Provenance Documentation to Enable Explainable and Trustworthy AI

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
with a
Major in Computer Science
in the
College of Graduate Studies
University of Idaho
by

Amruta Suresh Kale

Approved by:
Major Professor: Xiaogang Ma, Ph.D.
Committee Members: Min Xian, Ph.D.; Tin Nguyen, Ph.D.; Jia Song, Ph.D.
Department Administrator: Terence Soule, Ph.D.

May 2023

Abstract

Although Artificial Intelligence/Machine Learning (AI/ML) systems have outperformed humans in a variety of sectors, the inability to explain their autonomous decisions and actions has created a new challenge in the research community. The need for explainability has shifted the focus of AI research from complex black-box models to explainable and interpretable models. Recently the topic of Explainable AI (XAI) and Trustworthy AI (TAI) has become a hotspot and is widely acknowledged by academia, industry, and government. The basic principle of TAI is to build an AI system that is lawful, ethical, and robust so that humans could rely on them. One of the feasible steps in building a TAI system is through is XAI. XAI aims to make AI/ML results more understandable and explainable to humans. While there are a variety of explainability approaches and methodologies designed for providing explanations and user-friendly decisions, each has its benefits and drawbacks as well as several unsolved challenges. Through this Ph.D. research, our objective is to analyze the inter-relationship between provenance (which is the origin or source of something), XAI, and TAI, build a software package to document provenance and extend reproducibility of AI/ML workflows, and test the package in real-world applications to support XAI and TAI. We want to demonstrate that provenance holds great promise for the new state-of-the-art AI/ML solutions; and adopting provenance documentation is increasingly important for illustrating the details of AI/ML workflows and guiding human decision-making.

In order to achieve our objective, we proposed five research topics and the corresponding activities:

1. Identify the inter-relationship between provenance, XAI and TAI through a literature review.
2. Study different software tools/applications and workflow management system (WfMS) to understand how provenance is documented.

3. Highlight the importance of workflow standardization like Common Workflow Language (CWL) which provides a standardized framework to describe the AI/ML workflows and enable computational reproducibility and portability.
4. Build a software package to document provenance and describe AI/ML workflows into CWL-compliant format to extend reproducibility of workflows.
5. Test the package in real-world applications to support XAI and TAI.

To address the first research topic, we investigated a variety of research papers, techniques, tools, and WfMS that support provenance, XAI, and TAI together. An extensive literature study was carried out with the Scopus database from 2010 - 2020 to discover records indicating provenance, XAI, and TAI, and identify the inter-relationship between them. To address the second research topic, we examined and demonstrated various WfMS, packages, and software applications that capture provenance to make AI/ML models transparent, explainable, and understandable. In order to address the third and fourth research topics, we developed a python package called ***geoweaver_cwl***, which translates Geoweaver AI/ML workflows into the standardized format known as CWL. This package not only ensures that all the essential details of the workflows are documented, but it also enhances the computational reproducibility and portability of workflows. To demonstrate the practical application of ***geoweaver_cwl*** in our fifth research topic, we conducted a series of tests on various use cases ranging from simple to complex, drawn from Geoweaver and other domains.

The need for explainability in AI/ML models has attracted great attention in recent years. However, it is not sufficient to explain AI/ML models using post-hoc explanations alone. Provenance documentation is one of the means to accomplish transparency, traceability, explainability, and reproducibility in AI/ML models. In this research, we provide a community-driven solution ***geoweaver_cwl***, which addresses the current struggles in attaining portability, reproducibility, transparency, and scalability of AI/ML workflows. We evaluated ***geoweaver_cwl*** using various use cases from different domains. The study indicates that the ***geoweaver_ cwl*** package can greatly assist the students, researchers, and geoscience community in translating their AI/ML workflows into CWL-compliant WfMS software applications.

We hope this Ph.D. research not only serves as a starting point for future research advances but also as a reference material that encourages experts and professionals from all disciplines to embrace the benefits of reproducible workflows, provenance, XAI, and TAI.

Acknowledgements

I am deeply indebted to my advisor Dr. Xiaogang (Marshall) Ma for his invaluable advice, continuous support, and patience throughout my Ph.D. journey. Dr. Ma's immense knowledge and abundant experience have always encouraged me in my academic, research and daily life. He gave me every bit of guidance and expertise that I needed during my first few semesters; and has always encouraged me to learn, explore by providing insightful feedback and suggestions. I would also like to take this opportunity to thank Dr. Ziheng Sun for his technical support and mentorship in my study. My appreciation also goes to all the committee members Drs. Min Xian, Tin Nguyen, Jia Song for their support and guidance in making this study a triumph.

It has been an amazing experience working across two research groups from different universities, and I owe a huge debt of gratitude to everyone on the TickBase and Geoweaver team, from whom I have learned so much. I appreciate the assistance and efforts of everyone with whom I have collaborated externally, especially Drs. Frederick Harris, Tin Nguyen, Ziheng Sun, and Chao Fan. My sincere thankfulness also goes to the funding organization that supported my research. This work was made possible through the generous support of the National Aeronautics and Space Administration (NASA) , and the National Science Foundation (NSF) who provided financial support for this project.

I'd like to express my appreciation to my friends in the Computer Science department at the University of Idaho for all the great times that we have shared. I am particularly thankful to Joel Oduro-Afriyie, Sana Algaraibeh, Rayan Alshamrani, Nuzhat Yamin, Drs. Amal Aljohani, Parinaz Ghafari, and Sanaz Salati for their infinite support, motivation, and encouragement throughout this journey.

Dedication

I am deeply thankful to my family and in-laws for their love, support, and sacrifices. Without them, this thesis would never have been written. This final word of gratitude is reserved for my beloved husband Sumit, who has consistently inspired, motivated, and supported me throughout my journey and made these years the best years of my life.

I dedicate this thesis in the loving memory of my father Mr. Suresh Kale, whose role in my life was, and will remain immense forever.

This thesis is also dedicated to all the wonderful women in my family who selflessly prioritized the needs of their loved ones and family over their personal aspirations.
(My Grandmothers, My Mother, My Aunts, and My Mother-Law)

Table of Contents

List of Figures

List of Tables

Statement of Contribution

The research presented in this dissertation proposes a novel approach to AI/ML workflow translation from Geoweaver to Common Workflow Language (CWL) using the Python package ***geoweaver_cwl***. By leveraging the power of CWL, this work significantly improves the interoperability of geospatial data analysis workflows across different platforms and domains. The proposed solution provides a simple and automated way to migrate complex geospatial workflows to a standardized and widely adopted format, facilitating collaboration and reproducibility. This research contributes to the advancement of the field of geospatial data analysis by providing a practical solution to a long-standing problem and offers a valuable tool for researchers and practitioners in this field.

Here is the list of primary responsibilities and contributions of all the co-authors included in this dissertation. Some chapters in the dissertation are based on multi-authored articles. For all those articles, I have taken primary responsibility and authorship. The details are listed below.

Chapter 1 and Chapter 2: Myself and my advisor (Dr. Xiaogang Ma) proposed the topic for the literature review and designed the framework. I conducted the literature review and wrote the first draft. All co-authors (Dr. Tin Nguyen, Dr. Fredrick Harris, Chenhao Li, and Jiyin Zhang) contributed to the discussion and revision of the manuscript.

Chapter 3: In this chapter of the dissertation, I studied and demonstrated different workflow management systems, tools, and software applications and wrote the first draft. Dr. Xiaogang Ma contributed to the discussion and revision of the book chapter.

Chapter 4: This chapter demonstrates the technical framework of our python package ***geoweaver_cwl*** where I proposed the concept, methodology, build a software package, and wrote the original draft, and contributed to the review and editing of this paper. Dr. Ziheng Sun, Dr. Chao Fan, and Dr. Xiaogang Ma contributed to the methodology, resources, review, editing, and validation of this chapter.

Chapter 5: Chapter 5 demonstrates the usability of the python package ***geoweaver_cwl***. In this chapter, we tested ***geoweaver_cwl*** with multidisciplinary use cases from simple to complex once's. I wrote the first draft and tested all the use cases in geoweaver_cwl. Dr. Xiaogang Ma contributed to the discussion and revision of this chapter. Rayan Alshamrani and Zhe Wang provided us with the use case. We also collaborated with geoweaver to get a complex use case.

List of Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
XAI	Explainable Artificial Intelligence
TAI	Trustworthy Artificial Intelligence
DL	Deep Learning
ANN	Artificial Neural Network
DNN	Deep Neural Network
SVM	Support Vector Machine
FATE	Fairness, Accountability, Transparency and Explainability
CWL	Common Workflow Language
DARPA	Defense Advanced Research Project
GDPR	General Data Protection Regulation
HLEG	High-Level Expert Group on AI
FAT/ML	Fairness, Accountability and Transparency in Machine Learning
CNN	Convolution Neural Network
RNN	Recurrent Neural Network
RF	Random Forest
LIME	Local Interpretable Model Agnostic ExPLanations
SHAP	SHapely Additive ExPLanations
TSHAP	Tree SHapely Additive ExPLanations
IG	Integrated Gradients
W3C	World Wide Web Consortium
ACM	Association for Computing Machinery

OWL	Web Ontology Language
MetaCLip	METAdata for CLImate Products
EU	European Union
GIS	Geographic Information Systems
ESIP	Earth Science Information Partners
NASA	National Aeronautics and Space Administration
NCA	National Climate Assessment
PCA	Principal Component Analysis
SSH	Secure Shell
YAML	Yet Another Markup Language
WfMS	Workflow Management Systems
EOSDIS	Earth Observing System Data and Information System
DAG	Directed Acyclic Graph
WfMC	Workflow Management Coalition
OASIS	Organization for the Advancement of Structured Information Standard
BPEL	Business Process Execution Language
BPMN	Business Process Model and Notation
SCUFL	Simple Conceptual Unified Flow Language
YAWL	Yet Another Workflow Language
XML	Extensible Markup Language
CMAQ	Community Multi-scale Air Quality Model
MS	Multiple Sclerosis

Chapter 1 Introduction

This chapter is adapted from the published paper:

Amruta Kale, Tin Nguyen, Frederick C. Harris, Chenhao Li, Jiyin Zhang, Xiaogang Ma; Provenance documentation to enable explainable and trustworthy AI: A literature review. Data Intelligence 2022; DOI: https://doi.org/10.1162/dint_a_00119

1.1 Background of XAI and TAI

Over the past decade, the rapid rise of applications in Artificial Intelligence (AI) has raised the discussion of explainable AI (XAI) and trustworthy AI (TAI) among data science practitioners (Wing, 2020). We have seen remarkable progress in AI algorithms and facilities for high-performance computation, and applications of AI are thriving in various domains, such as virtual assistants, healthcare, autonomous vehicles, criminal justice, human resource, and environmental science. In many applications, the results generated by AI/ML models have a huge impact on human decision-making. However, existing models are insufficient to certify how and why the results were obtained, which leads to growing concerns that these AI/ML models are unfair, opaque, or non-intuitive (Goodman and Flaxman, 2017**)**. For example, ML and Deep Learning (DL) are the most representative technologies in AI and are widely used by data science practitioners. ML is a powerful tool and can identify patterns and examine correlations on large datasets. DL is a subset of ML that achieves great power and flexibility (Goodfellow et al., 2016). It uses a vast amount of labeled data and multiple layers of algorithms to imitate the neural network in our brain, with the aim to achieve human-like cognitive abilities. The most representative DL technology is the Artificial Neural Network (ANN) or Deep Neural Network (DNN). DNN comprises a large number of neurons or nodes with each layer. These nodes are interconnected in a complex manner and activate multiple combinations at each layer. However, it is debatable how this complex network works and derives its output which leads to the "black-box" problem (Castelvecchi, 2016). Although these models perform complex computational tasks with high predictive accuracy, we need to ensure that the steps, workflows, and results of these models are transparent, interpretable, unbiased, and trustworthy. One of the approaches for increasing transparency is to explain these complex models through XAI, which in turn is a feasible step in building TAI (Adadi and Berrada, 2018).

The goal of XAI is to provide algorithmic transparency that can be understood by the average human being (Ribeiro et al., 2016). XAI will help to answer questions like how the system made certain predictions, why the system fails, or what biases are present in the system or data (Guidotti et al., 2018; Murdoch et al., 2019). However, not all AI applications need explanation. Some practitioners and academics discussed that explaining a black-box model is difficult to achieve or perhaps unnecessary. Instead, they suggested that these models should be designed inherently interpretable (Rudin, 2018, 2019; Rudin and Radin, 2019). This approach is highly debatable, as in most of the applications the accurate predictive solutions are provided by complex ML models. Some of the ML models such as rule-based learning, K-nearest neighbor, and linear regression have high interpretability and their workflows are easy to understand. However, many other AI models such as DNN, support vector machine (SVM), and Bayesian models have complex structures and workflows, which are mysterious to the outside observers. On some occasions, even the programmers of these models are incapable of explaining why a model behaves in a certain way and generates a specific output. With the growing use of AI applications in every aspect of our modern life, there is also an increased risk of unanticipated behavior. The danger is in creating and using decisions that are not justifiable, legitimate, or that merely do not allow obtaining detailed explanations of their behavior. In that sense, XAI and TAI will be qualified to reveal the strengths, limitations, and/or weaknesses of AI/ML models. They are also an important means to establish user engagement and trust in AI applications.

1.2 A Brief Reflection of Provenance

The technical approaches for XAI and TAI are under quick development, at which some researchers highlighted that provenance is an evolving field to explain AI-based systems (Liu et al., 2017; Jentzsch and Hochgeschwender, 2019; Frost, 2019). Provenance answers the question of who-what-when-where by documenting the process at each step, such as entities, agents, and activities. By portraying transparency, the documented provenance helps trace back the origin of data, demonstrate the steps of data processing, and determine the trustworthiness of results (Jaigirdar et al., 2019; Amalina et al., 2019; Jaigirdar et al., 2020). Given the non-intuitive nature of many AI/ML algorithms, tracking provenance in AI/ML workflows will be helpful since it is an effective technique to highlight significant components in the process and allows scientists to understand how the result was obtained (Samuel et al., 2020). To achieve repeatability and comparability in AI/ML experiments, one must first understand the metadata and most importantly the provenance of the artifacts in the ML process (Kumar et al., 2016). Very recently, Werder and Balasubramaniam (2021) also suggested that data provenance assists and improve fairness, accountability, transparency, and explainability (FATE) in AI/ML algorithms and enables trust. Several other researchers suggested that provenance documentation is an emerging approach toward XAI and TAI (Liu et al., 2017; Jentzsch and Hochgeschwender, 2019; Frost, 2019). Nevertheless, the work in this field is still limited and there is no systematic discussion or road map for those topics in multi-disciplinary data science.

We anticipate that provenance documentation is an important factor in building XAI and TAI as it not only provides metadata of a workflow but also confirms the authenticity and reproducibility of results. This chapter aims to conduct a literature review of existing research on XAI, TAI, and provenance, with a focus on their applications in data science. We started our literature search by scrutinizing academic papers from Scopus as it is one of the largest and most reliable literature databases for scientific research. The search was conducted based on keywords to select papers. We used generic search strings to get more search results like “explainable ai”, “trustworthy ai”, “artificial intelligence”, “explainable artificial intelligence”, “machine learning”, and “provenance”. Our objective was to focus on recent advances. Therefore, we restricted our search from 2010 to 2020. We followed the standard systematic literature review method with backward and forward snowballing strategies (Wohlin, 2014). Snowballing strategy uses a reference list of the paper or citations of the paper to identify additional papers. The gathered papers were then scanned based on the title, abstract, and keywords to verify whether the reported work includes work on XAI, TAI, and provenance. We did not aim to survey all research papers. Instead, we divided our search based on two standards; 1) selection based on a higher level of citation and 2) high-quality papers including good coverage and technicality in the field. Irrelevant articles were excluded, and the remaining articles were examined in detail to understand whether they provide enough information about the proposed methodology, technical approaches, and results. In addition to the literature found on Scopus, in the review and discussion, we also incorporated a number of other publications that deliver a good definition of fundamental concepts and illustrate successful applications.

1.3 Motivation of this PhD Research

At present, AI/ML research in earth science is lacking in efficient management, and it is difficult to share, replicate, track provenance, and scale ML workflows. Most of the time, scientists manage their ML workflows on their own. Due to the uncertainty, complexity, and variety of ML models, researchers struggle with a solo management strategy to track and control ML workflows, especially when big Earth data is involved. To make AI-based earth scientific workflows more shareable, replicable, reusable, and most importantly provenance-enabled, we collaborated with George Mason University for scaling up the power of cloud-based workflow platforms, automating provenance documentation for open science, and advancing the development to support explainable AI and trustworthy AI.

1.4 Objective of this PhD Research

To address the above-mentioned challenges, the objective this Ph.D. research is to analyze the inter-relationship between provenance, XAI, and TAI, build a software package to document provenance and extend reproducibility of AI/ML workflows, and test the package in real-world applications to support XAI and TAI. In order to ensure that any researcher or practitioner who wishes to upgrade their existing work can reproduce, share, and maintain it without any obstacles while maintaining the integrity of provenance. The dissertation answers the following research questions.

1. How can provenance contribute to the explainability and transparency of AI/ML models to support the goals of XAI and TAI?
2. Would adopting domain-specific provenance standards be necessary, or can we rely on universal standards such as PROV-DM (Provenance Data Model) to document all the necessary complex details?
3. What software tools and WfMS are available for documenting provenance?
4. What sets Geoweaver apart from other WfMS?
5. What are the long-term benefits of standardizing workflows using CWL?

In achieving this goal, the following hypotheses are taken into consideration to help with this study's process of investigation in order to address the aforementioned research questions.

1. Provenance documentation is one of the approaches in providing explainability and transparency in AI/ML models with the help of proper documentation and metadata information.
2. Provenance will not only help users to trace, evaluate, understand, and reproduce the AI/ML results but will also enhance users' decisions about how much trust to place in data and results generated from the original sources.

3. Adding domain-specific documentation standards can help the community to grasp and begin employing appropriate practices routinely.
4. Documenting the necessary detail of a workflow will help researchers with troubleshooting in the event of errors, shedding light on the behavior of the model.
5. Automating the process of provenance tracking by using workflow platforms, tools, and packages will benefit to avoid the risk of manual documentation.
6. Adopting open standard like CWL to describe large scale workflows will enable reusability and collaboration.

1.5 Dissertation Outline

The dissertation consists of six chapters. The four of which (Chapter 2 – Chapter 5) provide solid answers to the aforementioned questions. These chapters are either been published or will be submitted for publication as a peer-reviewed papers. The highlights of these chapters are briefly described below:

Chapter 2: Background and related work

This chapter describes the fundamental concepts of provenance, XAI, and TAI and demonstrates the inter-relationship between the concepts through the bibliometric analysis highlighting the recent development in the research area. This chapter also summarizes the fundamental challenges that specify research questions in the research objective and outlines the structure of the dissertation.

Chapter 3: Provenance in Earth AI

This chapter highlights the importance of documenting provenance in the earth science domain. While demonstrating different tools and WfMS that support reproducible results and provenance tracking like MetaClip, Kepler, and Geoweaver to illustrate the state-of-the-art technologies for ensuring data quality and the workflow process in earth and environmental sciences.

Chapter 4: Geoweaver_cwl: Transforming Geoweaver AI workflows to Common Workflow Language to extend interoperability

This chapter describes and demonstrates the technical framework of our python package ***geoweaver_cwl.*** The proposed work in this chapter features standardizing AI/ML workflows into CWL format in order to enable reproducibility. To verify the usability of this package, it has been tested on the complex workflow provided by Geoweaver. The ***geoweaver_cwl*** Python package is made open access at: https://pypi.org/project/geoweaver-cwl/0.0.1/.

Chapter 5: Utility of the Python package Geoweaver_cwl for improving workflow reusability: An illustration with multidisciplinary use cases

This the chapter briefly demonstrates and tests the usability of ***geoweaver_cwl*** with five different use cases from diverse domains. These use cases were created in Geoweaver and later tested on the package in order to translate them into CWL scripts. The translation was successful, easy, and fast. The exemplar code of the demonstrated use cases is accessible at GitHub https://github.com/amrutakale08/geoweaver_cwl-usecases.

Chapter 6: Conclusion, limitations, and future directions

This chapter briefly reviews all the preceding chapters, demonstrates the overall significance and unique findings, underlines the limitations, and provides recommendations for future studies.

Chapter 2 Background and Related Work

This chapter is adapted from the published paper:

Amruta Kale, Tin Nguyen, Frederick C. Harris, Chenhao Li, Jiyin Zhang, Xiaogang Ma; Provenance documentation to enable explainable and trustworthy AI: A literature review. Data Intelligence 2022; DOI: https://doi.org/10.1162/dint_a_00119

2.1 Fundamental Concept of XAI and TAI

2.1.1 Background of Explainability and Trustworthiness in AI

AI/ML models have achieved rapid progress and worldwide adoption, and many of them can be seen on our streets and at our homes. However, despite the successful AI applications, we still lack a scientific understanding of their workflows. To gain more benefit out of these AI-based systems they first need to explain to humans why they made a certain decision and which important features they considered in the process (Montavon et al., 2017; Adadi and Berrada, 2018; Miller, 2019). There are numerous reasons why these systems should be understandable, interpretable, and explainable. It will not only gain trust in humans but will also give confidence that the system works well. In recent years there have been several controversies where the outcomes generated by AI/ML models were biased or discriminatory (Osoba and Welser IV, 2017; Chen et al., 2019). These models have become so dominant that they are raising doubts about future humanity and demand an explanation. For example, in 2016 Microsoft launched a Twitter bot called "Tay", which was designed to entertain and engage people. In less than 24 hours, Tay's talk extended to racist and offensive comments, forcing Microsoft to take it offline (Tennery and Cherelus, 2016; Vincent, 2016). There were even life-threatening incidents caused by AI. In 2015, a self-driving Tesla was involved in a deadly accident in China when it was in autopilot mode and failed to identify a road-sweeping truck (Boudette, 2016). In another incident reported in 2018, a self-driving Uber killed a woman in Arizona. It turned out that the automatic car's software had no capability to classify an object as a pedestrian until that object was near a crosswalk (Mcfarland, 2018; McCausland, 2019). The IBM Watson system once failed to recommend correct treatments for cancer patients (Ross and Swetlitz, 2018). Also, Amazon's AI recruiting tool displayed a gender bias. It was demonstrated that the new recruiting tool was trained to screen applicants by looking for patterns in applications submitted to the company. The majority of the submissions were from men candidates, reflecting male dominance in the tech industry. Accordingly, the AI recruiting tool trained itself that male candidates were preferable, which eventually led to the gender inequality in its recommendations (Dastin, 2018). There are several more examples mentioned in the literature where AI-based systems malfunctioned (e.g., Tan et al., 2017; Adadi and Berrada, 2018). Accordingly, there is a growing need for tools to check vulnerabilities and flaws in AI-based systems, as well as to help developers and users understand why the machine makes a certain decision.

The basic principle of TAI is to build AI-based systems that are lawful, ethical, and robust to ensure that humans can rely on them (Floridi, 2019; Thiebes et al., 2020; Jain et al., 2020). The key to establish TAI is by using XAI, which refers to the series of frameworks and techniques used to ensure that the results generated by AI-based systems are easily understandable and interpretable to humans (Gunning and Aha, 2019). Explainability plays a crucial role in achieving trust and transparency in AI algorithms. To improve explainability, data science practitioners have developed many approaches and strategic plans on XAI. For example, the National Academies of Sciences and the Royal Society organized a forum in 2017, which reported that trust, transparency, interpretability, and fairness are the most significant societal challenges in AI-based systems (NAS, 2018). Simultaneously, the Defense Advanced Research Projects Agency (DARPA) funded the "Explainable AI (XAI) Program" to improve the explainability of AI results (Gunning, 2019). Also, in July 2017, "The New Generation Artificial Intelligence Development Plan" was sanctioned by China's State Council, to encourage explainability and extensibility (Roberts et al., 2021). In May 2018, the European Parliament set the law of General Data Protection Regulation (GDPR) to award citizens a "Right to Explanation" in cases where their activities are affected by AI (Goddard, 2017). Soon after that, in June 2018 a High-Level Expert Group (HLEG) on AI was set up in the European Commission to design the guideline for TAI (AI HLEG, 2019). The government of Finland published a final report on Finland's artificial intelligence programs in June 2019 in order to position Finland as a leader in the application of AI (MEAEF, 2019). To encourage public trust and promoting the use of AI in the federal government, the White House signed an executive order on TAI in December 2020 (White House, 2020). Along with those efforts, the topics of XAI and TAI have received great attention in the academic, industrial, and governmental sectors.

