper/view/10619>.