Very recently, Wing (2021) outlined research agendas that combine the concepts of trustworthy computing, AI, and formal methods for ensuring trustworthiness. In her view, the previous discussion on trustworthy computing covers a set of topics: reliability, safety, security, privacy, availability, and usability. The AI/ML systems especially DL models add a dimension of complexity to traditional computing systems and raise more topics of interest, such as accuracy, robustness, fairness, accountability, transparency, interpretability/explainability, ethics, and more. She also pointed out that although the ML community takes accuracy as a gold standard, XAI and TAI will require trade-offs among the topics mentioned above. In recent years, XAI and TAI topics have also been increasingly discussed in workshops and conferences. For instance, the Fairness, Accountability, and Transparency in Machine Learning (FAT/ ML) conference series are a unique venue for those topics (Rakova et al., 2020). The records of search queries and publications also reflect the increasing attention to XAI and TAI. The graph in Figure 2.1 shows the popularity of keywords on Google Trends from 01/2017 to 12/2020. For the same period, we found 772 publications on Scopus whose title, abstract, or keywords refer to XAI or TAI. Figure 2.2 shows the distribution of those publications in each year.

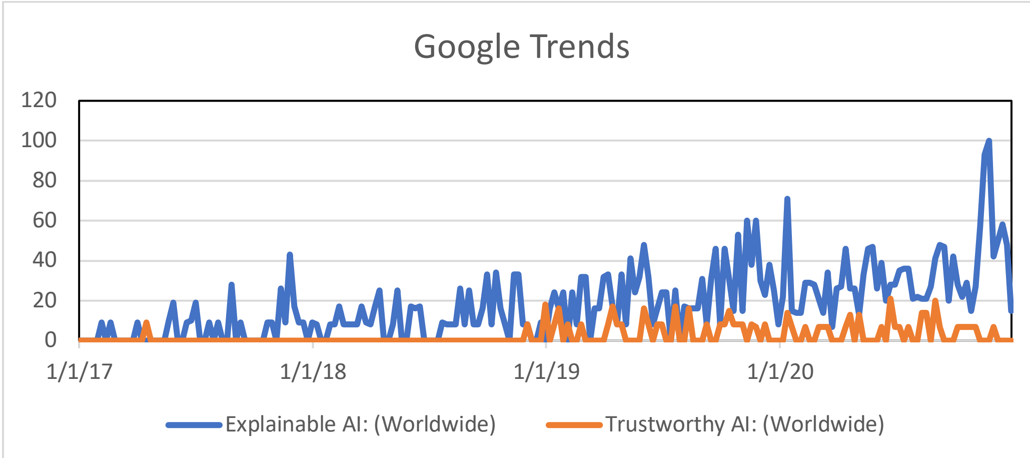

Figure 2.1: Interest over time for the terms "Explainable AI" and "Trustworthy AI"

(Distribution of publications (01/01/2017 – 12/31/2020) whose title, abstract, or keywords include "Explainable AI" or "Trustworthy AI". This query was used to extract the results from Scopus: (TITLE-ABS-KEY ("Explainable AI") OR TITLE-ABS-KEY ("Trustworthy AI")) AND PUBYEAR > 2016 AND PUBYEAR < 2021. The query was conducted on Aug 1st, 2021.)

	Year of Publication				
Document Type	2017	2018	2019	2020	Grand Total
Article			21	53	74
Article in Press				6	6
Book Chapter				5	5
Conference Paper			57	158	215
Conference Review			1	11	12
Editorial			1		1
Final	5	80	127	234	446
Letter				1	1
Note			1	1	2
Review			3	7	10
Grand Total	5	80	211	476	772

Number of records: 1 – 234

Figure 2.2: Distribution of publications

2.1.2 Technical approaches for XAI and TAI

There have been several advances in explanation methods and strategies to make AI-based systems more ethical, transparent, and explainable (Singh et al., 2018). In particular, there have been many discussions on technical approaches to enable XAI and TAI in ML models. ML models are classified into two types: transparent and opaque (Belle and Papantonis, 2020). The transparent ML models are recognized as understandable and capable of explaining to some degree by themselves, such as logistics/linear regression, decision tree, k-nearest neighbors, and Bayesian models (Holzinger et al., 2017; Murdoch et al., 2019). These models can fit well when the primary dataset is not complex. In contrast, opaque ML models are “black-box” in nature, making them complex and tricky to understand. Despite obtaining high predictive accuracy, they lack explainability or interpretability of how the results are generated (Montavon et al., 2017; Adadi and Berrada, 2018). Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Support Vector Machine (SVM) and Random Forest (RF) are the algorithms that fall under opaque models. For instance, RF was initially introduced as a technique to improve accuracy using a single decision tree. In that situation, RF can be treated as a ‘transparent’ model. However, this technique often suffers from overfitting and poor generalization. To address this issue RF combines multiple trees in which each individual tree is trained on a different part of the training dataset and captures different characteristics to calculate the final outcome. This whole process is far more challenging to explain and lacks interpretability than a single tree, forcing the user to apply a post-hoc explainability approaches to gain more insights from it (Belle and Papantonis, 2020, Arrieta et al., 2020). A post-hoc explainability approach is often employed to extract information about what the model has learned (Guidotti et al., 2018). It means that, when an ML model is unable to explain the intricate method, a separate model is applied to provide an explanation. The post-hoc explainability is categorized into two different techniques: model-agnostic and model-specific (Miller, 2019). The model-agnostic technique can be applied to any type of ML model no matter how complex they are. For instance, some model-agnostic techniques such as Local Interpretable Model-Agnostic Explanations (LIME) (Ribeiro et al., 2016) and SHapely Additive exPLanations (SHAP) (Lundberg and Leen, 2017) are widely used to explain DL models. While model-specific technique is only applicable to a single model or a class of models, Tree SHAP (TSHAP) (Lundberg et al., 2020) and Integrated Gradients (IG) (Sundararajan et al., 2017) are some of the popular techniques used for explaining the ML models. When compared with the model-specific techniques, the model-agnostic techniques are more flexible (Ribeiro et al., 2016). Figure 2.3 and Table 2.1 depicts the classification of ML models and the corresponding XAI approaches, in which we have taken the motivation from Arrieta et al. (2020) and Belle and Papantonis (2020), but we adapted the organizational structure to better match the topics discussed here.

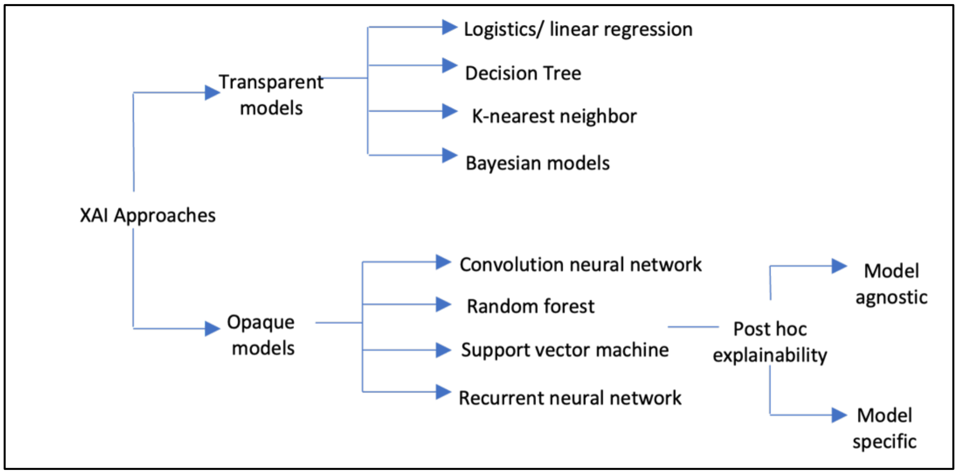

Figure 2.3: Classification of ML models and XAI approaches

Table 2.1: Classification on ML models and explainability method

	Types of ML models	***Classification of model***	***Post-hoc explainability approach required***	***Explainability method***
XAI Approaches	Linear /Logistics regression	Transparent model	No	-
	Decision Tree			
	K-nearest neighbor			
	Bayesian models			
	Convolution neural network	Opaque model	Yes	Visual explanation
	Random Forest			Local/ global explanation
	Support vector machine			Explanation by example
	Recurrent neural network			Explanation by simplification

Although these XAI approaches can generate results to explain an ML model, many metadata and context information are still missing. To increase transparency and explainability in AI-based systems, applying provenance documentation can be a complementary technology to the existing XAI approaches (cf. Singh et al., 2018; Jentzsch and Hochgeschwender, 2019). Provenance documentation shows promise in increasing transparency as it can be used for many purposes, such as understanding how data were collected, determining ownership and rights, tracing steps in data analysis, and making judgments about resources to use. Section 3 presents a detailed bibliometric analysis to demonstrate how provenance, XAI, and TAI are interconnected to each other.

2.2 Provenance, XAI, and TAI: Bibliometric Analysis from Different Aspects

Bibliometric analysis is an effective way to measure the influence of publications in a research area. Our objective behind the bibliometric analysis is to demonstrate evidence of how provenance, XAI, and TAI are interconnected to each other in the publications. To collect the appropriate literature, we compared several databases, such as Google Scholar, PubMed, Web of Science, and Scopus. Although Google Scholar can provide diversified literature, it lacks quality control which makes it inefficient for publication search and analysis. In our work, we decided to focus on only the Scopus database as it provides wide coverage of literature from all major disciplines and all records are organized with good quality measures. A number of terms were used to query the title, abstract, or keywords of publications. As the query script (see below) shows, besides "provenance", we required at least one of the other search terms to be present in the title, abstract, or keywords of a publication. The query was executed in Scopus on August 30th, 2021, and a total of 426 publications between 01/2010 and 12/2020 were found.

Query:

(TITLE-ABS-KEY (machine AND learning)

OR TITLE-ABS-KEY (explainable AND ai)

OR TITLE-ABS-KEY (trustworthy AND ai)

OR TITLE-ABS-KEY (artificial AND intelligence)

OR TITLE-ABS-KEY (explainable AND artificial AND intelligence)

AND TITLE-ABS-KEY (provenance))

AND PUBYEAR > 2009 AND PUBYEAR < 2021

To analyze the results, we used two tools: Bibliometrix and VOS Viewer. Bibliometrix is an open-source tool designed in the R environment for quantitative research, including all the key bibliometric methods of analysis. It allows importing bibliographic data directly from Scopus and other databases. Besides the general bibliometric analysis functions, other measures such as co-citation, coupling, and co-word analysis are also enabled (Aria and Cuccurullo, 2017). VOS Viewer is a software tool for constructing and visualizing bibliometric networks such as authors, journals, and/or individual publications. More sophisticated conditions such as co-occurrences of words or co-citation based on authors can also be used in the network construction (Eck et al., 2010). Below is a list of results generated in our analysis to the 426 publications found on Scopus.

Analysis by timeline of publications: The line graph in Figure 2.4 shows the number of publications per year from 2010 to 2020. The interesting pattern is an exponential growth in publications from 2016. It shows that the studies related to XAI, TAI, and provenance have received increasing attention in the past four years. Figure 2.5 is a word growth graph, which shows the cumulative appearance of authors' keywords (i.e., keywords given by authors in a publication) over time among the 426 publications. While overall it shows a trend similar to Figure 2.4, it is noteworthy that artificial intelligence, machine learning, learning systems, and provenance are the words that stand out as the most predominant among all the authors' keywords.

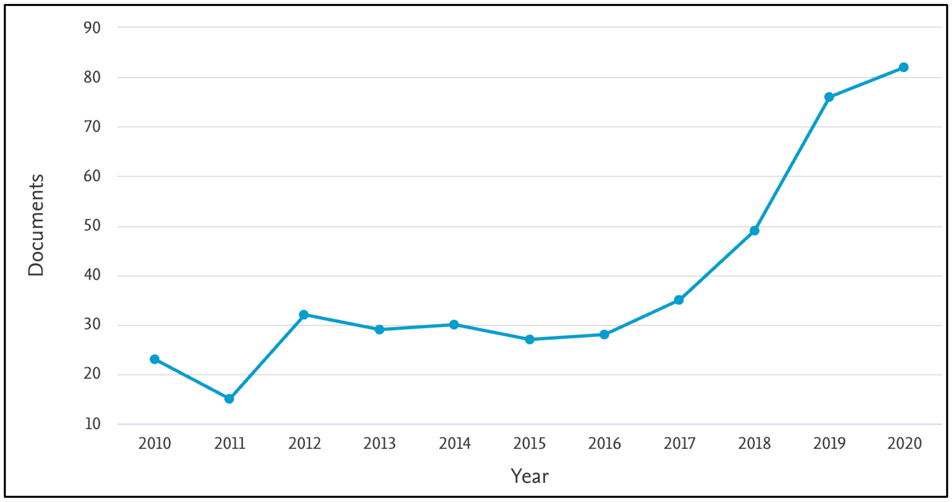

Figure 2.4: Annual number of publications among the 426 records retrieved from Scopus

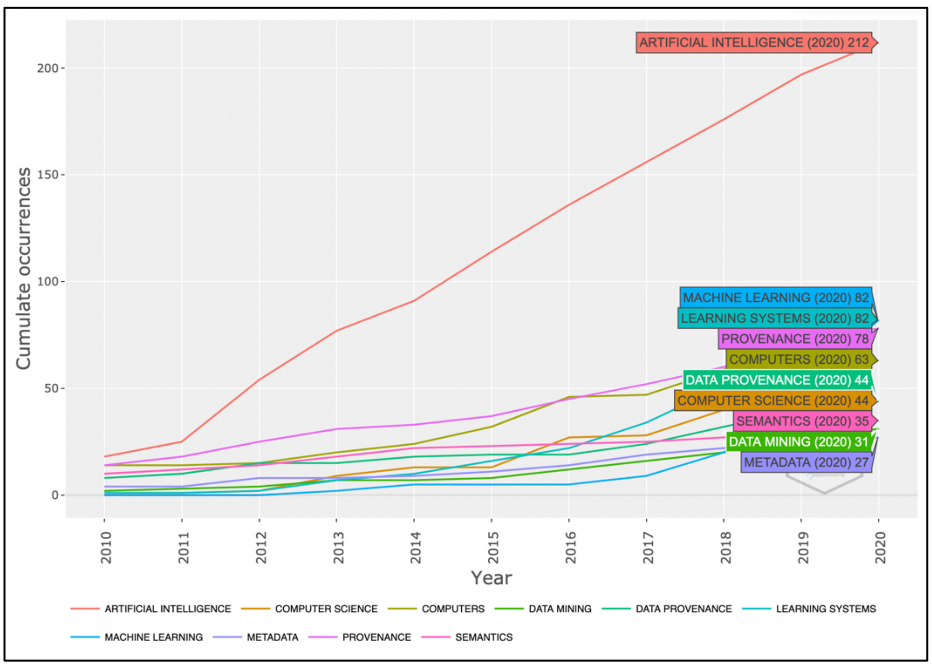

Figure 2.5: Line graph representing cumulative appearance of word growth among authors' keywords of the 426 publications

Analysis by subject keywords of references: The references cited by a publication are also a good way to reflect the subject of the publication itself. Keyword Plus collects words or phrases in the titles of a publications references, which provides greater depth and variety for bibliometric analysis (Garfeild, 1990). With Keyword Plus data of the 426 publications retrieved from Scopus, we created a word cloud to visualize the frequency of keywords (Figure 2.6). The bigger the word or phrase appears in the word cloud, the more often it appears in the Keyword Plus data. Machine learning, learning systems, provenance, data provenance, semantics, and metadata are the most prominent words standing out in the figure.

Figure 2.6: A word cloud illustrating the most frequent keywords in the Keyword Plus data of the 426 publications

Analysis by subject area and document type: Another advantage of Scopus data is to show the disciplinary background of the publications. The pie chart in Figure 2.7 illustrates the proportions of different disciplines among the 426 publications. It is clear that most publications are in the fields of computer science and mathematics. Also, it is interesting to see that about a quarter of the publications have a background in other disciplines, such as engineering, decision science, and Earth and planetary sciences, which means XAI, TAI, and provenance have also received attention in those disciplines. The donut chart in Figure 2.8 represents the proportions of document types. Conference papers are more than half and journal articles are about a quarter of the 426 publications.

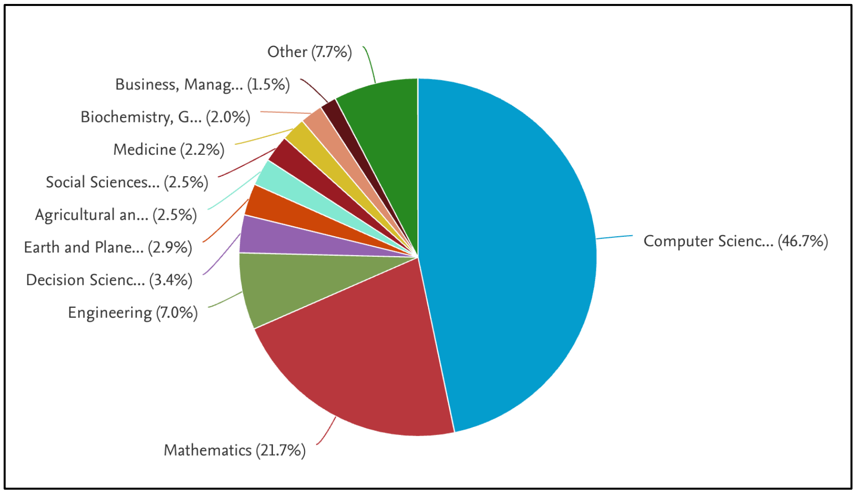

Figure 2.7: Proportions of disciplines among the 426 publications

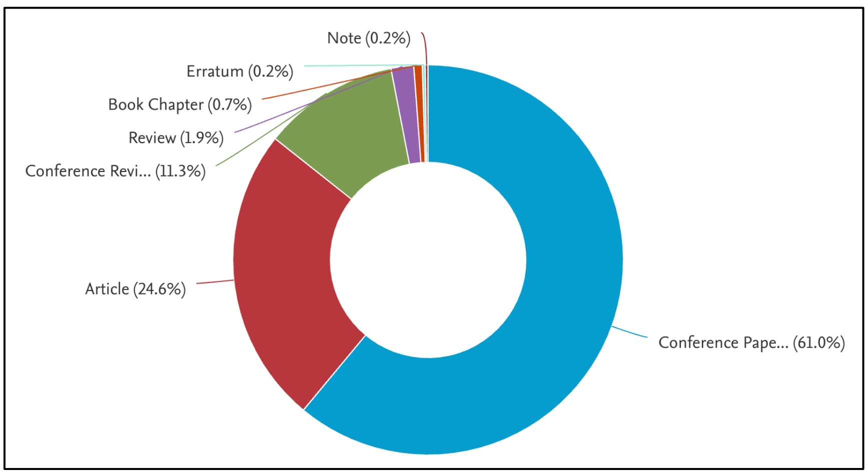

Figure 2.8: Document types among the 426 publications

Analysis by co-relationship of authors' keywords: The co-occurrence of authors' keywords shows how different research topics are relevant to each other in a publication. For all the authors' keywords in the 426 publications from Scopus, we first ranked them by frequency of appearance. Then, we took the top 15 keywords in the list and used VOS Viewer to draw a co-occurrence graph (Figure 2.9). In the figure, the size of each node represents the frequency of appearance of the corresponding keyword. Also, it shows that the 15 keywords are divided into four clusters based on their interconnections, and their frequency of co-occurrence is reflected in the size of lines between the nodes. Among all the 15 keywords and four clusters, provenance and machine learning have the highest appearances. They are closely interconnected with each other and also co-occur with a large number of other keywords.

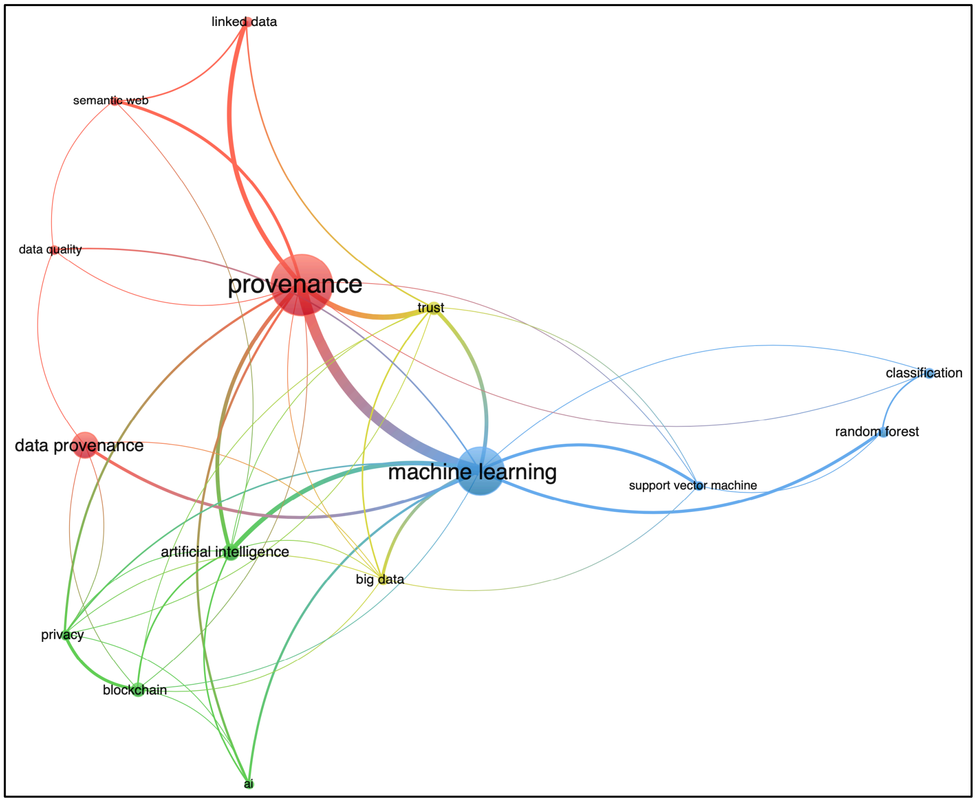

Figure 2.9: Co-occurrence of authors' keywords among the 426 publications. Here only the top 15 keywords with the highest frequency of appearance are shown

2.3 A Reflection on the Relationship Between Provenance, XAI, and TAI

2.3.1 Increasing Attention and Community Works on Standards for Provenance Documentation

The bibliometric analysis in the above section shows an increasing trend of research on provenance, XAI, and TAI. This subsection will incorporate the review of a number of other publications to demonstrate their inter-relationships at a finer scale. Experts and researchers are interested in capturing provenance for several reasons, among which the most important is that well-documented provenance confirms the authenticity of scientific outputs (Moreau et al., 2008). Provenance is the origin or history of something in its literal meaning (Cheney et al., 2009). Some researchers (Jentzsch and Hochgeschwender, 2019) discussed that provenance can be understood as a subset of metadata. We would like to add that provenance not only present the metadata of various objects in a workflow but also the interrelationships between them to show the history of derivation (Ma, 2018). According to PROV Family Documents of the World Wide Web Consortium (W3C), provenance is described as "*information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness*" (Groth and Moreau, 2013; Missier et al., 2013). As such, provenance can answer questions such as how the quality of the data is, what is the data source, when was the data created, what were the steps involved in creating a result, what were the steps in a model used for the data analysis, and who developed and/or ran the workflow (Moreau et al., 2008; Moreau and Groth, 2013).

As AI continues to expand with more diverse information the need of documenting provenance also increases. AI systems need to include provenance as it enables trust and provides users with tools that allow them to access, record, and further investigate resources and steps in a workflow (Chari et al., 2020). The Association for Computing Machinery (ACM) Policy Council set principles for transparency and accountability, in which data provenance is one of the key principles (Garfinkel et al**.**, 2017). Although their comments are on the generic transparency and accountability, their approaches and methods are also insightful for the work of XAI and TAI. Kirkpatrick (2016) stated that regular supervision is necessary for AI-based systems as they can cause harm to many people by generating bias or discriminatory results. Even if the predictions generated by AI/ML models deliver high accuracy, it is crucial to know the very roots before concluding any decision, especially in critical domains such as human activities (Buneman and Tan, 2019; Shaw et al., 2019). Jentzsch and Hochgeschwender (2019) stated that adopting the established methods from the field of provenance to describe ML models will lead to more transparent AI-based systems. A few other researchers also discussed how provenance can increase the reproducibility of ML models (Miles et al., 2007; Davidson and Freire, 2008; Alahmari et al., 2020). Recently, Sarpatwar et al. (2019) described how blockchain allows users to trace the provenance of training models resulting in more transparent and fair AI-based systems. For example, users will be able to discover biases or unclear sourcing of data and see what exactly leads to an action or decision made by AI-based systems. Several other researchers also proposed that provenance is essential to hold AI-based systems to the same standards of accountability as humans (Goodman and Flaxman 2017; Lucero et al., 2018). Based on a review of those publications, in Figure 2.10 we present the research topics involved in provenance, XAI, and TAI, and illustrate the overlapped parts.

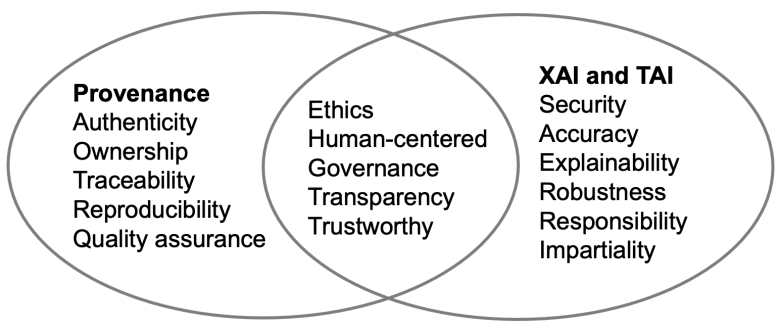

Figure 2.10: The similarity of topics involved in provenance, XAI, and TAI

There are many existing models, languages, and tools designed and developed by researchers to enable provenance documentation, and some are developed specifically for AI/ML models. The W3C PROV Ontology (PROV-O) is a representation of the PROV Data Model (PROV-DM) using the Web Ontology Language 2 (OWL2) (Lebo et al., 2013). It allows creating new classes, properties and exchange provenance information generated from different systems. ProvStore is the first online public provenance repository supporting the standards of W3C PROV. It allows users to store, access, integrate, share, organize, visualize, and export provenance documents in various formats, such as PROVN, JSON, Turtle, and XML (Huynh and Moreau, 2014). There are also tools supporting the validation and browsing of provenance documents. ProvValidator is an online tool for validating provenance documents, ensuring that the documents have consistent history and are safe to use for analysis (Moreau et al., 2014). Prov Viewer is a visualization tool that allows users to explore provenance data through zooming, collapsing, filtering to provide different levels of granularity in the analysis (Kohwalter et al., 2016). For workflow platforms and AI/ML models, there are also ongoing activities on specific standards and tools for provenance documentation. The Common Workflow Language (CWL) is a standard designed to provide specifications and semantics for workflows and tools in data-intensive science. The goal is to make scientific results portable and scalable across software and hardware environments, and thus support reproducibility (Amstutz et al., 2016). OpenML is an online platform that allows machine learning researchers to share the code and results (e.g., model, prediction, and evaluation) and organize it in an effective way for easy access (Vanschoren et al., 2014). ModelDB is an open-source end-to-end system for the management of ML models and has libraries available for Scikit-Learn and Spark ML. It also allows data scientists to perform experiments and build ML models, while the metadata such as pre-processing steps, hyperparameters, quality metrics, and training are automatically captured in the background. ModelDB uses a relational database to store all the extracted metadata and a branching model to track each model's history over time (Vartak et al., 2016).

2.3.2 Real-world practices of provenance documentation and the support to XAI and TAI

In real-world practice, the scope of provenance differs from user to user and is also dependent on the research needs and technologies used (cf**.** Simmhan et al., 2005; Buneman et al., 2008; Cheney et al., 2009). To formalized provenance documentation, Groth et al. (2012) outlined the characteristics of the provenance model into several categories, such as content, management, and use. The purpose is to support engineers to categorize the components and dimensions according to the functionality they are involved in. The W3C PROV is a set of documents that defines various aspects necessary to achieve, exchange, and make use of provenance information amongst diverse environments (Groth and Moreau, 2013)**.** For example, the PROV-DM is structured in six components: 1) entities and activities, and the time at which they were created, used, or ended, 2) derivations of entities from other entities, 3) agents bearing responsibility for entities that were generated and activities that happened, 4) a notion of the bundle as a mechanism to support the provenance of provenance, 5) properties to link entities that refer to the same thing, and 6) collections forming a logical structure for its members (Moreau and Missier, 2013). Those models, categories, and guidelines are further adapted to match needs in real-world applications. For instance, Branco and Moreau (2006) attempted to build a large-scale provenance model for an eScience experiment enabling provenance to be made available as metadata. Pimentel et al. (2016) presented a unique approach for analyzing and tracking provenance collected from scripts. This tool helps scientists record, reproduce, and compare all information and supports decision-making. Huynh et al. (2018) proposed a provenance network analysis method by applying ML techniques on the network metrics to generate provenance information automatically from application data/logs. To provide sufficient information on the decisions made by AI-based systems to the end-users, Jaigirdar (2020) proposed a six-W framework (which, what, who, where, when, and why).

There have been many successful applications of provenance documentation in recent years, and some of them show good performance with AI/ML models in workflow platforms. Renku is an open online platform that can track every version of data, code, and results, and help researchers evaluate, reproduce, and reuse data and algorithms (Krieger et al., 2021). WholeTale is a similar platform that enables reproducibility by allowing researchers to capture and share data, code, and workflow environment in research (Brinckman et al., 2019). Tilmes et al. (2013) and Ma et al. (2014) adapted PROV-O in an ontology to capture provenance of workflows in global change research. Based on those earlier works of provenance documentation, Ma et al. (2017) developed an experiment to capture fine-granular provenance of workflows in Jupyter. Schelter et al. (2017) proposed a lightweight system that allows storage, extraction, and management of provenance and metadata from ML experiments. Dataset, models, predictions, evaluations, hyperparameters of the models, schemas of the dataset, and layout of the deep neural network are some of the common artifacts that can be achieved. Spinner et al. (2019) designed a visual analytics system named "exlpAIner" which allows users to understand all steps of an ML model, diagnose the limitation using XAI methods, and then refine and optimize the model. Agu et al. (2019) developed a guideline provenance ontology (G-Prov), with the intent to represent provenance of treatments at different granularity levels and share the information with healthcare practitioners. Provenance of scientific workflows has been a long-term concern in research **(**Davidson et al., 2008). Recently, with the wide usage of Jupyter and RMarkdown in different scientific disciplines, there has also been solid progress on provenance documentation in workflow platforms. For instance, Samuel (2019) designed a tool named ProvBook, which captures and stores the provenance of a notebook in Jupyter and allows users to compare results. ProvPy is a Python library with an implementation of the W3C PROV-DM. It allows to import and export of provenance information in different formats, such as PROV-JSON and PROV-XML (Huynh, 2020).

Some recent projects also leverage the technical advances in semantics, data visualization and cloud computing. For example, MetaClip (METAdata for CLImate Products) (Bedia et al., 2019) develops vocabularies and an R package to capture the provenance of climate research in PROV-O format. The provenance is recorded in JSON-LD format and appended inside the image file of a climate research output. Then, an interactive web portal can load the image and then read and visualize the provenance information into a graph. The nodes and edges in the graph are interactive, where an end user can click and browse the detailed attributes. Another example is Geoweaver (Sun et al., 2020, 2022), which is an open-source and cloud-based application that allows AI practitioners in earth science to integrate, write, and share workflows. In the cloud-based environment, other users can easily find and trace shared workflows of interest and replicate the code in their own work.

2.4 A Vision on the Trends of Provenance, XAI, And TAI in the Next Decade

It is evident that provenance can help us address issues associated with transparency, explainability, accountability, and authenticity in XAI and TAI. The above bibliometric analysis and reflection highlighted many existing studies, and we believe there will be more advancement in the joint research of XAI, TAI, and provenance in the coming years. Below is a list of our thoughts on future work.

Although AI/ML models have made profound advances, many of them are still deficient in preventing biased and discriminative results. Biases might be caused by many reasons, such as incomplete data, data labelling, adversarial manipulation, missed steps in an ML model, or a workflow guided by a bad hypothesis. Adapting provenance methods will lead to more traceability and transparency of AI applications. A comprehensive description of methods, models, algorithms, and data should be recorded with the aim that they can be further reviewed. Rigorous validation and testing should be done on AI/ML models, and those test results should also be well documented. These steps in provenance documentation can help researchers build explainable and trustworthy systems. Even though documented provenance cannot immediately determine the cause of a bias or error, the complete information can support researchers in tracing all components in the workflow to find the likely cause.

As data are the primary source for any results generated by an AI-based system, studies of XAI and TAI can benefit from many existing mature technologies of metadata and data provenance. Data are suspect when the origin cannot be verified. If a company is using data that are not traceable but concluding an important decision, then this decision is not reliable and will raise concerns amongst users. Provenance provides the flexibility of documenting data at every single step in a data science workflow, ranging from data collection, data cleansing, data analysis, derived data, to the final result. The documented data provenance will be a solid component for XAI and TAI in AI-based systems.

The granularity of provenance (i.e., level of details) depends on the real-world needs. It is crucial to understand that different stakeholders have different requirements on the details of provenance in AI/ML models. Not all people are interested in detailed workflow documentation, while some critical domains such as healthcare, government, and criminal justice require diligent information as the results generated by AI/ML models can have a serious impact on human life, environment, and/or policy making. For AI-based systems, there should be a detailed user survey to clarify the needs of stakeholders before the functions for provenance documentation are developed.

More automated technologies and tools should be developed for recording and sharing provenance information of AI-based systems. We need efficient tools to document provenance and a better-digitized environment to archive, share, and distribute the provenance information to a broad community. Those tools will document the provenance in standard structures and make the information accessible and queryable. In particular, we hope packages can be used for popular workflow platforms such as Jupyter and RMarkdown to automatically document provenance. Several recent studies mentioned in Section 4 have already made solid progress in that direction. Once those packages are in place, there can be a lot of adoptions and adaptations in various scientific domains.

Moreover, we need to understand XAI and TAI as a socio-technical issue, and we need a comprehensive approach to tackle the issue from both social and technical aspects. The GDPR released by the European Parliament is a good example to help understand this topic. GDPR introduces the standardized data protection law, aiming to create consistent protection of users' data. It states that the data cannot be used without user consent. To assist the implementation of this regulation, provenance information can be used to track down all the activities, which can help to clarify if the data are used in the right way or not. In the world of AI, more work is required to increase awareness and fully establish users' rights and obligations on their data.

2.5 Conclusion

The need for explainability in AI/ML models has attracted great attention in recent years. However, it is not sufficient to explain AI/ML models using post-hoc explanations alone. Provenance documentation is one of the means to accomplish transparency, traceability, explainability, and reproducibility in AI-based systems. This study presented a systematic literature review of recent work and advances in the field of XAI, TAI, and provenance. First, we provided the fundamental concepts of XAI and TAI and listed the latest discussions on these topics. Second, we analyzed the inter-relationships between XAI, TAI, and provenance through a bibliometric analysis. We specified how provenance documentation plays a crucial role in building explainability and trustworthiness in AI-based systems, and briefly introduced a few tools and platforms such as Renku, WholeTale, MetaClip, and Geoweaver. Third, we presented a vision on the trends of research on XAI, TAI, and provenance in the next decade. We hope this literature analysis highlights the importance of provenance in AI-based systems and encourages AI practitioners/researchers to start documenting provenance. We expect to see more AI/ML models become explainable, providing enough details to the end-user, and we believe that provenance documentation will be one of the significant approaches to accomplish that.

Chapter 3 Provenance in Earth AI

This chapter is adapted from the published work:

Amruta Kale, Xiaogang Ma; "Provenance in Earth AI", 2022. Sun, Z. (ed.) Earth Science Artificial Intelligence, Elsevier.

3.1 Introduction

The growing use of Artificial Intelligence (AI) and Machine learning (ML) has become an indispensable part of our modern life. This is due to the advancement of technologies (e.g., deep learning (DL)) that have largely contributed to the enormous success of AI/ML systems in terms of prediction and accuracy. Even with such unprecedented success, there are still a few challenges that slow down or hinder the AI/ML systems, such as the inability to explain their decisions. Although the black-box nature of these AI/ML systems allows complex computational tasks with powerful prediction, the processes are opaque and lack explainability. Even though we understand the mathematical architecture of a machine learning system, getting insight into the interior workings of these models is sometimes challenging. Accordingly, there is a strong desire for explicit modeling and reasoning tools to explain how and why a result was attained.

Due to the vast availability of big data (volume, variety, and velocity), DL algorithms are now frequently used in most domains. The huge success of DL models like Deep Neural Networks (DNN) and Artificial Neural Networks (ANN) comprises a combination of multiple layers and millions of parameters that extract important features from raw data. Yet, this complex process also makes DNN applications into black-box models (Castelvecchi, 2016). Even though these models deliver high predictive accuracy they often lack transparency. As the black-box model is increasingly used, the demand for explanation is also increasing from various stakeholders in AI (Preece et al., 2018). Another risk lies in making and implementing judgments that are not reasonable, lawful, or simply do not allow an extensive explanation of their actions especially in critical domains (Gunning and Aha, 2019). It is common to believe that focusing purely on performance will make AI/ML systems opaque, unfair, and non-intuitive. As the demand and awareness for ethical AI are increasing, people are hesitant to apply AI/ML techniques that are not transparent, interpretable, reproducible, and traceable (Goodman and Flaxman, 2017; Zhu et al., 2018). It is a common understanding that there are trade-offs in terms of the model's performance/accuracy and transparency, but as we are moving towards a more automated world, AI/ML models should also be human-understandable. To facilitate human understandability users often require explanations from AI/ML models as to how these systems arrived at their conclusions; this however is often lacking in the existing system (Montavon et al., 2017; Adadi and Berrada, 2018; Miller, 2019).

Recently researchers have acknowledged the increasing need for explainable artificial intelligence (XAI) and trustworthy artificial intelligence (TAI) into AI/ML systems (Wing, 2020). As a result, several survey papers have highlighted the significance of XAI and TAI (Adadi and Berrada, 2018; Arrieta et al., 2020; Belle and Papantonis, 2020: Wing, 2021). This research field holds substantial promises to address the challenges mentioned above. XAI refers to the methods and techniques in the applications such that the results generated by AI/ML models are easily explainable, and understandable to humans (Ribeiro et al., 2016, Gunning and Aha, 2019). It may include general information about how the system operates, why does the system failed, what underlying features were considered, and information about training and test dataset (Guidotti et al., 2018; Lipton, 2018; Murdoch et al., 2019). However, we also believe that explanation is user-focused, and the type of explanation depends on the user's role, former knowledge, and domain. Some safety-critical applications may require comprehensive knowledge to make judgments, while others may not require a detailed description of the systems and how they arrive at their conclusions. For example, a meteorologist anticipating weather forecasts and other weather occurrences, such as where a hurricane would impact, may require thorough information about the factors that influence atmospheric conditions, and weather patterns over time. However, general users may require information regarding the circumstances and safety precautions to be undertaken. To make AI/ML systems transparent, explanation is key as it provides extensive knowledge about the system and builds user engagement in the AI/ML systems.

Despite the revolutionary success of AI/ML models, there have been many post-hoc explainability approaches (Guidotti et al., 2018; Lipton, 2018; Arrieta et al., 2020; Belle and Papantonis 2020) designed to provide explanations to AI/ML models that are not transparent by design. More preciously, these post-hoc methods are an interpretable model (e.g., linear model or decision tree) which is used to train on the black-box model to get a better understanding. These techniques contain explanations about the results in the form of natural language explanations (Krening et al., 2016), visualizations of learned models (Mahendran and Vedaldi, 2015), and explanations by example (Mikolov et al., 2013) to understand the underlying model. However, we believe XAI is a diverse topic, and it cannot be solved by a single disciplinary approach. Consequently, some academics stated that provenance is also an emerging field that can be used to explain AI/ML systems (Liu et al., 2017; Jentzsch and Hochgeschwender, 2019; Frost, 2020). Provenance has the capability of explanation, that has been neglected or has not received the attention it deserves. The inclusion of provenance can address what and why aspects by documenting the entire process. Several researchers further discussed that enabling provenance is essential for determining authenticity, building trust, and ensuring reproducibility in AI/ML models (Jaigirdar et al., 2019; Amalina et al., 2019; Jaigirdar et al., 2020). In one of our previous literature reviews, we found that including provenance in AI/ML models will bring strength in explanation and improve transparency (Kale et al., TBD). We believe that adding provenance in AI/ML systems will help generate resourceful and sufficient explanations for users along with reproducibility.

In this chapter, we will first provide an overview of basic concepts in provenance, XAI, and TAI. Second, we will discuss the related work in the field of provenance and AI in earth science by mentioning the state-of-the-art progress. Third, we will present several tools from the communities designed for capturing provenance to support explainability and transparency. Finally, we will discuss the progress and trends, and wrap up the chapter.

3.2 Overview of Relevant Concepts in Provenance, XAI, and TAI

3.2.1 Guidelines for Building Trustworthy AI

AI has enormous potential in revolutionizing everyone's lives. It has spread across all facets of society bringing profound changes individually, societally, and environmentally. However, even with such unprecedented advancement, they still face challenges in addressing trustworthiness, transparency, and intelligibility. In order to build a transparent and fair AI system, the High-level Expert Group (HLEG) on AI prepared a document on ethics and guidelines on TAI (AI HLEG, 2019). This guideline listed seven fundamental criteria that AI systems must achieve to be considered trustworthy. TAI is built on three key components: an AI system should be (1) lawful, adhering to all applicable laws and regulations; (2) ethical, respecting the principal and ethical values; and (3) robust, both technically and socially (Floridi, 2019; Thiebes et al., 2020; Jain et al., 2020).

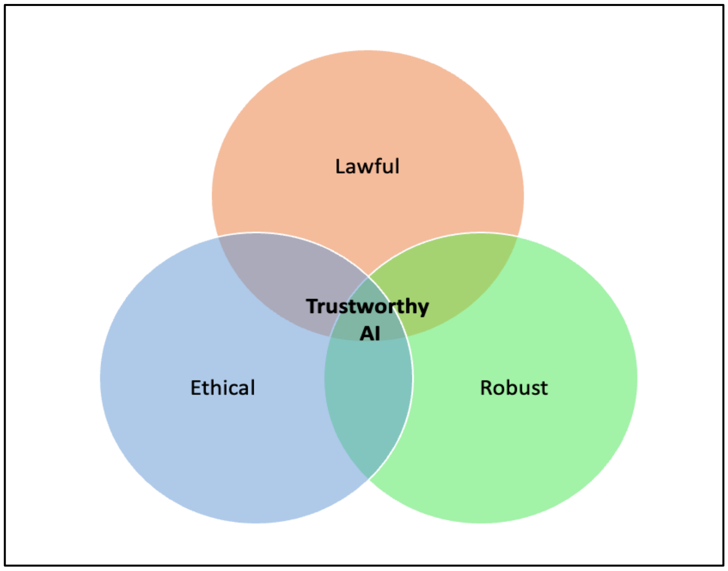

Figure 3.1: Trustworthy AI with three key components

As AI/ML systems are increasingly used, the necessity to explain the results has led to new discussions and actions in scientific communities. It is essential that these systems must be transparent, unbiased, and reliable, which is why these guidelines are so important. These guidelines will help newcomers to attain a basic understanding of what is TAI and how to realize it. Here are the seven European Union (EU) guidelines for defining TAI (AI HLEG, 2019):

- **Human agency and oversight:** AI systems should support human agency and fundamental rights and not limit or mislead human freedom.
- **Technical robustness and safety:** Trustworthy AI demand algorithms to be safe, consistent, and robust enough to deal with errors or irregularities throughout the AI system's life cycle.
- **Privacy and data governance:** Throughout the entire lifecycle, AI systems must maintain privacy and data protection where users should have complete control over their own data.
- **Transparency:** AI systems should be traceable, explainable, and well communicable even if the system has flaws or limitations in it.
- **Diversity, non-discrimination, and fairness:** AI systems should be fair to all stakeholders regardless of their age, gender, abilities, or characteristics.

- **Societal and environmental well-being:** AI systems should promote social transformation as well as enhance environmental sustainability and accountability.
- **Accountability:** Mechanisms should be put in place to ensure ownership, accountability, and potential compensation for AI systems and their outcomes.

3.2.2 Understanding Explainable AI

The inability of AI systems to provide comprehensive information has raised social, ethical, and legal pressure for the development of new AI techniques that are capable of making explainable and understandable decisions. TAI and XAI are often mentioned together. XAI suggests a transition toward more transparent AI. However, XAI is not a new field; the term was first coined by Lent et al., (2004) to highlight the ability of their system to explain the behavior of AI-controlled entities in simulation games. Recently, the topic has received great attention from both academia and industry. As a result, several survey papers have highlighted the noteworthy importance of XAI (Adadi and Berrada, 2018; Singh et al., 2018; Lecue, 2019; Arrieta et al., 2020; Belle and Papantonis, 2020; Das and Rad, 2020). This research field aims to develop a set of strategies that will make the result of AI/ML systems understandable to humans. XAI will be essential if the user needs to understand what, why, and how aspects of the models. To address this, the Defense Advanced Research Projects Agency (DARPA) funded the "Explainable AI (XAI) Program" to improve explainability through the local and post-hoc interpretation methods (Gunning, 2019; Arrieta et al., 2020). The program focuses on building explainable models while maintaining high predictive accuracy (Gunning and Aha, 2019), with the goal of creating new ML techniques that combine explanations, and enable users to understand, manage and effectively gain trust. These ML techniques will have the ability to identify flaws and predict how the machine will behave in the future.

The key objective of XAI is to address trustworthiness and intelligibility in AI/ML models. Figure 3.2 characterizes the visual representation of DARPA's XAI program. However, for greater clarity, we have simplified the diagram based on the types of ML models. Traditionally there are two types of ML models, and the choice of these models depends upon the different application purposes. Transparent ML models (e.g., linear regression, k-nearest neighbors, bayesian models, decision trees) have the ability to figure out what went wrong in the system or explain how they arrived at a particular decision (Holzinger et al., 2017; Murdoch et al., 2019). These models have a substantial percentage of training and test accuracy and work well with simple datasets. However, when dealing with complex applications, transparent models are insufficient for analysis which is why opaque models are required. Opaque ML models (e.g., deep learning and neural networks) are black boxes in nature, despite high predictive accuracy, these models cannot be easily examined or understood (Montavon et al., 2017; Adadi and Berrada, 2018). But with the new XAI approach, the system takes input from the current task and makes decisions, recommendations, and actions that allow users to understand and evaluate based on system explanation. This technique will help users regulate their decisions by providing a reason or justification, particularly when unexpected assumptions are made. XAI will provide insights about the behavior of systems and their unknown flaws, improve models' transparency, and verify predictions, which will lead towards TAI.

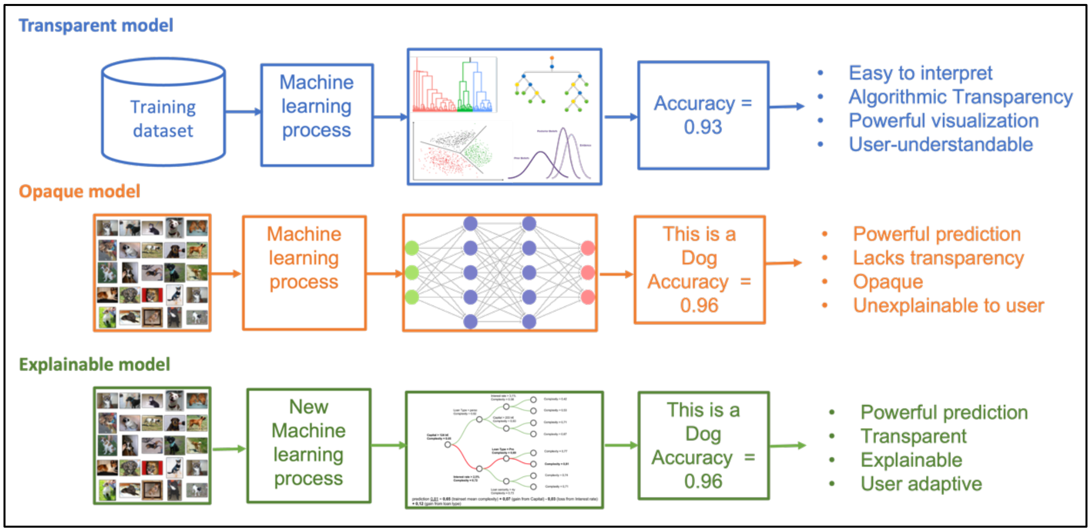

Figure 3.2: Transparent vs Opaque vs Explainable model (Above image is adapted from DARPA's XAI program, (Gunning and Aha, 2019))

3.2.3 Provenance and Provenance Documentation

XAI provides transparency and increases the intelligibility of a system using post-hoc explanation methods (Köhl et al., 2019). In our opinion, gaining transparency using post-hoc methods can be useful, but many metadata and context information about these systems are still widely neglected. To achieve accurate explanation, provenance documentation should be an essential component for the XAI approaches (cf. Singh et al., 2018; Jentzsch and Hochgeschwender, 2019). Experts and researchers are interested in documenting provenance for several reasons. Most importantly, well-documented provenance confirms the credibility of the scientific results and enables reusability (Moreau et al., 2008; Zeng et al., 2019). Provenance determines ownership, as it provides the historical context of who has owned the work and when. The term provenance has an exceptionally long history, it is "the origin or source of something" (Cheney et al., 2009). In this way, it is like metadata, which is data about data (Ma, 2018). Metadata is a crucial component of data collection and distribution. It provides information like the author's details, date created, date modified, and data file versions in a structured and standardized form in such a way that the dataset can be potentially reused.

According to the definition of World Wide Web Consortium (W3C), provenance is defined as "*information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness*" (Groth and Moreau, 2013; Missier et al., 2013). That definition means provenance can be used to document not only metadata but also other entities and steps in a workflow. The W3C PROV family consists of twelve documents (PROV-Overview, PROV-PRIMER, PROV-O, PROV-DM, PROV-N, PROV-CONSTRAINTS, PROV-XML, PROV-AQ, PROV-DICTONARY, PROV-DC, PROV-SEM, and PROV-LINKS) which give details to help understand and implement provenance documentation. Figure 3.3 illustrates the basic elements of PROV-O (Prov Ontology). It is also known as the starting point terms and is built on three fundamental classes (1) *prov:Entity,* which is a physical, digital, conceptual, or other kinds of thing with some fixed aspects; entities may be real or imaginary; (2) *prov:Activity,* which is something that occurs over a period of time and acts upon or with entities; it may include consuming, processing, transforming, modifying, relocating, using, or generating entities; and (3) *prov:Agent*, which is something that bears some form of responsibility for an activity taking place, for the existence of an entity, or another agent's activity. We believe adding provenance in AI/ML systems will help address the issues associated with reproducibility, transparency, explainability, accountability, and authenticity.

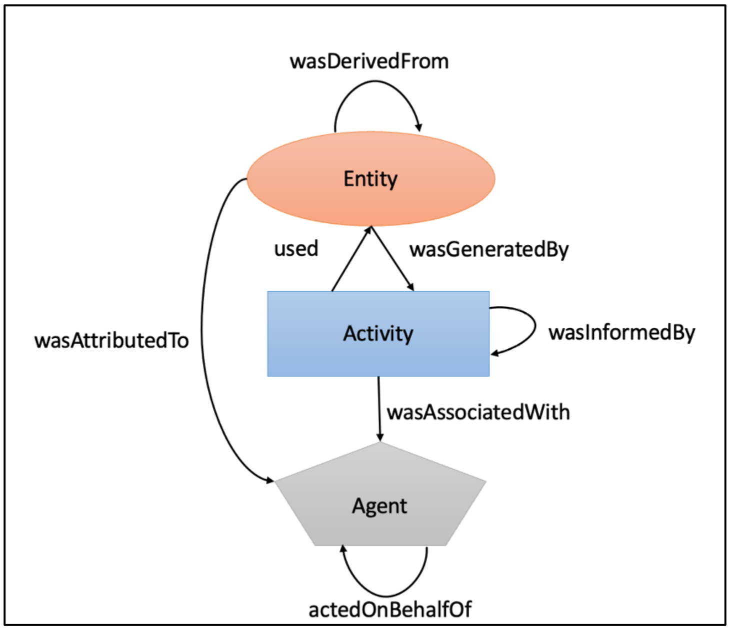

Figure 3.3: The three top classes of PROV-O model and properties. (Above image adapted from W3C PROV family of Documents, (Groth and Moreau, 2013))

3.3 Need for Provenance in Earth AI

3.3.1 Use of AI in the Earth Science Domain

In recent years, AI has been widely used to improve or replace conventional tasks in earth science domains. These methods have proven to be effective in performing various tasks like climate models, anomaly detection, weather prediction, event classification, and space weather, raising the expectation that AI could address some of the major challenges in earth science (Rasp et al., 2018). According to Intel's survey from 2018, 74% of the respondents indicated that AI would support solving long-term environmental concerns (Intel study, 2018). As a response, Intel is on board, pledging to restore 100% of its global water use by 2025. According to the article published by Columbia University's Earth Institute (Cho, 2021), AI has assisted farmers in India in increasing groundnut yields by providing knowledge on how to prepare the field, apply fertilizer, and choose sowing dates, resulting in a 30 % increase in yields per hectare. According to the same article, Norway uses AI in the development of a flexible and autonomous power grid that incorporates more renewable energy. Microsoft's AI for Earth program, launched in 2017, seeks to provide 200 research grants totaling $50 million to projects that use AI to address environmental challenges (Microsoft's Earth AI program, 2017). IBM's Green Horizon project in China utilized an AI system to predict pollution levels, track pollution sources, and generate potential solutions to drastically lower the pollutants (IBM Green Horizon Project, 2016). The increasing attention to AI/ML and earth science has also been reflected in the records of publications for the past 10 years. Figure 3.4 illustrates the exponential growth of relevant publications found in Scopus.

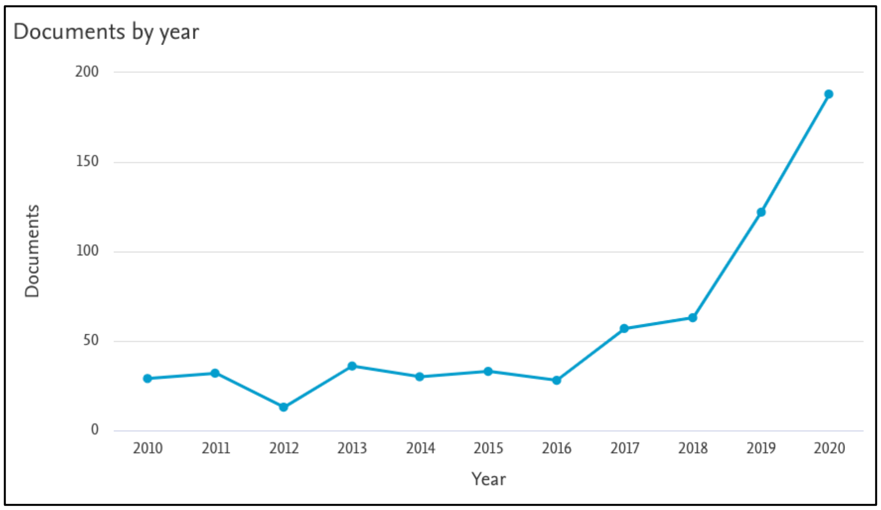

Figure 3.4: Distribution of publications line graph

(Distribution of publications (01/2010 – 12/2020) whose title, abstract and keywords include "earth science" and "artificial intelligence" or "machine learning". The query below was used to extract results from the Scopus database on Dec 29th, 2021.

(TITLE-ABS-KEY (earth AND science) AND TITLE-ABS-KEY (machine AND learning)

OR TITLE-ABS KEY (artificial AND intelligence)) AND PUBYEAR > 2009 AND PUBYEAR < 2021)

3.3.2 Related Work in Provenance and Earth Science

As the amount of data is increasing in the earth science domain, there are numerous initiatives taken to extend and improve practices in preserving provenance. For instance, Lanter (1991, 1993) developed a meta-database system for tracking the process of workflow and a system (Geolineus) for recording Geographic Information System (GIS) operations. Governments and other funding organizations have expressed the need for provenance and have been developing policies related to documenting provenance. In 1998, NASA established the Federation of Earth Science Information Partners (ESIP) to involve a larger group of stakeholders in improving techniques for storing, searching, accessing, and using earth science data (Showstack, 1998). ESIP has also initiated standard practices for reusability of data by addressing the issues of provenance in earth science (e.g., Duerr et al., 2011; ESIP Stewardship Committee, 2019; Downs et al., 2015; Mayernik et al., 2015). The Third US National Climate Assessment (NCA3) published in 2014 undergo a rigorous review process to ensure transparency and credibility (Garfin et al., 2013; Tilmes et al., 2013; Ma et al., 2014a,b). Earth scientists have proposed standards to document the provenance of both data and scientific workflows (Sun et al., 2020, 2022). Among all those endeavors, the development of Semantic Web technologies, especially those together with the W3C PROV, has been proven to be an efficient way for representing and documenting provenance (Di et al., 2013, Moreau and Groth, 2013).

Table 3.1: New initiative taken by the research community to extend provenance in Earth Science

References	***New Initiatives taken to extend provenance in Earth Science***
Lanter 1991, 1993	Developed a meta-database system for tracking the process of workflow and a system for recording GIS system.
Showstack, 1998	NASA established the Federation of (ESIP) in improving techniques for storing, searching, accessing, and using earth science data
Duerr et al., 2011; ESIP Stewardship Committee, 2019; Downs et al., 2015; Mayernik et al., 2015	ESIP initiated standard practices for reusability of data by addressing the issues of provenance in earth science.
Garfin et al., 2013; Tilmes et al., 2013; Ma et al., 2014a,b	The Third US National Climate Assessment (NCA3) published in 2014 undergo a rigorous review process to ensure transparency and credibility.
Sun et al., 2020, 2022	Earth scientists have proposed standards to document the provenance of both data and scientific workflows
Di et al., 2013, Moreau and Groth, 2013	The development W3C PROV, has been proven to be an efficient way for representing and documenting provenance

The growing use of data across all domains is leading to an era of data abundance. Abundant data can certainly be helpful for AI/ML models to generate predictions and results, but only when their roots are known. Provenance is the bloodline of data; it provides historical context and ensures authenticity. In real-world practices, provenance allows users to understand where the data came from, how it was collected, what important steps were involved in creating a result. Further, it allows AI/ML models to be effective and trustworthy (Moreau et al., 2008; Moreau and Groth, 2013). Despite this, compared with the widespread of AI/ML systems and data science applications, only a limited number of research projects have fully implemented provenance in their system architecture. Moreover, according to the recent poll, most academics are still unfamiliar with metadata standardization in their field (Tenopir et al., 2020). We believe data is a suspect when origins cannot be verified. In a recent example (Eisenman et al., 2014), an unreported change was discovered in the Antarctic sea ice cover that seemed to be increasing at a significant rate. Later, it turned out that the volume of Antarctic sea ice does not seem to be rising at the same rate as previously anticipated. In reality, much of the past artifact could have been caused by a satellite observation error that was undocumented. This mistake could have been avoided if the provenance of the data was fully recorded. Transparency and verifiability are the essential components in research and scientists must always verify if the data is trustworthy before analysis. Earth science community must keep up the rapid pace in building AI/ML models that are more transparent by design. As this research field is still at its early stages of development, it provides lots of scopes for the research community to design new models that are explainable, transparent, and most importantly reproducible. Therefore, with this chapter, we want to emphasize that provenance has the potential and needs to be considered in future earth science applications.

3.4 Technical Approaches

There is a wide variety of software tools developed to support reproducible results and provenance tracking in scientific research, such as Kepler (Altintas et al., 2004), DataRobot (DataRobot, 2012), Neptune (Talia, 2013), Datatron (Datatron, 2016), Metaclip (Bedia et al., 2019), Geoweaver (Sun et al., 2020), Amazon SageMaker (Das et al., 2020) and Collibra (Hilger and Wahl, 2022). These tools provide several ways to explore the provenance repository by tracking model activity, recording all changes in the data and the model, and outlining best practices for data management and disposal. This section gives a detailed introduction to three platforms, Metaclip, Kepler, and Geoweaver to illustrate the state-of-the-art technologies for ensuring the quality of data along with the workflow process in Earth and environmental sciences.

3.4.1 Metaclip (METAdata for CLImate Products)

Metaclip is a framework designed for documenting and presenting the provenance of climate products (Iturbide et al., 2019). An important part of it is the metaclipR package, which is developed as an additional component of the climate4R framework for climate data analysis in the R environment. The documented provenance is based on the Resource Description Framework (RDF) and focuses on semantic description of climate products such as maps, plots, or datasets, allowing each product and its provenance information to be delivered jointly. Thus, Metaclip enables climate data users to assess the quality, reliability, and trustworthiness of the data that they are using (Bedia et al., 2019). The documented provenance is understandable to a wide range of users from professionals who require technical details to common users who are interested in higher-level information. To realize that Metaclip has developed four core vocabularies (Datasource, Calibration, Verification, and Graphical_output). These vocabularies extend the basic PROV-O classes and properties, providing appropriate descriptions in order to offer meaningful information involved in the creation of specific data products.

- Datasource: This vocabulary indicates the source of input data as well as the change it has experienced throughout, like subsetting, aggregation, anomalies, principal component analysis

(PCA), and climate indices. It also provides the connections between the various transformation commands and arguments used in each stage.

- Calibration: This vocabulary translates the metadata describing bias correction, downscaling, and other forms of statistical adjustment. The calibration vocabulary is based on the framework designed by VALUE (Gutiérrez et al., 2019; see also: http://www.value-cost.eu/), which aims to systematically validate and improve downscaling methods in climate research.
- Verification: This vocabulary includes metadata related to the verification of seasonal forecast products, providing a description of the verification measures applied, as well as describing the verification aspect that each measure addresses. In addition, this vocabulary also includes a theoretical model for describing other forms of climate validation.
- Graphical_output: This vocabulary aims at describing graphical products, like charts and maps, including a characterization of uncertainty types and how they are communicated.

The example below is taken from the Metaclip interpreter (http://www.metaclip.org/interpreter) to highlight the main functionality of metaclipR in describing a complete workflow. The architecture of the interpreter is designed based on two components: (1) a backend service to extract and parse provenance information and (2) a front-end component that handles interactive visualization. The image on the left of Figure 3.5 is a climatological map of specific indexes. All the provenance information of the workflow generating that image is encoded in RDF and embedded in the JPG file. For such a provenance-embedded climate product, the Metaclip interpreter can extract and visualize the provenance information at different levels of granularity (right part of Figure 3.5). For instance, it can provide high-level provenance information about the climate product like how the image was generated, who was involved in producing the image, and steps in the workflow process. Then, by clicking on the corresponding node for a step in the workflow, a user can get additional technical details and a full description of the data or code.

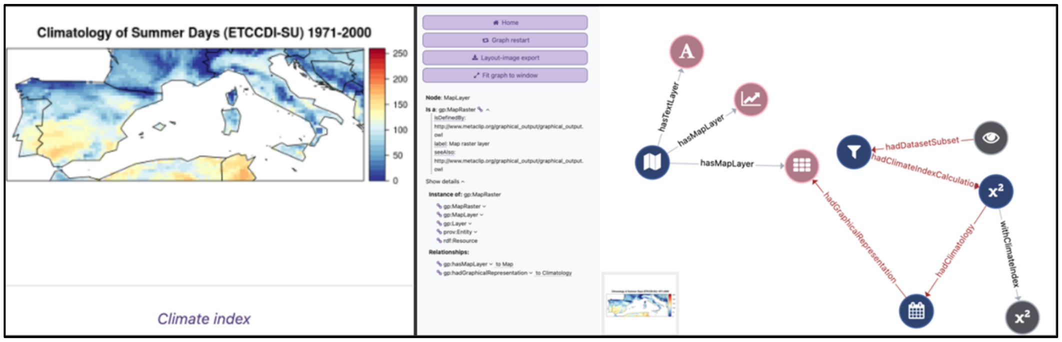

Figure 3.5: Metaclip web interface

(Metaclip interpreter displaying a climate index product on the left and provenance visualization on the right. By double-clicking on each of the nodes in the provenance network the user can get additional details and sub-properties for the node. The image is captured from the Metaclip interpreter website: http://www.metaclip.org/interpreter)

3.4.2 Kepler Scientific Workflow System

Kepler is an open-source, Java-based collaborative platform for scientists from all disciplines (Altintas et al., 2004). It is one of the most widely used scientific workflow systems in a variety of projects to manage, process, and analyze scientific data. The goal of Kepler is to simplify the workflow creation and execution process so that scientists can quickly design, monitor, re-run, and discuss analytical methods with minimal effort. Kepler is built upon the Ptolemy II system which is an actor-oriented design methodology (Eker et al., 2003). From the underlying Ptolemy II system, Kepler inherits several advanced functionalities such as variable (director-based) execution models, nested workflows, and the Vergil graphical user interface (GUI) which makes the system more versatile. Kepler's GUI is powerful yet user-friendly for both engineers and end users allowing efficient solutions for scientific problems in any domain (Figure 3.6).

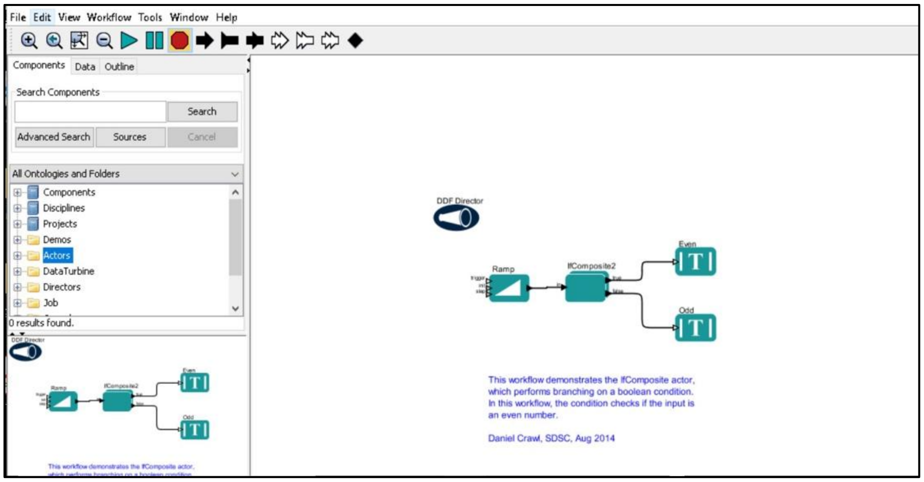

Figure 3.6: The demo workflow from Kepler's software

(The demo workflow from Kepler's software. The sample workflows can be found on the website: https://kepler-project.org/users/sample-workflows.html)

3.4.3 Geoweaver

Geoweaver is a web-based workflow system that helps AI practitioners to integrate multiple steps such as preprocessing, training, testing, and post-processing into a single automated workflow (Sun et al., 2020). It provides great benefits to earth scientists and has the capability to run, modify, reproduce, share, track, and reuse AI workflows in a single or distributed environment. Geoweaver runs on a web browser and can be installed by any individual or group to manage their resources. The core design of Geoweaver is divided into three modules (Host, Process, and Workflow) which help AI practitioners sort their experiments and allow reusability.

- Host: This module is the foundation of the entire framework. It allows users to connect different resources like SSH (Secure Shell), Jupyter Lab, and third-party computing platforms such as Jupyter Notebook Server, Google Earth Engine, and Google Colab.
- Process: This module supports Python, Jupyter Notebook, Shell script (bash), and SSH for running different programs. It allows users to create new processes and edit existing ones.

- Workflow: This module enables users to create workflows from the existing processes. The graphical panel allows users to drag and drop different processes and link them into a workflow. Once the workflow is created it can be downloaded, uploaded, and edited.

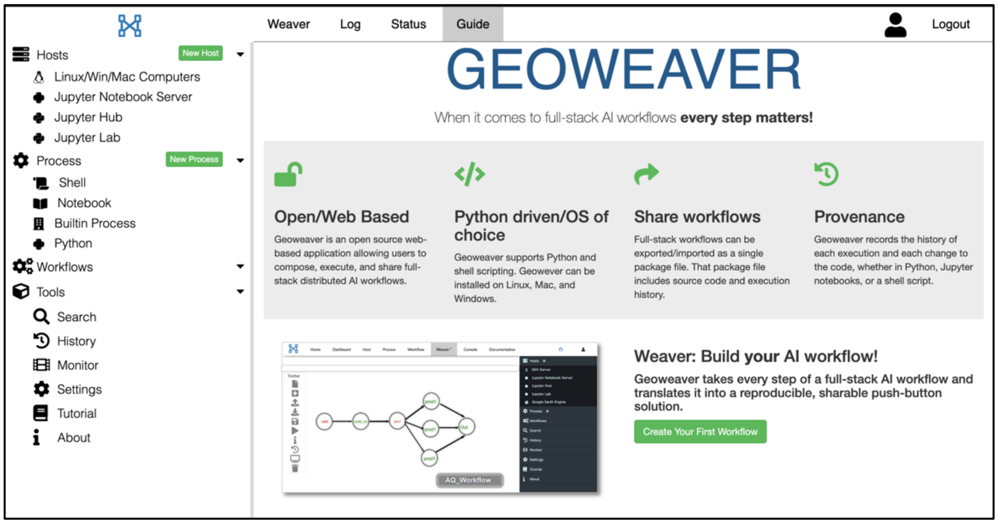

Figure 3.7: The user interface of Geoweaver

(The user interface of Geoweaver. It consists of a workspace canvas, a top main menu, a left-side menu panel, and a logout window. There are three main folders in the Left menu panel: Host, Process, and Workflow. They each include a list of child nodes with multiple options. The Host consists of a machine that can be either physical or virtual. The Process consists of four child folders: Shell scripts, Python code, Jupyter Notebooks, and built-in processes. The Workflow folder gives users the flexibility to create a new workflow based on the existing processes. Geoweaver makes all these entities manageable in one place and can help the AI community to share, track and reproduce results. The live demo of Geoweaver can be found on the website: https://geobrain.csiss.gmu.edu/Geoweaver/)

Here is the step-by-step guide to connecting to a host, creating processes, managing workflows, and viewing recorded provenance in Geoweaver.

Step 1: Connect to the host

In Geoweaver, hosts are the computing platforms where we connect to different recourses, in this example, we have established the connection through localhost. Once the connection is successfully established, we may proceed to our next step i.e., creating a process.

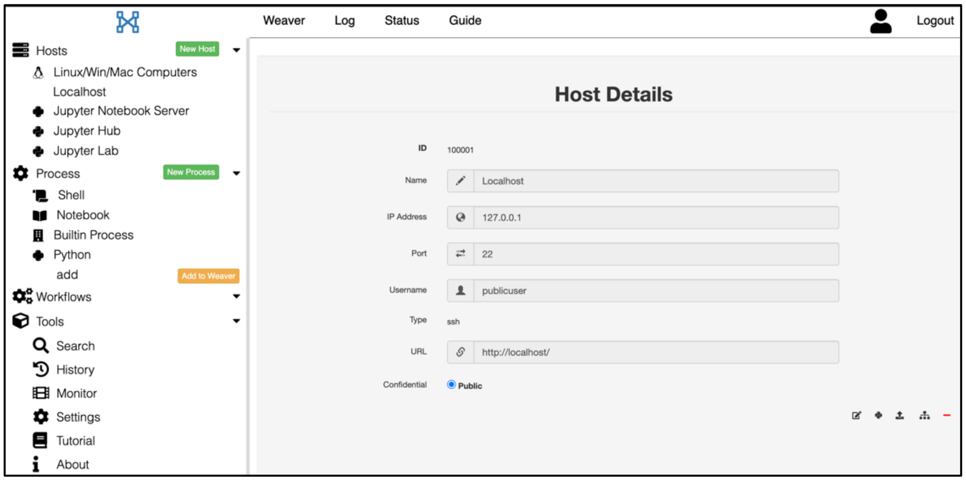

Figure 3.8: Connecting to host in Geoweaver

Step 2: Create a Process

Processes are programs like Python code, shell scripts, Jupyter notebook, and Google Earth Engine scripts. All of the code can be managed in one place and executed on different hosts. In this example, we will be creating all our processes using Python. Once you finish writing the code click on the run button to execute the process. In the popup window, select "Localhost" and click "Execute". In the popup Python environment dialog, click "Confirm" to the default and add your password. If you see the output printed in the logging window, it means you have successfully created and run your first process in Geoweaver! Congrats!

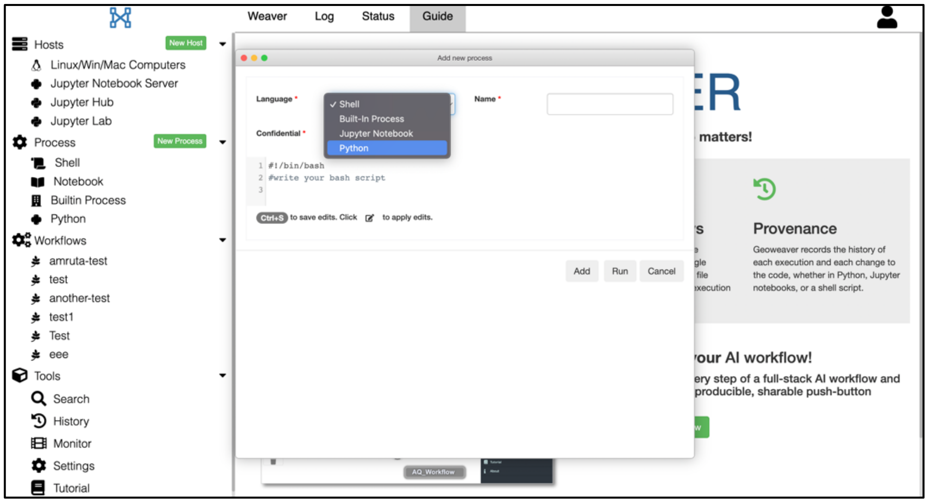

Figure 3.9: Writing the first Python program in Geoweaver

All the history of each process execution will be saved in the Geoweaver database, even if the used hosts are no longer available. This is how the provenance is documented in Geoweaver. The below image (Figure 3.10) displays the execution ID of each process, at what time the process was run, and information about whether the process ran successfully or failed.

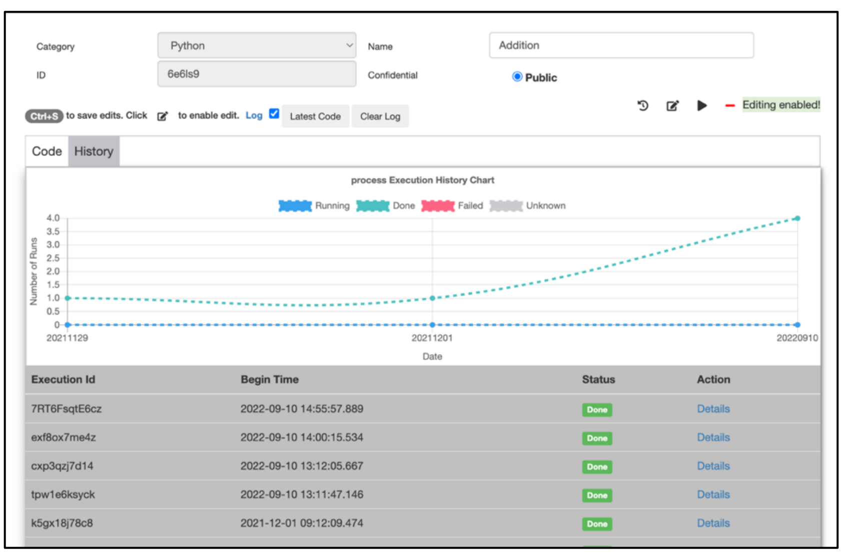

Figure 3.10: A Geoweaver dashboard to browse the provenance recorded for each process

(A Geoweaver dashboard to browse the provenance recorded for each process. Here category is the type of process we are executing, the name is the given program name and ID is the unique identification of each program.)

Step 3: Create a Workflow

Normally, workflows are the connected pipelines of several (>2) processes, but Geoweaver also supports isolated processes, that is, processes not connected to each other are also allowed. Complex scientific experiments can simply be broken down into a number of workflows, which can then be executed and managed here. The weaver workspace allows users to create new workflows or edit the existing ones. The “add to weaver” button in the process module allows users to add different processes to the weaver workspace. In Figure 3.11 we have three processes created in Python. For better understandability we have demonstrated simple Python programs “addition” (addition of two numbers), “for_loop” (running a for loop to print a series of numbers), and “if_else” (print the greater number). To create a workflow the processes can be linked with each other by pressing and holding the SHIFT key on the keyboard. Once the workflow is created you can click on the “plus” button in the top floating bar, then add details in the popup window “Input Workflow Name” (write the name of your workflow) and a simple description in the “Description”, once you add all the information click confirm to complete. To run the workflow, click on the “play” button in the top floating bar, in the pop window select the “one-host” option, then choose localhost and set the environment to default. Finally, click Run, enter the password for localhost and confirm.

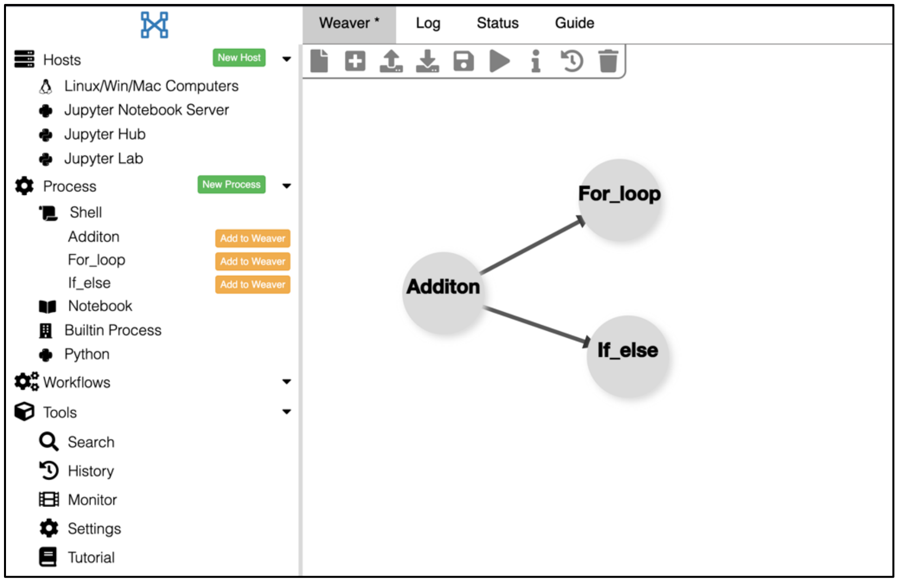

Figure 3.11: Demonstration of workflow created in Geoweaver using different process

While the workflow is in the execution mode, you will notice different colors: blue means the process is waiting, yellow means the corresponding process is running, green means the process execution is finished, and red means the process execution is failed for some reason.

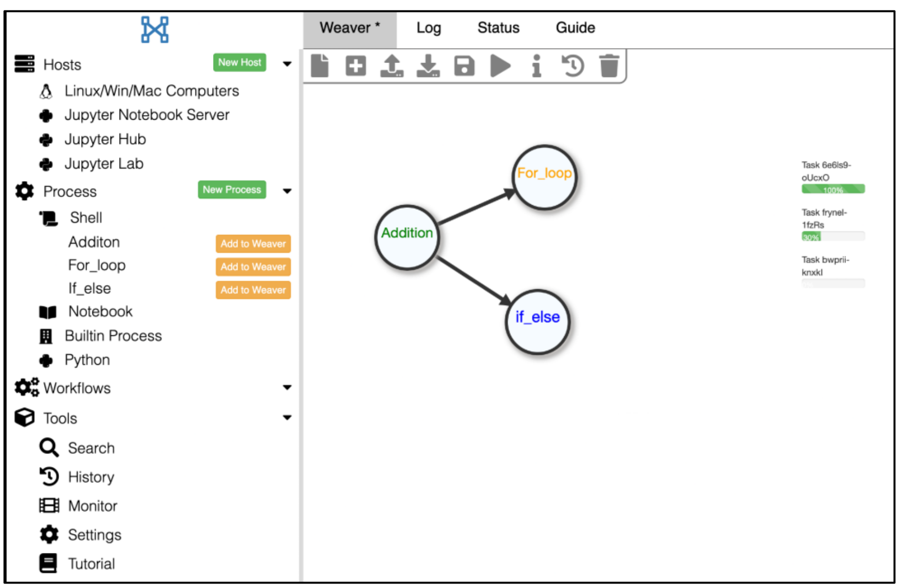

Figure 3.12: Web interface representing different colors in the execution mode. The image is captured from Geoweaver in-browser software

To export the workflow from Geoweaver click on the downward icon button in the top floating bar. The workflow exportations will provide two options “Workflow with process code” and “Workflow with process code and history”. The first option will only download source code and workflow json, but the latter will download source code, workflow json along with the detail history of the previous execution. The second option is recommended as it is provenance enabled. Click “Confirm” and a ZIP file will be automatically downloaded to your machine. Another best feature of Geoweaver is the ability to reproduce and edit the existing workflows. To import the shared workflow, click on the upward icon in the top floating bar and drag and drop the Geoweaver ZIP file and click “Start”. Once the uploading is finished and if the workflow file is valid, it will ask “The upload workflow is valid. Do you want to proceed to save it into the database?” “You can click OK, then the workflow will be automatically loaded in the workspace and ready to reuse.

All three platforms introduced above have their websites where detailed tutorials and sample workflows can be accessed, including examples in Earth and environmental sciences. Interested readers are suggested to go to their websites (see links in captions of Figures 3.7 to 3.10) to see and practice the different technical approaches for provenance or metadata documentation.

3.5 Discussion

Provenance in Earth AI is closely relevant to reproducibility. In our opinion, one of the most important aspects of making AI/ML more reproducible is to record or document all the core primitives such as hyperparameters, model architecture, code commits, datasets, and all the metadata associated with the training process. We understand there are plenty of different factors such as data changes, different software environments or versions, and numerous other small variations that can result in a reproducibility crisis. As it is not necessary to document all the detailed information, AI practitioners must prioritize documenting the most important elements of a project from day one to enable other researchers to easily reproduce their work when necessary. Adding a standardized format of reproducibility also ensures efficiency and accuracy. This will not only help researchers to reproduce results but will also ensure transparency and trust. Beyond documenting the fundamental components of an AI/ML system, the concept of reproducibility can be viewed as a systematic way of working in data-intensive Earth AI.

Another advantage enabled by provenance is interpretability. Although interpretability and explainability are often discussed interchangeably, interpretability is concerned with the factors that influence a model's decisions, while explainability deals with the reasoning process that a model follows to arrive at a final decision. The need for interpretability has been highlighted in many studies, especially when the decisions made by AI/ML algorithms have generated unintended biased, discriminatory, and even harmful outcomes. This issue has raised concerns of transparency and ethics for AI practitioners, such as when algorithms are deployed in critical domains like healthcare. In our opinion, interpretability is a prerequisite for humans to trust AI/ML models. It allows users to understand the causes behind decisions of real-world AI/ML applications, and thereby improve the fairness of the models. Enabling interpretability in AL/ML models will improve confidence and trust in the model. It will help data scientists to draw explanations from the black-box model for why certain decisions or predictions have been made. Recent techniques such as LIME (Local Interpretable Model-Agnostic Explanation) and SHAP (SHapley Additive exPLanations) show great promise for model interpretability. However, there is still ample room for improvement from a data science and engineering perspective. In order to understand opaque models, we need new initiatives and techniques to design systems that are safe, robust, ethical, and most importantly, interpretable.

As AI/ML systems expand to include more diverse applications, the need for capturing provenance is gaining traction in the research community. We believe that the inclusion of provenance will not only strengthen AI/ML systems but will also improve transparency and explainability. Adding domain-specific documentation standards can help the community to grasp and begin employing appropriate practices routinely. In our opinion, data only adds value when it is accompanied by provenance information (for example, Wikipedia is not considered a trustworthy source due to the fact that many of its sources cannot be verified). Relying purely on data without verifying its source could be an unhealthy practice. On the other hand, documenting the necessary detail of a workflow will help researchers with troubleshooting in the event of errors, shedding light on the behavior of the model. It is worth noting that, manual documentation can put model provenance at risk, particularly when working with large datasets. For this reason, we encourage automating the process of provenance tracking by using workflow platforms, tools, and packages to limit manual operation. Fortunately, more adequate tools for recording and sharing provenance information are already being developed, which will greatly facilitate the automation of provenance documentation. The future of provenance documentation looks promising because governments and funding organizations are already recognizing the need for data preservation and provenance and are increasingly providing guidelines and support for works in that direction.

Looking into the future, we propose a few points for discussion. The first is that AI/ML systems need to be adaptive and interactive, providing explanations based on users' needs, expertise, and requirements. The success of W3C PROV is a perfect example of making AI/ML processes reproducible or repeatable. However, we believe that as the research progresses, we will need more adoption of provenance standards to enable open science in various disciplines, including those in earth science. Our second point is that more functions of data management and provenance documentation need to be enabled in workflow platforms. To speed up scientific research and comprehension, open data will allow researchers to share data, information, and expertise (with clear licensing) enabling transparency and reproducibility. The increasing availability of open data comes with the need for data management. However, the humongous amount of data available today makes manual management an unrealistic approach. There is the need for more platforms like Geoweaver which automate the AI/ML workflow and enable users to perform all tasks in one place more efficiently. The third point is about leveraging cloud service in data-intensive Earth AI. Many of the existing automated data management and analysis platforms are cloud-based, and we will undoubtedly see a continuation of rapid adoption and growth of cloud platforms. This will likely result in a shift in focus from complex high-cost local computation to cloud computing. This fact will be the primary motivator for the researchers to migrate to cloud platforms.

3.6 Conclusion

With the increasing adoption of AI/ML systems, there is an increasing need for the results to be interpretable, reproducible, traceable, and explainable. Although the post hoc explainability approaches could be one way to explain a black-box model, we believe they are still in their infancy stage and not completely reliable. We suggest adopting established methods from the field of data and software provenance will be an ideal solution for providing explanations to AI/ML systems. Provenance will not only help users to trace, evaluate, understand, and reproduce the AI/ML results but will also enhance users' decisions about how much trust to place in data and results generated from the original sources. In this study first, we presented a summary highlighting the fundamental concepts of XAI, TAI, and provenance. Second, we discussed how AI/ML models have advanced in the earth science discipline and the related work in provenance. Third, we illustrated three different tools that support reproducible results and provenance tracking. Lastly, we presented a research outline in the discussions to analyze the challenges and suggest further research opportunities. We hope this chapter provides enough justice to the importance of provenance and gives insight into new tools and the progress that can be made in AI/ML systems. We believe that provenance remains an important topic and has much more to offer the earth science community.

Chapter 4 Geoweaver_cwl: Transforming Geoweaver AI Workflows to Common Workflow Language to Extend Interoperability

This chapter is adapted from the preprint:

Amruta Kale, Ziheng Sun, Chao Fan, Xiaogang Ma; "Geoweaver_cwl: Transforming Geoweaver AI Workflows to Common Workflow Language to Extend Interoperability", 2022. SSRN Elsevier Preprint available. DOI: http://dx.doi.org/10.2139/ssrn.4284586

4.1 Introduction

We are witnessing a widespread adoption of AI/ML in our everyday life. The recent success of DL has largely contributed to the huge success of AI/ML models. DL algorithms are widely used in mission-critical applications like healthcare, autonomous robots and vehicles, image classification, and detection. Despite the significant improvement in performance and predictions, the black-box nature of DL algorithms can raise social and ethical questions about their operations and results. Even the programmer designing the complex AI/ML model finds it difficult to gain insight into an internal system that is often opaque. This issue has extended the research focus from improving accuracy to explainable and interpretable ML models (Doshi-Velez et al., 2017; Gilpin et al., 2018; Adadi and Berrada, 2018; Wing, 2020; Sun et al., 2022).

Recent interests in XAI and TAI have achieved great momentum in making AI/ML models more explainable, interpretable, and transparent (Adadi and Berrada, 2018; Rudin, 2018; Rudin, 2019; Wing, 2020). XAI proposes a shift toward more transparent AI. It aims to develop a set of strategies to make ML models more explainable while maintaining their high predictive accuracy (Ribeiro et al., 2016, Gunning and Aha, 2019). Several strategic plans have shown the growing dynamics in the field of XAI. For instance, Local Interpretable Model-Agnostic Explanations (LIME) (Ribeiro et al., 2016) and SHapely Additive exPLanations (SHAP) (Lundberg and Lee, 2017) are two popular approaches to explaining ML and DL models. SHAP values can be used to explain a variety of models including both classification and regression problems. While both LIME and SHAP are model-agnostic approaches, SHAP is widely used and more acceptable because it guarantees the fair distribution of contribution for each of the variables. Government agencies like GDPR, and the AI HLEG have also highlighted the importance of transparent and fair AI systems (Goddard, 2017, AI HLEG, 2019).

As the field of XAI continues to expand, it is important to develop new research strategies that include the provenance of upstream steps and history model runs. The diverse nature of AI/ML models in the field of XAI requires a multi-disciplinary approach, and in our previous papers, we highlighted the importance of provenance documentation and its benefit for AI/ML models (Kale et al., 2022; Kale and Ma, 2022 In press). We suggest that adopting approaches and methods from the field of provenance will help to generate resourceful explanations and improve reproducibility (Ma et al., 2017; Kale et al., 2022). When data is captured or generated, it is always accompanied by a plethora of additional information, which is generally referred to as metadata. Data type, document size, or any other information that specifies a property of the actual data can all be considered metadata (Baca, 2016; Ma, 2022). Provenance, on the other hand, might be considered an expansion of metadata. It includes information about the origin of the data and any other objects of interest, as well as information about the workflow (Cheney et al., 2009; Tilmes et al., 2013; Ma et al., 2014). The term provenance is defined by World Wide Web Consortium (W3C) (Moreau and Missier, 2013) as: *"Provenance is the information about entities, activities, and people involved in producing a piece of data or a thing which can be used to form assessment about its quality, reliability, or trustworthiness".* Provenance documentation answers questions like "Who is the author?", "How was the data created?", and "What earlier processing steps were done to the data before it reached the current form?".

There are a wide variety of computing environments and software tools designed to track and manage AI/ML experiments. These tools provide several ways to explore the provenance repository by tracking model activity, recording changes in the data and model, and outlining best practices for data processing. Kepler (Altintas et al., 2004), DataRobot (DataRobot, 2012), Datatron (Datatron, 2016), Metaclip (Bedia et al., 2019), and Amazon SageMaker (Das et al., 2020) are some of the tools which fall under these categories. However, to the best of our knowledge, none of these tools provides satisfying solutions for full-stack AI/ML workflow management, along with full access to the files, data flow, code, and most importantly, enabling reproducibility and portability of workflows. Therefore, in this paper, we emphasize the new workflow management system (WfMS) named "Geoweaver" which enables users to share, replicate, and reuse AI/ML workflows all in one place more efficiently. Geoweaver supports connecting all the preprocessing steps, training and testing of AI/ML models, and post-processing steps into a single automated workflow (Sun et al., 2020).

To further extend the interoperability of AI/ML workflows, in this paper we present the python package ***geoweaver_cwl***, a wrapper tool that transforms Geoweaver AI/ML workflows into Common Workflow Language (CWL) scripts in a way that makes them portable and scalable across a variety of software and hardware environments. This paper will describe how the tool is developed and implemented in our use cases and is organized as follows. In section 2, we first describe an overview of CWL, followed by the conceptual framework of Geoweaver, and then describe the architecture of ***geoweaver_cwl***. In section 3, we demonstrate our python package by applying a use case from the Geoweaver platform and assess the quality of the package and its influence in Geoweaver. In section 4, we discuss the importance of adopting CWL standards and highlight the future direction of our work. Finally, we conclude with a few additional remarks.

4.2 Technical Framework of the geoweaver_cwl Package

4.2.1 An Overview of the Common Workflow Language

CWL is a community standard to describe command-line-based workflows (Amstutz et al., 2016, Crusoe et at, 2022). It offers a typical but simplified set of generalizations that are commonly implemented in many popular WfMS. The language's declarative format enables users to describe the process of executing diverse software tools and workflows through their command-line interface. Previously, to link the command-line tools researchers need to write shell scripts. Although these scripts offer an efficient approach to accessing the tools, writing, and maintaining them requires specialized knowledge. As a result, researchers spend more time maintaining the scripts than conducting their research. However, with the increase in workflow popularity, the number of workflow management tools has increased, and each of them has its standards for specifying the tools and workflows. This has reduced the portability and interoperability of these workflows. CWL aims to reduce the barrier to researchers using these technologies by providing a standard to unify them. In order to ensure reproducibility, CWL standards explicitly support the usage of container technologies. These standards ensure portability, so the same workflow can be executed in both local and high-performance cloud environments.

4.2.2 Conceptual Framework of the Geoweaver Workflow Management System

Geoweaver is a stable practical platform for NASA's Earth Observing System Data and Information System (EOSDIS) that enables earth scientists to manage, share, replicate and reuse their AI/ML workflows. Geoweaver helps AI practitioners by providing more sophisticated AI workflows that not only include data preprocessing, training, and testing of AI algorithms but also post-processing of results into an ad hoc automated workflow (Sun et al., 2020). Earth scientists increasingly begin to manage their workflows, but because of the uncertainty and complexity of AI/ML models, scientists often find solo management problematic, particularly when massive data is involved. Geoweaver offers a great solution to these problems. The fundamental design of Geoweaver is organized into three modules (Host, Process, and Workflow), which enable AI practitioners to sort and reuse their AI/ML experiments.

- Host: This module serves as the cornerstone for the framework, differentiating it from other WfMS. It enables users to connect to several resources such as virtual machines, Jupyter server instances, Secure shell (SSH), and third-party computing platforms like Google Earth Engine, Jupyter Notebook Server, and Google Colab. Additionally, the file transfer services (file uploading from local computers to remote servers, and file downloading from remote servers to local computers) provided by the host module allow users to transfer their workflow from one platform to another.
- Process: This module includes five submodules and one database. As most of the current AI/ML experiments employ Python programming, the process module supports Python, Jupyter Notebook, Shell scripting (bash), and SSH for running system-level programs. All the dependent libraries like DeepLearning, Keras, PyTorch, and TensorFlow are easily accessible in the process. The process editor/creator interface allows users to create new processes and edit existing ones. Whenever a new process is created, it gets stored in a MySQL database. The process monitor is responsible for all the execution events in the process module and reports the real-time status. Once the execution is complete the input, output, and code that has been executed will be recorded and stored in a database. The provenance manager is responsible for evaluating the recorded history of each process in order to assess data quality and recover from failure.
- Workflow: The term "Workflow" is a wide-ranging phrase that can be interpreted in a variety of ways (WMP, 1998; S. Jablonski and C. Bussler, 1996). For instance, many geoscientists often refer to Jupyter Notebook or bash script as a workflow. In Geoweaver, workflow denotes a pipeline linking multiple processes together. The workflow module consists of two functions (1) Building workflows from the existing process and (2) Managing the query, edits, and execution of the workflows. The workflows in Geoweaver are Directed Acyclic Graphs (DAG), which means the

workflow follows a certain direction. Once the workflow is created it can run using localhost or by setting the default environment. The workflow module displays a color-coded real-time status of each process in the execution mode. Different colors represent the status of each process: blue means the process is waiting; yellow means the process is running; green means the process is finished running; and red means the process failed. A more detailed demonstration of Geoweaver is described in a previous paper (Kale and Ma, 2022 In press). Exporting and importing the existing workflows in Geoweaver is simple and easy. The downloaded workflow can be automatically loaded into the workspace and ready for execution and reuse. Figure 4.1 and Table 4.1 describes the framework of Geoweaver with the three core modules.

Table 4.1: Geoweaver demonstrating different modules and features

	Modules	***Features***
Geoweaver	Host (This module enables users to connect to several resources)	Virtual machines
		Jupyter server instance
		Secure Shell
		Google Earth Engine
		Jupyter Notebook Server
		Google Colab
	Process (This module includes five submodules and one database)	Python
		Jupyter notebook
		Shell Script
		Secure Shell
		MySQL
	Workflow (This module allows a pipeline linking multiple processes together)	Build workflows from existing process
		Manage query, edits, and execution of workflow

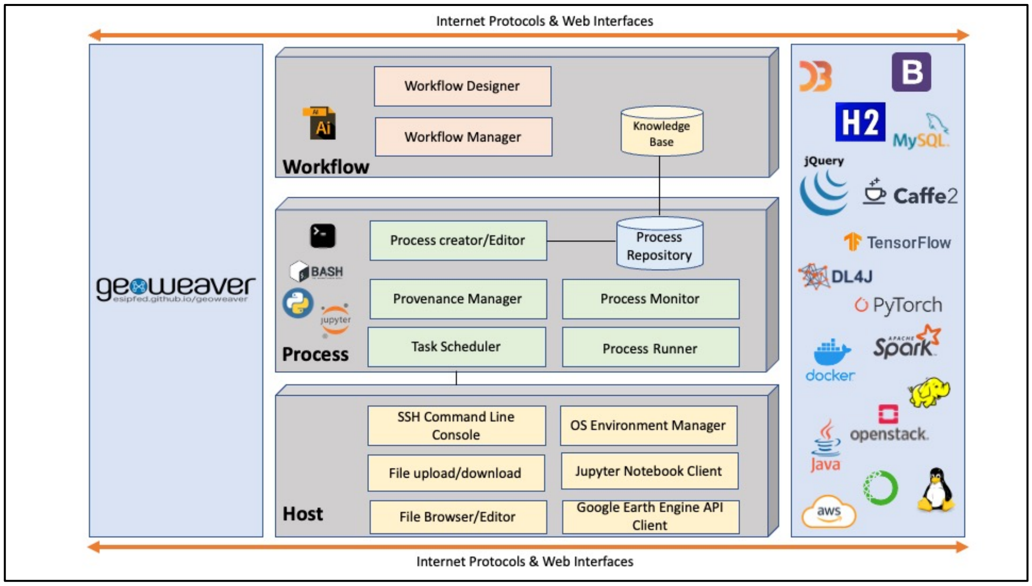

Figure 4.1: Workflow management framework of Geoweaver and its core modules (Host, Process, and Workflow), adapted from (Sun et al., 2020)

4.2.3 Architecture of the geoweaver_cwl Wrapper Tool

A wrapper tool is used to simplify the process of editing, running, and sharing tools in a digital environment. In this paper, we regard "wrapper tool" as a process of using CWL to describe command-line tools so that they can be run as an application or a tool in part of a larger workflow. Using the wrapper tool with CWL will make all the documents portable, sharable, and executable. The preliminary step for creating a workflow in Geoweaver is through the workflow module. The workspace allows users to compose a workflow using existing processes. Once the workflow is created it can be downloaded with two options "workflow with process code" or "workflow with process code and history". The first will simply download the workflow and source code. The latter will download all the history of the prior workflow executions in addition to the source code and workflow. The downloaded workflow comes with a Zip file that includes a code folder, a history folder, and a workflow file. The code folder contains the code files (processes) used to form the workflow, the history folder contains the historical details of each process like the begin_time, end_time, input, and output. The workflow file contains the information on the nodes and edges that link together to form the workflow.

To further extend the portability and interoperability of workflows built in the Geoweaver framework, we designed ***geoweaver_cwl***, a python package that captures inputs (source and target processes) from a Geoweaver workflow file and transforms them into CWL scripts. A key contribution to our work is an add-on functionality that dynamically generates corresponding CWL code without the user having to know the CWL syntax. The CWL file features text fields that comprehensively describe workflow commands and parameters.

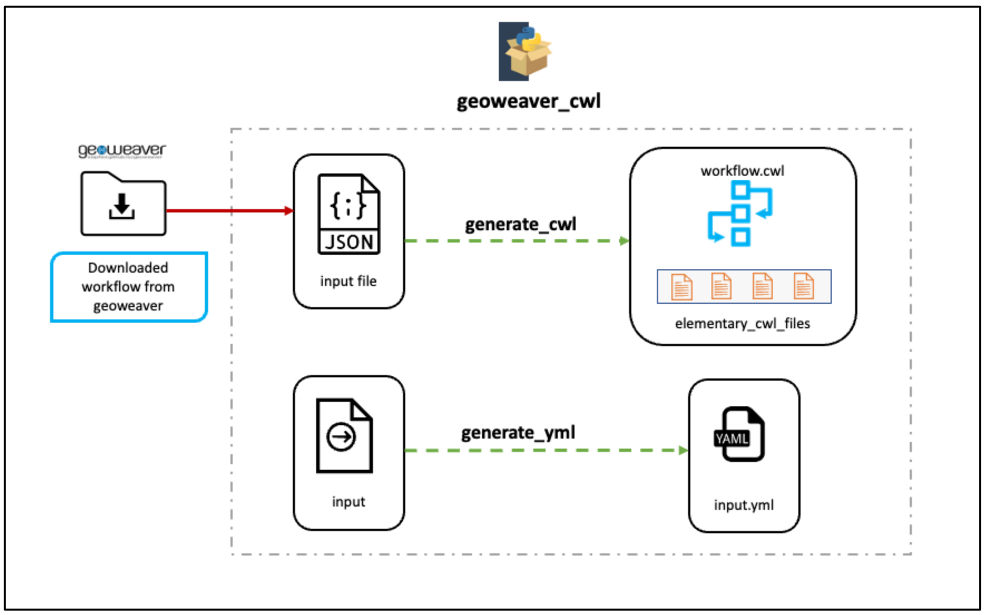

Figure 4.2: Architecture of geoweaver_cwl package with key functions

Figure 4.2 illustrates a brief architecture of ***geoweaver_cwl***. The package contains two main functions "generate_cwl" and "generate_yml". The generate_cwl function takes workflow.json from Geoweaver as the input, captures the nodes, and the edges from the workflow, and writes the steps that form the data flow into CWL scripts. To capture the source and target from the workflow file, we iteratively visit each node in the workflow, and each visited node that has not been previously processed becomes a source node. Then, for each source node, we compile the child nodes, and each child node serves as a target for the source node. Each source-target pair is processed by writing a CWL script that provides explicit inputs and outputs for each phase. Carrying out this procedure eventually enables us to generate the CWL scripts for the whole workflow. Equations 1 and 2 below describes the process of translating the workflow file into CWL.

$$PL = \rho(workflow.json) \tag{1}$$

$$[workflow.cwl, elementarycwlfiles] = \forall p: W\big(p, \in (p)\big), \qquad p \in PL \tag{2}$$

where PL = process_list

ρ = Graph edge extraction function

W, ∈ = file writing functions

Additionally, the function also generates a new subdirectory called "elementary_cwl_files" which stores new CWL files (the processes used in the workflow) translated from the code folder. Below is the pseudo-code of the generate_cwl and generate_yml functions.

Graph edge extraction function

```
Read edges from workflow.json
Let process_list, target_list be empty list
      For each edge in edges
      Let source be edge.source
      Let target be edge.target
            If source not in process_list
                  Append source to process_list
            If source is in target_list
                  Remove source from target_list
            If target not in target_list
                  Append target to target_list
Add elements from target_list to process_list
```

File writing function for workflow.cwl

```
read process_list -> workflow.json
for process in process_list
      write process_name
      write run command
      write input command
      write output command
      create an elementary_cwl_files
```

File writing function for elementary CWL files

```
Create a new elementary_cwl_files
      write baseCommand
      write input
      write output
```

The generate_yml function produces a Yet Another Markup Language (YAML) file, which writes the input to run the workflow.cwl file. The YAML file describes which input to run for the cwl files.

```
Class: Directory/file
Path: path of the file or directory
```

The ***geoweaver_cwl*** package is fully open access and the installation is simple. The package can be downloaded from: https://pypi.org/project/geoweaver-cwl/0.0.1/. Figure 4.3 demonstrates the installation steps for the ***geoweaver_cwl*** package along with the use of some functions. To facilitate reuse and adaptation, we have made the source code, a detailed user guide, and concrete self-contained examples file available on GitHub under an open-source license: https://github.com/amrutakale08/geoweaver_cwl and self-contained example on https://github.com/amrutakale08/geoweaver_cwl-usecases.

```
In [1]: pip install geoweaver_cwl

        Requirement already satisfied: geoweaver_cwl in /Users/amrutakale/opt/anaconda3/lib/python3.8/site-packages (0.0.9)
        Note: you may need to restart the kernel to use updated packages.

In [2]: from geoweaver_cwl import translator as tr

In [3]: tr.generate_cwl('workflow.json')

        Output file: workflow.cwl
        Writing header...
        Writing steps...
        CWL file written to workflow.cwl

In [4]: tr.generate_yml('input.yml')

        Writing YML file...
        YML file created ...
```

Figure 4.3: Installation and usage of the geoweaver_cwl package

Once the workflow files are described in CWL scripts, they can be executed using any other software that supports CWL, like cwltool, Arvados, Toil, CWL-Airflow, and more. Some advanced applications like Rabix are also available to run the process faster. Rabix is an open-source desktop application that allows researchers to create and edit CWL documents (Amstutz et.al, 2021). The application provides a text-based visual editor's interface which allows users to edit and create new workflows. However, many of the researchers who are unfamiliar with CWL may find Rabix's composer too hands-on and less intuitive. Therefore, in this dissertation, we are going to use the traditional cwltool. To run the newly generated CWL files from Geoweaver, we will use the below command. We invoke cwl_runner with workflow.cwl and input object input.yml on the command line.

```
cwl-runner workflow.cwl input.yml
```

The command will trigger all the functions inside the CWL and YAML files in the same order as Geoweaver and is supposed to get the same results. As mentioned above, the advantage of CWL is that it provides a solution for describing portable and reusable workflows. The transformation from Geoweaver to CWL through the ***geoweaver_cwl*** package allows geoscientists to easily share, exchange, modify, and reuse workflows. Additionally, CWL-compliant applications are highly portable and can be run in a variety of environments, including local or cloud infrastructures.

4.3 Use Case Implementation, Result, and Evaluation

Based on the ***geoweaver_cwl*** package, we tested a list of workflows from simple to complicated ones. Here we use a Geoweaver workflow available on GitHub (https://github.com/earth-artificial-intelligence/kenya-crop-mask-geoweaver) to demonstrate and verify the usability of our package. The scientific topic of that workflow is the annual and in-season mapping of cropland in Kenya (Tseng et al., 2020). The GitHub repository contains the code folder, history folder, and workflow.json file.

We installed the ***geoweaver_cwl*** package and followed the above-mentioned procedures to describe the workflow in the CWL text document. After using the functions ***generate_cwl*** and ***generate_yml***, we obtained the files "input.yml", "workflow.cwl", and "elementary cwl files folder", which included the cwl files used in creating the workflow. The workflow translation process was fast and easy, and we also noticed that using cwltool speeds up workflow execution compared to the original procedure in Geoweaver.

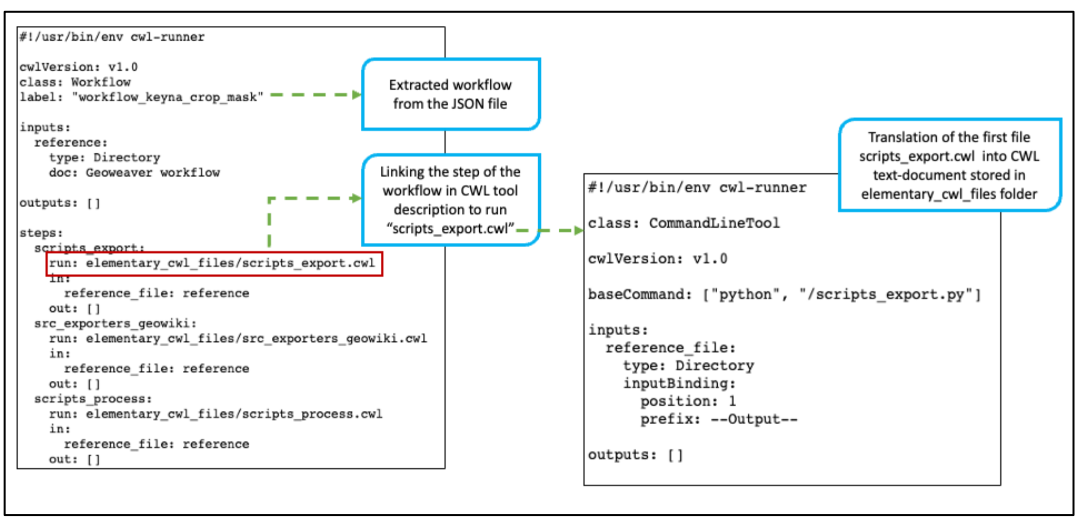

Figure 4.4: Exemplar scripts of workflow steps

(Exemplar scripts of workflow steps in the workflow.cwl from workflow.json file (left) and the CWL text document scripts_export.cwl describing computational steps present in the elementary_cwl_files folder (right))

We successfully transformed the Geoweaver workflow of Kenya cropland mapping into CWL format using the ***geoweaver_cwl*** package. The left part of Figure 4.4 shows the described workflow in CWL from the workflow.json in Geoweaver. The CWL file contains a ***cwlVersion*** section which indicates the version of the CWL document. The ***class*** section with a value of ***Workflow*** indicates that this document describes the workflow. The ***inputs*** and ***outputs*** sections describe the inputs and outputs of the workflow, respectively. The ***steps*** section describes the actual steps of the workflow. In this example, the first step is to run the "scripts_exports.cwl" present in the folder elementary_cwl_files. The code of "scripts_exports.cwl" is illustrated in the right part of Figure 4.4. The workflow steps in CWL do not always run in the written sequence. Instead, the order is determined by the dependencies across steps. To evaluate the result of the transformation we ran the CWL text document using cwltool, and we observed that it executed smoothly and generated the same result as in Geoweaver. The CWL result of this example is accessible on GitHub: https://github.com/amrutakale08/geoweaver_cwl-usecases. We are now transforming more Geoweaver AI workflows with this package and sharing the results on GitHub.

Geoweaver provides a unique combination of features, such as a user-friendly interface, full-stack code, a history of previous versions, and sharable AI/ML workflows. It is a user-friendly entry point to solve AI-related workflow issues for a variety of disciplines in geosciences as well as beyond. The ***geoweaver_cwl*** package developed in this work further extends the portability and interoperability of workflows created in Geoweaver. The package can quickly transform Geoweaver workflows into CWL format, and the result can be run on many CWL-compliant software applications. Moreover, the CWL result can be also executed on diverse computing platforms including local computers, cloud environments, or high-performance clusters. The transformation process is intuitive and new users will spend less time getting familiar with the package.

4.4 Discussion

We encourage geoscientists as well as other AI practitioners to use Geoweaver and the ***geoweaver_cwl*** package to increase the reproducibility and interoperability of their work. The developed package helps automatically transform Geoweaver AI/ML workflows to a community standard CWL. As an extension to Geoweaver, the CWL result can be executed on diverse computing platforms which gives users more opportunities to run the workflow without compromising provenance or having to recreate the workflow if they want to use another WfMS. CWL can formally describe inputs, outputs, and other execution details of the workflow in a text-based document. It supports workflows that specify dependencies among tools and use one device output as input to another. CWL documents are text-based so that they can be created manually, without or with less computer programming. However, ensuring that these documents adhere to the CWL syntax specification may restrict some users from adopting it. The developed ***geoweaver_cwl*** addresses this gap. It can automatically describe workflows into CWL to make it effortless for geoscientists to share data analysis workflows in varied formats without learning the technical details of the CWL syntax.

There are a wide variety of WfMS software tools available all over the research community, that are constantly being developed, revised, and improved every day. While the availability of such tools benefits the community, it also presents a great challenge: as more and more tools are created, a set of standards needs to be adopted in order to ensure the portability and reproducibility of the resulting workflows. CWL, as reflected in its name, aims to be such a community standard to harmonize the workflow formats proposed by various WfMS software tools. Reproducibility enables researchers to track and debug potential errors and validate the authenticity of the results, and as such it plays a vital role to make scientific research accurate, efficient, and cost-effective. Because CWL tracks code versions, inputs, outputs, and more, researchers can use it to pinpoint where the analysis went wrong, or where in the analysis the particular piece of data leads to new insights. Therefore, the transformation from Geoweaver workflows to CWL format is a necessary extension with regards to broad portability and reproducibility.

Portability is crucial when it comes to scientific research and analysis. When one workflow is designed for a type of computational environment such as a personal computer it may not function in a similar way as in the cloud. Therefore, researchers may spend more time and effort in debugging the tool to make it work in the desired environment. This could result in inconsistent outcomes or errors. In contrast, CWL enables portability by being explicit about inputs, outputs, data location, and execution models that can be executed on any of the CWL-compliant environments. CWL-based documents can be downloaded, edited, and executed on local infrastructure or uploaded and executed in the cloud.

The scientific provenance research community has evolved significantly in recent years to provide several strategic capabilities, to make AI/ML workflows more explainable and reproducible. The declarative approach to describe workflow in CWL scripts facilitates and encourages users to explicitly declare every single step, improving the white box view of reviewing process and potential provenance. Such workflows will eliminate the "black box" nature by offering insights into the entire process used to build artifacts. This will support the research community in carrying out thorough studies that will enable them to satisfy those essential requirements for building a transparent and explainable AI/ML application. Documenting provenance to support published research should be considered a best practice rather than an afterthought. The community should be encouraged to follow well-established and consensus best practices for workflow design and software environment deployment. The aim of Geoweaver and the ***geoweaver_cwl*** package is to promote the efforts in that direction.

In order to improve the efficiency of the developed ***geoweaver_cwl*** package, our plan is to continue using Geoweaver and the package with more AI research projects. So far, we have only tested our package on definite workflows created by Geoweaver, and we believe further analyses are necessary to validate the broad utility of the package. For our future work, we would like to collaborate with a diverse research team from different domains and collect complex use cases from them. Testing different use cases will confirm additional details and novel functions and also ensure that our package satisfies the end-user requirement.

4.5 Conclusion

In this chapter, we first introduce Geoweaver and present a wrapper tool that overcomes current challenges of achieving repeatability, reproducibility, and reusability of workflows. To assess the outcome, we tested ***geoweaver_cwl*** with multiple use cases provided by Geoweaver. The study demonstrates that the ***geoweaver_cwl*** package can bring great benefits to the earth science community. The code is publicly available on GitHub (https://github.com/amrutakale08/geoweaver_cwl) and is currently used for translating Geoweaver AI/ML workflows. We encourage the AI community to participate in the adoption of Geoweaver by integrating the ***geoweaver_cwl*** package into their projects and to address any issues whenever possible to facilitate the development of new functionality in future versions.

Code Availability:

The ***geoweaver_cwl*** Python package is made open access at: https://pypi.org/project/geoweaver-cwl/0.0.1/. The source code of the package is accessible at: https://github.com/amrutakale08/geoweaver_cwl and exemplar results are accessible at: https://github.com/amrutakale08/geoweaver_cwl-usecases. The source code of the Geoweaver platform is accessible at: https://github.com/ESIPFed/Geoweaver.

Chapter 5 Utility of the Python package Geoweaver_cwl for improving workflow reusability: An illustration with multidisciplinary use cases

5.1 Introduction

Scientific workflow management systems (WfMS) like Kepler (Altintas et al., 2004), VisTrails (Callahan et al., 2006), Apache Oozie (Apache Oozie, 2012), Apache Taverna (Apache Taverna, 2014), Apache Airflow (Apache Airflow, 2015), and Geoweaver (Sun et al, 2020) have become increasingly popular. Not only do they support the automation of repetitive tasks, but also capture complex details at various levels and systematically record provenance information for the derived data products (Gil et al., 2007, Kale et al., 2022). WfMS have emerged as an alternative to ad-hoc approaches for constructing data-intensive machine learning (ML) experiments and provenance tracking. In general, a WfMS can be thought of as a program that consists of a set of modules connected by data flow, where each module can take input data from previous modules, parameter settings, and data from external sources. The visual representation can be considered as a graph, where each node represents the modules and edges represent the flow of data between them. Once the processes are linked together, the WfMS enables users to execute workflows automatically and monitor the progress in real-time.

The growing popularity of workflows has also raised numerous concerns about reproducibility and portability among the scientific community. Ad-hoc methods of data exploration (e.g., Perl scripts) have been widely used in the scientific community but also have significant limitations. This could hamper the collaboration between several researchers unless we standardized computational reusability and portability of workflows. To enable reusability and interoperability, the WfMS communities, for example, the Organization for the Advancement of Structured Information Standard (OASIS) (OASIS, 1998), Workflow Management Coalition (WfMC) (WfMC, 2001), Kepler (Altintas et al., 2004), Galaxy (Goecks et al., 2010), and World Wide Web Consortium (W3C) (Missier et al., 2013) have proposed a series of workflow languages that describe and record these workflow links and the involved processes. The standard languages commonly used in the industrial sector include BPEL (Business Process Execution Language) (Akram et al., 2006), BPMN (Business Process Model and Notation) (Chinosi and Trombetta, 2012), and Common Workflow Language (CWL) (Crusoe et al., 2021). For scientific workflows, most WfMS define their own languages, such as Taverna SCUFL2 (Simple Conceptual Unified Flow Language), and YAWL (Yet Another Workflow Language). These workflow languages offer information about the models and describe the process in a portable and reusable manner. Despite the fact that there are numerous WfMS being developed in the community, only a handful of them use the standard languages to describe their workflows. Whereas other WfMS have their unique syntax or approach for describing workflows and infrastructure needs. This approach might restrict computational portability and reusability. As a result, the majority of workflows created cannot be shared among different WfMS. Therefore, choosing the WfMS should be exercised with attention because the process of transitioning the workflow could be complicated and time-consuming especially when qualifying reproducibility. Table 5.1 highlights the different workflow automation software representing the workflow language and the important features they support.

In the below Table 5.1, we demonstrated different workflow automation software and its feature to make scientific workflows reproducible, portable, and provenance-enabled. However, in this dissertation, we would like to draw attention to Geoweaver a WfMS that helps scientists to sort AI experiments (create, manage, execute, share and record) and improves automation and reproducibility of workflows (Sun et al., 2020). Geoweaver is a simple-to-learn and adaptable application that can be used by anyone having prior experience with python scripting. The accessibility barrier is minimal for reproducing the existing workflows and downloaded workflows can be carried out independently without the need for software installation. Geoweaver has the capability to automate the workflow, record provenance, and export history without worrying about the technical debts and potential loss of their experiment's history and source code. To ensure interoperability of the designed workflows, we went one step further and automatically translated Geoweaver AI/ML workflows into CWL by our very first python package ***geoweaver_cwl*** (Kale et al., 2022 Preprint). We firmly believe that employing a CWL standard can offer a great solution for describing portable, flexible, and reusable workflows while also reducing the software engineering burden accompanying large-scale data analysis.

Table 5.1: Different workflow automation software highlighting the important features

Tool	**Process**	**Workflow Language**	**Open Source**	**Features**
Kepler	Web services Unix commands Shell script	XML (Extensible Markup Language)	✓	Allows user to reuse data, workflow, and components Freely available under BSD (Berkeley Source Distribution) license
Apache Airflow	Bash Python	DAGs (Directed Acyclical Graphs)	✓	Highly scalable Allows user to monitor and manage task easily
Apache Taverna	Local and remote servers RESTful services Shell script R processor	SCUFL2 (Simple Conceptual Unified Flow Language)	✓	Wide range of services and extensible architecture Workflow provenance Secure
Apache Oozie	Java Hadoop jobs (MapReduce, Pig, Hive, Sqoop) Shell scripts	DAGs (Directed Acyclical Graphs)	***X***	Scalable Reliable Extensible Integrated
VisTrails	Python Local and remote servers Web services	XML (Extensible Markup Language)	✓	Flexible Provenance architecture Support collaborative exploration and visualization
Geoweaver	Python Shell script Local and remote servers Secure Shell	CWL (Common Workflow Language)	✓	Hybrid Workflow Full access to Remote files Process-oriented provenance Code machine separation Hidden data flow

Our objective is to highlight the importance of scientific workflows and encourage the research community to adopt the CWL standards. In this chapter, we first provide the brief methodology of a provenance-enabled platform named Geoweaver and the python package ***"geoweaver_cwl"*** which transforms Geoweaver AI workflows into CWL. Second, we present several use cases from different domains to test the useability of the derived package. Third, we demonstrate results from the use cases and underline that CWL standards can assist in overcoming major challenges when sharing workflows between institutions and users. Finally, we conclude with some future research directions.

5.2 Methodology of Geoweaver and Geoweaver_cwl

Reproducible analyses require sharing data, methodology, and computational algorithms (Peng, 2011). In recent years, methods for organizing big data analysis through computational workflow and workflow description language have become increasingly popular to enable reproducibility and interoperability in the earth science domain (Kale and Ma, 2022 In press). In this chapter, we work on Geoweaver (Sun et al., 2020) as a WfMS tool for researchers and students to improve their research productivity and workflow FAIRness. It is an in-browser software application that allows researchers to create and execute full-stack data processing workflows by utilizing online spatial data resources, high-performance computing environments, and open-source deep learning frameworks. Geoweaver offers a comprehensive solution that includes server management, a code repository, workflow orchestration tools, and a history logger (Sun et al., 2020). We consider Geoweaver as an ideal WfMS which captures crucial provenance data that can reliably trace the history of analytical results. In our previous chapter, we demonstrated how to successfully create and execute a workflow in Geoweaver with a simple example (Kale and Ma, 2022 In press). One of the major benefits of Geoweaver is that it does not take long to understand, and users can run the Geoweaver workflow package without having Geoweaver installed. Figure 5.1 demonstrates the workflow shared by user A to user B and how the workflows created in Geoweaver are shareable, reproducible, and standardized in CWL format.

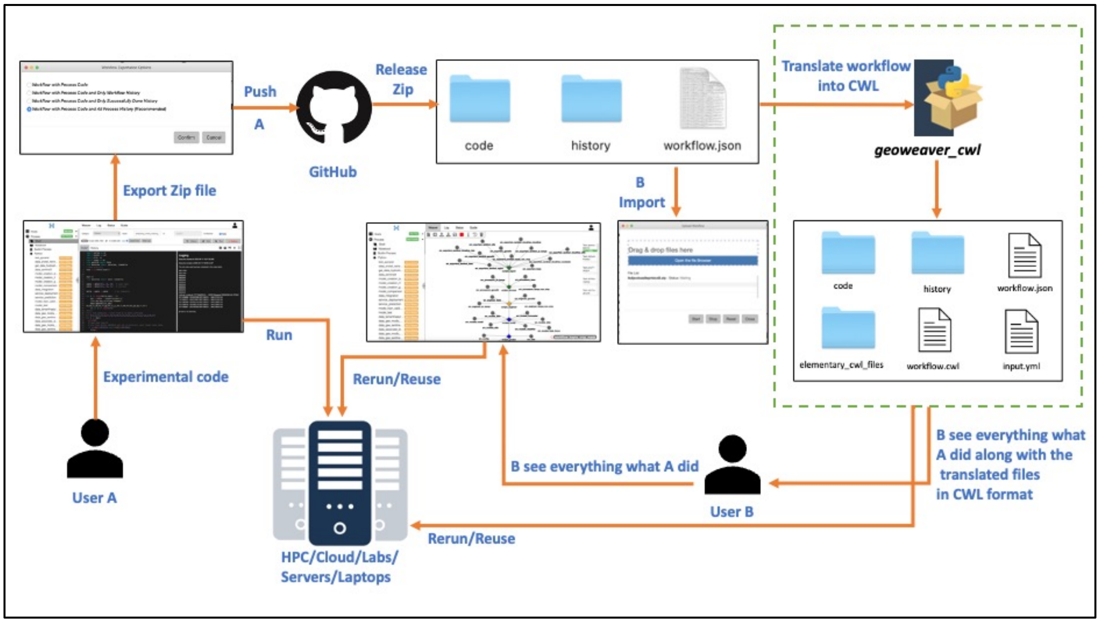

Figure 5.1: Working structure of Geoweaver and translation of AI/ML workflows shared between user A and user B to enable portability and reproducibility of workflows

The framework of Geoweaver is based on three core modules host, process, and workflow. The host module is the foundation of the entire framework which opens the entry to existing resources like servers, virtual machines, Jupyter servers, and third-party computing platforms like Google Earth Engine and Google Colab. The process module is widely used to write scripts, programs, commands, or code for the current AI/ML experiments (refer to experimental code in figure 5.1). As most of the AI/ML experiments commonly employ python, the process module supports Python, Jupyter Notebook, Shell Scripts, and Deep Learning libraries like Keras, TensorFlow, and PyTorch. Provenance support in scientific workflows is paramount, and Geoweaver is ideally positioned to record critical provenance information that can document the lineage of analytical results. The process monitor tracks all the events during the execution and stores all the execution results, inputs, and outputs in the database. The provenance manager is responsible for evaluating the quality of the stored history or fixing the process execution from failures. The workflow module is used to compose workflow by connecting multiple processes into the graphical workflow system. The workflow module also provides real-time status of each executed process indicating a progress bar with a different color. To make the workflow knowledgeable and shareable Geoweaver allows users to import and export (upload and download) the created workflow into a simple zip file with the intent that users can directly start working on the existing files. The zip file contains a code folder, history folder, and workflow.json file. The code folder contains all the experimental code written by a user, the history folder contains the history of all the executed processes in the workflow, and the workflow.json file contains the structured information of nodes and edges that form a workflow. Once the workflow is exported, a zip file can be shared on GitHub or another sharing medium in order for others to reuse the existing work. Any user who wishes to replicate or reproduce the existing work can import the zip file in Geoweaver or can also run the workflow package without having Geoweaver installed.

Despite the fact that workflow systems are popular prior to CWL standards, very few WfMS are compatible with each other. This implies that users who do not follow CWL standards must express their computational workflow differently every time they adopt a new workflow system, resulting in local success but global non-portability. This is due to the lack of standards or practices which makes it difficult for effective collaboration on computational methodologies. To overcome this challenge, we designed ***geoweaver_cwl*** a python package that automatically describes Geoweaver AI/ML workflows into CWL (Kale et al., 2022 Preprint). Describing these workflows into CWL will provide a structured and standard approach when sharing information or reproducing existing work. Once the zip file is exported from GitHub, users can install the ***geoweaver_cwl*** package and easily translate their workflows into CWL scripts. The ***geoweaver_cwl*** has two core functions generate_cwl and generate_yml that enables the easy translations of workflow into CWL files. The Generate_cwl function takes the input workflow.json file and describes the workflow into the workflow.cwl file. Additionally, the function also creates a subdirectory named elementary_cwl_files where all the processes involved in creating a workflow will be translated into CWL files. The generate_yml function produces a Yet Another Markup Language (YAML) file, which writes the input to run the workflow.cwl file. The declarative approach to describe the workflow into CWL scripts keeps provenance organized by documenting the inputs, outputs, workflow steps, and latest version. In addition to Geoweaver, the CWL result can be used on a variety of computer platforms, giving users more opportunities to run the workflow without sacrificing provenance or needing to redo the workflow if they want to use another WfMS. Geoweaver is the unique combination of hybrid workflows, remote access to files, hidden data flow, code-machine separation, well-documented provenance, and standardized CWL-complaint workflows which makes the user experience complete. Using Geoweaver is a long-term investment and will make every scientist's work preserved and be understandable even years after the original experiments.

5.3 Use Cases, Results and Evaluation

We have previously tested ***geoweaver_cwl*** with simple and complex workflows provided by Geoweaver. However, to verify the utility of our package, we decided to validate with more use cases from different domains. Below is a list of five different uses cases we tested on our package.

Use case 1: CMAQ-predict-geoweaver

The scientific topic of this workflow is to monitor and predict air quality index in California, that integrates the conventional air quality model, the Community Multi-scale Air Quality Model (CMAQ), and AI models. This workflow was created by Geoweaver and publicly available on GitHub (https://github.com/earth-artificial-intelligence/cmaq-predict-geoweaver). The GitHub repository contains the code folder, history folder, and workflow.json file.

We installed the ***geoweaver_cwl*** package and used the generate_cwl and generate_yml functions to translate this workflow into CWL. We obtained the files "input.yml", "workflow.cwl", and "elementary cwl files folder", which included the cwl files used in creating the workflow. The workflow translation process was fast and easy, and we also noticed that using cwltool speeds up workflow execution compared to the original procedure in Geoweaver. The translated code for this use case is also made available on GitHub: https://github.com/amrutakale08/geoweaver_cwl-usecases/tree/main/cmaq-predict-geoweaver-master.

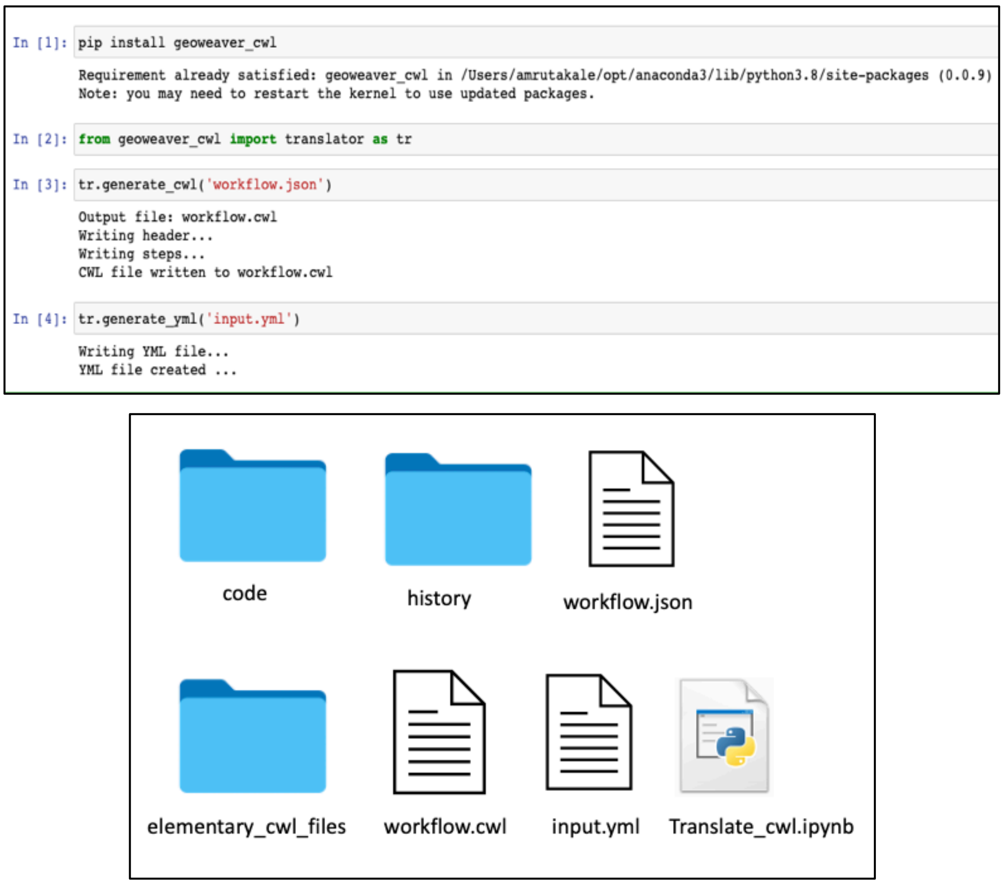

Figure 5.2: Installation of geoweaver_cwl package with the functions generate_cwl and generate_yml

(Installation of ***geoweaver_cwl*** package with the functions generate_cwl and generate_yml to successfully translate the Geoweaver workflows into CWL (top). Translated CWL files which consist of elementary_cwl_files folder, workflow.cwl, input.yml and the translate_cwl.ipynb file (python file where we have written the code to install the package and try to run functions generate_cwl and generate_yml) (bottom)).

Once the workflows are described into CWL scripts they can be executed on variety of software environments like cwltool, Arvados, Toil, CWL-Airflow, and more. Additionally, there are several cutting-edge applications like Rabix that can speed up the procedure. Rabix is an open-source desktop application that allows researchers to create and edit CWL documents (Kaushik et al., 2017; Amstutz et.al, 2021). In our previous chapter we used cwltool to execute the CWL scripts generated from the python package (Kale et al., Preprint). We invoke cwl_runner with workflow.cwl and input object input.yml on the command line.

```
cwl-runner workflow.cwl input.yml
```

The command is intended to provide the same outcomes as workflow files in Geoweaver by triggering all the CWL and YAML files' internal functions in the same order. The benefit of CWL offers a way to describe reusable and portable workflows. Researchers, students from any domain can easily share, exchange, edit, and reuse workflows by translating Geoweaver AI/ML workflows into CWL scripts using the ***geoweaver_cwl*** package. Another advantage of CWL is that applications written with CWL are portable and can be used in a multitude of environments, such as local or cloud infrastructures.

Use case 2: Emission-AI-geoweaver

This workflow is generated by Geoweaver to replicate the experiments done in Emission AI's published article, which is to build ML models to train on satellite observations (Sentinel 5), ground observed data (EPA eGRID), and meteorological observations (MERRA) data to directly predict the Nitrogen Dioxide (NO2) emission rate of coal-fired power plants. This workflow is publicly available on GitHub (https://github.com/earth-artificial-intelligence/emissionai-geoweaver). The GitHub repository contains the code folder, history folder, and workflow.json file. We followed similar steps as figure 5.2 to translate the workflow in CWL and successfully generated CWL files. The translated workflow code and files are available on GitHub: https://github.com/amrutakale08/geoweaver_cwl-usecases/tree/main/emissionai-geoweaver-main.

Use case 3: Eddy-detection-geoweaver

This workflow is generated by Geoweaver to replicate the experiment from the Jet Propulsion Laboratory (JPL) notebook. The workflow aims to train an ML model and use it to detect ocean eddies from remote sensing imagery automatically. This workflow is publicly available on GitHub (https://github.com/earth-artificial-intelligence/eddy_detection_geoweaver). The GitHub repository contains the code folder, history folder, and workflow.json file. We followed similar steps as figure 5.2 to translate the workflow in CWL and successfully generated cwl files. The translated workflow code and files are available on GitHub: https://github.com/amrutakale08/geoweaver_cwl-usecases/tree/main/eddy_detection_geoweaver-main.

Use case 4: Interdisciplinary use cases

The majority of the use cases we used in this dissertation and our previous chapter (Kale et al., Preprint) are acquired from Geoweaver. To confirm the portability and reliability of our package we tested a few other use cases from different domains. The first use case demonstrates the proposed framework of ML application in determining Multiple Sclerosis (MS) types and progression levels in MS patients. The second use case demonstrates the data preprocessing steps for the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data. Although these workflows are much simpler and smaller than the above use cases (1,2, and 3) we are certain that ***geoweaver_cwl*** successfully translates all types of workflows into CWL files. The workflows created in Geoweaver are publicly available on GitHub: https://github.com/amrutakale08/workflows. The GitHub repository contains the code folder, history folder, and workflow.json file for each workflow. We followed similar steps as figure 5.2 to translate the workflow in CWL and successfully generated CWL files. The translated workflow code and files are available on GitHub: https://github.com/amrutakale08/geoweaver_cwl-usecases/tree/main/MS-patients and https://github.com/amrutakale08/geoweaver_cwl-usecases/tree/main/ModisDataProcess .

Code Availability:

The ***geoweaver_cwl*** Python package is made open access at: https://pypi.org/project/geoweaver-cwl/0.0.1/. The source code of the package is accessible at: https://github.com/amrutakale08/geoweaver_cwl and exemplar results are accessible at: https://github.com/amrutakale08/geoweaver_cwl-usecases. The source code of the Geoweaver platform is accessible at: https://github.com/ESIPFed/Geoweaver.

5.4 Conclusion

The issue of standardizing computational workflows is becoming increasingly significant and has a prominent impact on the research community. To address this issue various domains in science, industry, and government have already transitioned to workflows, but initiative focusing on portability, scalability, and standardizing workflows still remains fragmented. In this dissertation, we call attention to this issue and provide a community-driven solution ***geoweaver_cwl***, which addresses the current struggles in attaining portability, reproducibility, and scalability of workflows. We evaluated ***geoweaver_cwl*** using various use cases from different domains. The study indicates that the ***geoweaver_ cwl*** package can greatly assist the students, researchers, and geoscience community in translating their AI/ML workflows into CWL-compliant WfMS software applications. In our future projects, we hope to maintain the current package and add new functionality as requested by the users. The package offers basic workflow translation of Geoweaver workflows into CWL scripts. In the future, we want to enhance our package by including more detailed workflow steps in order to provide more insight to the user. We encourage the research community to embrace WfMS and adopt the CWL standards in creating and sharing portable and complete workflow descriptions. The code for the python package is publicly available on GitHub (https://github.com/amrutakale08/geoweaver_cwl) and the translations of use cases into CWL is available on GitHub (https://github.com/amrutakale08/geoweaver_cwl-usecases).

Chapter 6 Conclusion, Limitations and Future Work

6.1 A Reflection on the Dissertation Objective and Research Topics

With the increasing adoption of AI/ML systems, there is an increasing need for the results to be interpretable, reproducible, traceable, and explainable. Although the post-hoc explainability approaches could be one way to explain a black-box model, we believe they are still in their infancy stage and not completely reliable. We suggest adopting established methods from the field of data and software provenance will be an ideal solution for providing explanations to AI/ML systems. Provenance will not only help users to trace, evaluate, understand, and reproduce the AI/ML results but will also enhance users' decisions about how much trust to place in data and results generated from the original sources. Our objective with this PhD research was to analyze the inter-relationship between provenance, XAI, and TAI, build a software package to document provenance and extend reproducibility of AI/ML workflows, and test the package in real-world applications to support XAI and TAI has been achieved successfully. We encourage the research community to participate in the adoption of Geoweaver by integrating the ***geoweaver_cwl*** package into their projects in order to enable reproducibility and well documented provenance.

There is also a need for more platforms like Geoweaver which automate the AI/ML workflows and enable users to perform all tasks in one place more efficiently. At the same time, we encourage the research community to adopt CWL standards when describing their AI/ML workflows to enable reusability. In order to facilitate the reproduction, reuse, and replication of computational methods, it is crucial to provide a complete and comprehensive description that encompasses what computer applications were utilized, how exactly they were used, and how they were connected to each other. This goal is prominently achieved by ***geoweaver_cwl.*** The package enables portability by being explicit about inputs/outputs to form the data flow, data locations, and execution models. Translating Geoweaver AI/ML workflows has not only supported users in getting detailed provenance information but has also offered more opportunities for students and researchers to reuse and reproduce the existing work. Furthermore, CWL supports using software container technologies, such as Docker and Singularity, to enable portability and delivers open standards, open-source code, and an open community. We expect to see more WfMS adopting CWL standards describing their workflows in order to have effective collaboration and reproducible research. We hope this research highlights the importance of provenance in AI-based systems and encourages AI practitioners/researchers to start documenting provenance. We expect to see more AI/ML models become explainable, providing enough details to the end-user, and we believe that provenance documentation will be one of the significant approaches to accomplish that.

6.2 Summary of Results and Their Inter-Relationship

Based on this Ph.D. research, we have successfully conducted literature reviews, built software packages, and implemented use case studies towards the overall objective. Below is a short summary of the activities and results following the layout of the research questions as stated in Chapter 1.

1. **How can provenance contribute to the explainability and transparency of AI/ML models to support the goals of XAI and TAI?**

In order to answer this question, we did a thorough literature review to understand how provenance, XAI, and TAI are interrelated with each other in Chapter 2. We also demonstrated the inter-relationship through bibliometric analysis which is simple but effective enough to demonstrate the research highlights for the past 10 years. Based on the analysis we comprehended that the key to establishing TAI is by documenting provenance in order to ensure that the results generated by AI-based systems are easily understandable and interpretable to humans. Explainability plays a crucial role in achieving trust and transparency in AI algorithms. Moreover, to improve explainability, data science practitioners have developed many approaches and strategic plans in order to support provenance documentation.

To increase transparency and explainability in AI-based systems, applying provenance documentation can be an essential technology to the existing XAI approaches. Provenance documentation is extremely important in the world of data and geospatial data production. Inclusion of provenance promises in increasing transparency as it can be used for many purposes, such as understanding how data were collected, determining ownership and rights, tracing steps in the data analysis process, and making judgments about resources to use.

2. **Would adopting domain-specific provenance standards be necessary, or can we rely on universal standards such as PROV-DM (Provenance Data Model) to document all the necessary complex details?**

In Chapters 2 and 3, we highlighted that despite the increasing attention and community effort dedicated to establishing standards for documenting provenance, the development of these standards for workflows involving AI/ML models is still in its early stages and has yet to reach full maturity. PROV-DM provides a standard data model for representing provenance information in a consistent and interoperable way, which defines a basic universal core data model for almost all kinds of the provenance information. However, there is uncertainty about using such a recommendation to record the provenance generated by AI/ML workflows that incorporate big geospatial data. Chapter 2 and 3 emphasize that there is need for domain-specific provenance standards. Adding domain-specific documentation standards will help the community to grasp essential details. It will help to capture the nuances and specific details of the data and processes involved in a particular domain, such as in the case of medical or environmental data. The decision to use a universal web provenance data model or a domain-specific provenance standard depends on the specific requirements and characteristics of the data and processes being documented and should be carefully evaluated on a case-by-case basis.

3. What software tools and WfMS are available for documenting provenance?

We presented a wide range of software tools, including Kepler, DataRobot Neptune, Datatron, Metaclip, and Geoweaver, in Chapter 3 and Chapter 5. These software tools are designed to promote reproducible outcomes and provenance tracking in scientific research. These tools provide several ways to explore the provenance repository by tracking model activity, recording all changes in the data and the model, and outlining best practices for data management and disposal. In Chapter 3, we also provided examples for the three tools Kepler, MetaClip, and Geoweaver to help readers better comprehend the various features that each tool has to offer. Following our analysis of these tools, we find that Geoweaver offers excellent capabilities for users to create, edit, and reuse AI/ML workflows all in one place.

4. What sets Geoweaver apart from other WfMS?

Geoweaver is a web-based workflow system that helps AI practitioners to integrate multiple steps such as preprocessing, training, testing, and post-processing into a single automated workflow. It provides great benefits to earth scientists and has the capability to run, modify, reproduce, share, track, and reuse AI workflows in a single or distributed environment. The strongest feature of Geoweaver is its capability to record the history of each process in order to assess data quality, recover from errors, and give users a complete sense of what time the file as open, what changes were made, by whom it was made, did the file run successfully or failed. Geoweaver offers great solutions to earth scientists in managing complex workflows when massive, big earth data is involved. The comprehensive demonstration and technical framework of Geoweaver is described in Chapter 3 and 4. We also compared Geoweaver with several other workflow automation software based on the number of features and workflow language they supports in Chapter 5, Table 5.1. We found out that none of WfMS supports standardizing workflow language in order to enable reproducibility and reusability except Geoweaver.

The current AI processing is done by a combination of software, scripts, libraries, and command-line tools. AI practitioners sometimes rely on their own proprietary ways to maintain their workflows, which are frequently developed without much consideration for accessibility or other people's ability to utilize or comprehend them. As many disparate and distributed entities are involved, it becomes a challenge to streamline all the processes to help scientists organize their deep learning projects into a manageable and organized manner. In the meanwhile, it is still challenging for the geoscientific community to exchange and reuse the developed AI workflows and findings, which leads to low efficiency in AI model training and deployment. Because scientists do not have access to the original AI processes, it is challenging for them to replicate the results due to the lack of information regarding workflow and platform conditions.

The majority of WfMS define their own languages for scientific processes, including Kepler, Taverna SCUFL2, and YAWL (Yet Another Workflow Language). These workflow languages provide abstractions and information models for processes. WfMS uses these abstractions and models to execute the corresponding workflows. Nevertheless, even though hundreds of WfMS have been created, only a handful of them implement the standard workflow language like CWL in order to enable reproducibility and reusability of workflow and one of them is Geoweaver.

5. What are the long-term benefits of standardizing workflows using CWL?

Chapter 4 demonstrates the working technical development of our python package ***geoweaver_cwl*** which translates Geoweaver AI/ML workflows into CWL scripts. CWL is a community standard to describe command-line-based workflows. It provides a standard but condensed set of generalizations procedures that are easily understood by every user. The language’s declarative format enables users to describe the process of executing diverse software tools and workflows through their command-line interface. Previously, to link the command-line tools researchers need to write shell scripts. Although these scripts offer an efficient approach to accessing the tools, writing, and maintaining them requires specialized knowledge. CWL aims to reduce the barrier to researchers using these technologies by providing a standard to unify them. In order to ensure reproducibility, CWL standards explicitly support the usage of container technologies. These standards ensure portability, so the same workflow can be executed in both local and high-performance cloud environments.

6.3 Scientific Contribution of this Ph.D. Research

- With this research we have proven that provenance documentation is a functional approach towards explainable and trustworthy AI.
- The translation of Geoweaver AI/ML workflows into the Common Workflow Language (CWL) scripts have further improved the interoperability of the documented provenance and provides flexibility to users to run their AI/ML workflows on local or cloud environment.
- The real-world use cases built from the TickBase and Geoweaver projects has demonstrated the scalability and portability of our python package.

6.4 Limitations

Although our Python package provides useful functions and features, there are several limitations that should be noted.

Performance: Sometimes the package may require significant computational resources and could experience slow run times when working with large datasets and complex workflows.

Platform Compatibility: Our Python package is primarily designed to translate Geoweaver AI/ML workflows into CWL scripts and only support python programs for now and does not support R or other programming languages. In order to work on Geoweaver platform you need to translate your code into python and then creates workflows in Geoweaver .

Documentation: While we have made every effort to provide comprehensive and up-to-date documentation, some areas of the package may be poorly documented or difficult to understand for users without prior experience in the field.

Dependencies: Our Python package only translates workflow generated by Geoweaver WfMS, which may cause issues with version conflicts or installation errors, particularly for users who are not using Geoweaver .

6.5 Future Work/Recommendation

The issue of standardizing computational workflows is becoming increasingly significant and has a prominent impact on the research community. To address this issue various domains in science, industry, and government have already transitioned to workflows, but initiative focusing on portability, scalability, and standardizing workflows still remains fragmented. In this research, we call attention to this issue and provide a community-driven solution ***geoweaver_cwl***, which addresses the current struggles in attaining portability, reproducibility, and scalability of workflows. We evaluated ***geoweaver_cwl*** using various use cases from different domains. The study indicates that the ***geoweaver_ cwl*** package can greatly assist the students, researchers, and geoscience community in translating their AI/ML workflows into CWL-compliant WfMS software applications. In our future projects, we hope to maintain the current package and add new functionality as requested by the users. The package offers basic workflow translation of Geoweaver workflows into CWL scripts. In the future, we want to enhance our package by including more detailed workflow steps in order to provide more insight to the user. We encourage the research community to embrace WfMS and adopt the CWL standards in creating and sharing portable and complete workflow descriptions. The code for the python package is publicly available on GitHub: https://github.com/amrutakale08/geoweaver_cwl and the translations of use cases into CWL is available at: https://github.com/amrutakale08/geoweaver_cwl-usecases.

List of Publications, Presentation and Awards

Publications

1. Utility of the Python package Geoweaver_cwl for improving workflow reusability: An illustration with multidisciplinary use cases. This chapter will be submitted to the Journal Earth Science Informatics.
2. Geoweaver_cwl: Transforming Geoweaver AI Workflows to Common Workflow Language to Extend Interoperability. SSRN Electronic Journal, Preprint 2022.

   Amruta Kale, Ziheng Sun, Chao Fan, Xiaogang Ma
3. Provenance in Earth AI. Elsevier Book Chapter

   Amruta Kale, Xiaogang Ma
4. Provenance documentation to enable explainable and trustworthy AI: A literature review. Data Intelligence journal, MIT Press, 2022. In press.

   Amruta Kale, Tin Nguyen, Frederick Harris, Chenhao Li, Jiyin Zhang, Xiaogang Ma
5. A knowledge graph and service for regional geological time

   Chao Ma, Amruta Kale, Jiyin Zhang, Xiaogang Ma

Presentations

1. Provenance documentation to enable explainable AI and trustworthy AI, National Science Foundation Annual Meeting, Coeur D Alene, ID, March 2022.
2. Geoweaver_cwl: A tool wrapper to translate Geoweaver AI workflows into Common Workflow Language. Earth Science Information Partners Summer Meeting, Pittsburg, PA, July 2022.
3. Semantics of FAIR geoscience data: A key factor to facilitate the data science workflow. GSA Annual Meeting, Denver, CO, October 2022

Xiaogang Ma, Amruta Kale, Chenhao Li, Que Xiang, Sanaz Salati, Anirudh Prabhu, Robert Hazen

4. Geoweaver_cwl: A python package for reproducible and interoperable workflows. American Geological Union Fall Meeting, Chicago, IL, December 2022.
5. Knowledge graphs for global and regional geologic time scales and an associated R package. AGU Fall Meeting 2020.

   Chao Ma, Xiaogang Ma, Amruta Kale
6. Approaches to improve semantic description and reasoning capability in the deep time knowledge base. GSA 2020 Annual Meeting.

   Xiaogang Ma, Chao Ma, Amruta Kale

Awards

1. Awarded a $3000 mini research grant for pitching a research plan at the FUNding Friday Poster Competition organized by Earth Science Information Partners (ESIP) in Summer Meeting 2022, Pittsburg, PA.
2. Early Career NSF Earth Cube travel award by American Geological Union (AGU) to present the research at the AGU Fall 2022 Meeting, in Chicago, IL.
3. Graduate Program and Student Association (GPSA) travel award to attend the Grace hopper celebration 2022 world's biggest women's conference in technology, Orlando, FL.

Code and Dataset Availability

1. Python package ***geoweaver_cwl*** is made open access at: https://pypi.org/project/geoweaver-cwl/0.0.1/
2. The source code of the package is accessible at the GitHub repository: https://github.com/amrutakale08/geoweaver_cwl
3. Exemplar results of the tested use cases are accessible at the GitHub repository: https://github.com/amrutakale08/geoweaver_cwl-usecases
4. New workflows created in Geoweaver are accessible at the GitHub repository: https://github.com/amrutakale08/workflows
5. The source code of the Geoweaver platform is accessible at: https://github.com/ESIPFed/Geoweaver

References

Adadi, A. and Berrada, M., 2018. Peeking inside the black box: a survey on explainable artificial intelligence (XAI). IEEE Access, 6, pp.52138-52160.

Agu, N.N., Keshan, N., Chari, S., Seneviratne, O., McCusker, J.P. and McGuinness, D.L., 2019. G-PROV: Provenance Management for Clinical Practice Guidelines. In: Proceedings of the Semantic Web Solutions for Large-scale Biomedical Data Analytics Workshop at the 2019 International Semantic Web Conference, Auckland, New Zealand, pp. 68-75.

AI HLEG (High-Level Expert Group on AI), 2019. Ethics guidelines for trustworthy AI. European Commission, Brussels, 39pp. Available: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai. Accessed on: September 21, 2021.

Akram, A., Meredith, D. and Allan, R., 2006, May. Evaluation of BPEL to scientific workflows. In Sixth IEEE International Symposium on Cluster Computing and the Grid,Vol. 1, pp. 269-274.

Alahmari, S.S., Goldgof, D.B., Mouton, P.R. and Hall, L.O., 2020. Challenges for the Repeatability of Deep Learning Models. IEEE Access, 8, pp.211860-211868.

Altintas, I., Berkley, C., Jaeger, E., Jones, M., Ludascher, B. and Mock, S., 2004, June. Kepler : an extensible system for design and execution of scientific workflows. In: Proceedings of the 16th International Conference on Scientific and Statistical Database Management. Santorini, Greece. pp.423-424.

Amalina, F., Hashem, I.A.T., Azizul, Z.H., Fong, A.T., Firdaus, A., Imran, M. and Anuar, N.B., 2019. Blending Big Data Analytics: Review on Challenges and a Recent Study. IEEE Access, 8, pp.3629-3645.

Amstutz, P., Crusoe, M.R., Tijanić, N., Chapman, B., Chilton, J., Heuer, M., Kartashov, A., Leehr, D., Ménager, H., Nedeljkovich, M. and Scales, M., 2016. Common workflow language, V1.0. Figshare. Available: https://doi.org/10.6084/m9.figshare.3115156.v2.

Apache Software Foundation., 2015. Apache Airflow. Available on: https://airflow.apache.org. Accessed on: January 9, 2023.

Apache Software Foundation 2014. Apache Taverna. Available on: https://taverna.apache.org. Accessed on: January 9, 2023.

Apache Software Foundation., 2012. Apache Oozie. Available on: https://oozie.apache.org. Accessed on: January 9, 2023.

Aria, M. and Cuccurullo, C., 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics, 11(4), pp.959-975.

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R. and Chatila, R., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, pp.82-115.

Baca, M. (ed.) 2016. Introduction to Metadata. Getty Publications. Accessible at: https://www.getty.edu/publications/intrometadata/. Accessed on: 17 November 2022.

Bedia, J., San-Martín, D., Iturbide, M., Herrera, S., Manzanas, R. and Gutiérrez, J.M., 2019. The METACLIP Semantic Provenance Framework for Climate Products. Environmental Modelling & Software, 119, pp. 445-457.

Belle, V. and Papantonis, I., 2021. Principles and Practice of Explainable Machine Learning. Frontiers in Big Data, 4, 25pp. Available: http://dx.doi.org/10.3389/fdata.2021.688969.

Boudette, N.E., 2016. Autopilot Cited in Death of Chinese Tesla Driver. The New York Times, 14 September. https://www.nytimes.com/2016/09/15/business/fatal-tesla-crash-in-china-involved-autopilot-government-tv-says.html. Accessed on: September 21, 2021.

Branco, M. and Moreau, L., 2006. Enabling provenance on large scale e-science applications. In: Proceedings of the 2006 International Provenance and Annotation Workshop, Chicago, IL, USA, pp. 55-63.

Brinckman, A., Chard, K., Gaffney, N., Hategan, M., Jones, M.B., Kowalik, K., Kulasekaran, S., Ludäscher, B., Mecum, B.D., Nabrzyski, J. and Stodden, V., 2019. Computing environments for reproducibility: Capturing the "Whole Tale". Future Generation Computer Systems, 94, pp.854-867.

Buneman, P. and Tan, W.C., 2019. Data provenance: What next? ACM SIGMOD Record, 47(3), pp. 5-16.

Buneman, P., Cheney, J., Tan, W.C. and Vansummeren, S., 2008, June. Curated databases. In: Proceedings of the 27th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, Vancouver, Canada, pp. 1-12.

Callahan, S.P., Freire, J., Santos, E., Scheidegger, C.E., Silva, C.T. and Vo, H.T., 2006, June. VisTrails: visualization meets data management. In Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data, pp. 745-747.

Castelvecchi, D., 2016. Can we open the black box of AI? Nature News, 538(7623), pp. 20 - 23.

Chari, S., Gruen, D., Seneviratne, O. and Mcguinness, D., 2020. Foundations of Explainable Knowledge-Enabled Systems. Available at: https://arxiv.org/pdf/2003.07520.pdf

Chen, L., Cruz, A., Ramsey, S., Dickson, C.J., Duca, J.S., Hornak, V., Koes, D.R. and Kurtzman, T., 2019. Hidden bias in the DUD-E dataset leads to misleading performance of deep learning in structure-based virtual screening. PloS One, 14(8), e0220113.

Cheney, J., Chiticariu, L. and Tan, W.C., 2009. Provenance in Databases: Why, How, and Where. Now Publishers Inc. Hanover, MA, USA. 100p. Available: http://dx.doi.org/10.1561/9781601982339.

Chinosi, M. and Trombetta, A., 2012. BPMN: An introduction to the standard. Computer Standards & Interfaces, 34(1), pp. 124-134.

Cho, R., 2021. Artificial Intelligence a game changer for climate change and the environment. https://news.climate.columbia.edu/2018/06/05/artificial-intelligence-climate-environment/. Accessed on: December 29, 2021.

Crusoe, M.R., Abeln, S., Iosup, A., Amstutz, P., Chilton, J., Tijanic, N., Ménager, H., Soiland-Reyes, S. and Goble, and The CWL community 2022. Methods included: Standardizing computational reuse and portability with the Common Workflow Language. Communications of ACM, pp 54 – 65.

Dastin, J., 2018. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G. Accessed on: September 23, 2021.

Das, A. and Rad, P., 2020. Opportunities and challenges in explainable artificial intelligence (xai): A survey. 24pp. Available: https://arxiv.org/abs/2006.11371.

Das, P., Ivkin, N., Bansal, T., Rouesnel, L., Gautier, P., Karnin, Z., Dirac, L., Ramakrishnan, L., Perunicic, A., Shcherbatyi, I. and Wu, W., 2020, June. Amazon SageMaker Autopilot: a white box AutoML solution at scale. In Proceedings of the Fourth International Workshop on Data Management for End-to-End Machine Learning, pp.1-7.

DataRobot AI Cloud, 2012. https://www.datarobot.com/. Accessed on: January 18, 2022.

Datatron MLOps: Machine learning Operations, 2016. https://datatron.com/. Accessed on: January 18, 2022.

Davidson, S.B. and Freire, J., 2008. Provenance and scientific workflows: challenges and opportunities. In: Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, Vancouver, Canada, pp. 1345-1350.

Di, L., Yue, P., Ramapriyan, H.K. and King, R.L., 2013. Geoscience data provenance: An overview. IEEE Transactions on Geoscience and Remote Sensing, 51(11), pp.5065-5072.

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J., Weinberger, D., Weller, A., 2017. Accountability of AI under the law: The role of explanation. arXiv preprint. arXiv:1711.01134.

Downs, R. R., R. E. Duerr, D. J. Hills, and H. K. Ramapriyan., 2015. Data stewardship in the Earth sciences, D-Lib, 21(7/8). doi:10.1045/july2015-downs.

Duerr, R. E., R. R. Downs, C. Tilmes, B. Barkstrom, W. C. Lenhardt, J. Glassy, L. E. Bermudez, and P. Slaughter, 2011. On the utility of identification schemes for digital earth science data: an assessment and recommendation. Earth Science Informatics, pp.139–160, doi:10.1007/s12145-0.

Eisenman, I., Meier, W.N. and Norris, J.R., 2014. A spurious jump in the satellite record: has Antarctic sea ice expansion been overestimated? The Cryosphere, 8(4), pp.1289-1296.

Eker, J., Janneck, J.W., Lee, E.A., Liu, J., Liu, X., Ludvig, J., Neuendorffer, S., Sachs, S. and Xiong, Y., 2003. Taming heterogeneity-the Ptolemy approach. In: Proceedings of the IEEE, 91(1), pp.127-144.

ESIP Data Preservation and Stewardship Committee., 2019. Data Citation Guidelines for Earth Science Data, Version 2. ESIP. doi: https://doi.org/10.6084/m9.figshare.8441816.v1.

Floridi, L., 2019. Establishing the rules for building trustworthy AI. Nature Machine Intelligence, 1(6), pp. 261-262.

Frost, L., 2019. Explainable AI and other questions where provenance matters, IEEE IoT Newsletter. https://iot.ieee.org/newsletter/january-2019/explainable-ai-and-other-questions-where-provenance-matters. Accessed on: September 21, 2021.

Garfield, E., 1990. KeyWords Plus-ISI's breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette. Current Contents, 32, pp. 5-9.

Garfin, G., Jardine, A., Merideth, R., Black, M. and LeRoy, S. eds., 2013. Assessment of climate change in the southwest United States: a report prepared for the National Climate Assessment. Island Press/Center for Resource Economics, pp.1 –533.

Garfinkel, S., Matthews, J., Shapiro, S.S. and Smith, J.M., 2017. Toward algorithmic transparency and accountability. Communications of the ACM, 60(9), pp. 5–5.

Gil, Y., Deelman, E., Ellisman, M., Fahringer, T., Fox, G., Gannon, D., Goble, C., Livny, M., Moreau, L. and Myers, J., 2007. Examining the challenges of scientific workflows. Computer, 40(12), pp.24-32.

Gilpin, L.H., Bau, D., Yuan, B.Z., Bajwa, A., Specter, M., Kagal, L., 2018. Explaining explanations: An overview of interpretability of machine learning. In: Proceedings of the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), Turin, Italy, pp. 80-89.

Goecks, J., Nekrutenko, A. and Taylor, J., 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome biology, 11(8), pp.1-13.

Goddard, M., 2017. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. International Journal of Market Research, 59(6), pp. 703-705.

Goodfellow, I., Bengio, Y. and Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA, 800p.
Goodman, B. and Flaxman, S., 2017. European Union regulations on algorithmic decision-making and a "right to explanation". AI Magazine, 38(3), pp. 50-57.

Groth, P. and Moreau, L., 2013. An Overview of the PROV Family of Documents, W3C. https://www.w3.org/TR/prov-overview/. Accessed on: September 21, 2021.

Groth, P., Gil, Y., Cheney, J. and Miles, S., 2012. Requirements for provenance on the web. International Journal of Digital Curation, 7(1), pp. 39-56.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F. and Pedreschi, D., 2018. A survey of methods for explaining black box models. ACM Computing Surveys, 51(5), pp. 1-42.

Gunning, D. and Aha, D., 2019. DARPA's explainable artificial intelligence (XAI) program. AI Magazine, 40(2), pp. 44-58.

Gunning, D., 2019. DARPA's Explainable artificial intelligence (XAI) program. In: Proceedings of the 24th International Conference on Intellegent Users interface. Marina del Ray, CA, USA. pp. ii-ii.

Gutiérrez, J.M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Rössler, O., Wibig, J., Wilcke, R., Kotlarski, S. and San Martin, D., 2019. An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. International journal of climatology, 39(9), pp.3750-3785.

Holzinger, A., Biemann, C., Pattichis, C.S. and Kell, D.B., 2017. What do we need to build explainable AI systems for the medical domain? 28pp. Available: https://arxiv.org/abs/1712.09923.

Huynh, T.D. and Moreau, L., 2014. ProvStore: a public provenance repository. In: Proceedings of the 2014 International Provenance and Annotation Workshop, Cologne, Germany, pp. 275-277.

Huynh, T.D., 2020. Prov 2.0.0 Python Package. https://pypi.org/project/prov/. Accessed on: September 21, 2021.

Huynh, T.D., Ebden, M., Fischer, J., Roberts, S. and Moreau, L., 2018. Provenance Network Analytics. Data Mining and Knowledge Discovery, 32(3), pp. 708-735.

IBM Green Horizon Project, China. https://www.ibm.com/blogs/internet-of-things/air-pollution-green-initiatives/. Accessed on: December 29, 2021.

Intel Newsroom: Applying emerging technology to solve environmental challenges. https://newsroom.intel.com/editorials/intel-study-applying-emerging-technology-solve-environmental-challenges/#gs.8c0dbs. Accessed on: December 29, 2021.

Iturbide, M., Bedia, J., Herrera, S., Baño-Medina, J., Fernández, J., Frías, M.D., Manzanas, R., San-Martín, D., Cimadevilla, E., Cofiño, A.S. and Gutiérrez, J.M., 2019. The R-based climate4R open framework for reproducible climate data access and post-processing. Environmental Modelling & Software, 111, pp.42-54.

Jaigirdar, F.T., Rudolph, C. and Bain, C., 2019. Can I trust the data I see? A Physician's concern on medical data in IoT health architectures. In: Proceedings of the Australasian Computer Science Week Multiconference, Sydney, Australia. pp. 1-10.

Jaigirdar, F.T., Rudolph, C., Oliver, G., Watts, D. and Bain, C., 2020. What Information is Required for Explainable AI?: A Provenance-based Research Agenda and Future Challenges. In: Proceedings of the IEEE 6th International Conference on Collaboration and Internet Computing (CIC). Atlanta, GA, USA. pp. 177-183.

Jain, S., Luthra, M., Sharma, S. and Fatima, M., 2020. Trustworthiness of Artificial Intelligence. In: Proceedings of the 6th International Conference on Advanced Computing and Communication Systems (ICACCS). Coimbatore, India. pp. 907-912.

Jentzsch, S.F. and Hochgeschwender, N., 2019. Don't Forget Your Roots! Using Provenance Data for Transparent and Explainable Development of Machine Learning Models. In: Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering Workshop (ASEW). San Diego, CA, USA. pp. 37-40.

Kale, A., Nguyen, T., Harris Jr, F., Li, C., Zhang, J., Ma, X., 2022. Provenance documentation to enable explainable and trustworthy AI: A literature review. Data Intelligence, In Press. doi:10.1162/dint_a_00119.

Kale, A., Ma, X., TBD. Provenance in Earth AI. In: Sun, Z. (ed.) Earth Science Artificial Intelligence. Elsevier. In Press.

Kale, A., Sun, Z., Fan, C. and Ma, X., Geoweaver_Cwl: Transforming Geoweaver AI Workflows to Common Workflow Language to Extend Interoperability. Available at: http://dx.doi.org/10.2139/ssrn.4284586

Kaushik, G., Ivkovic, S., Simonovic, J., Tijanic, N., Davis-Dusenbery, B. and Kural, D., 2017. Rabix: an open-source workflow executor supporting recomputability and interoperability of workflow descriptions. In Pacific Symposium on Biocomputing 2017 (pp. 154-165).

Kiepuszewski, B., Barros, A.P., Van Der Aalst, W., Ter Hofstede, A., 2003. Workflow patterns. Distributed and Parallel Databases, 14(1), 5-51.

Kirkpatrick, K., 2016. Battling algorithmic bias. Communications of the ACM, 59(10), pp. 16–17.

Köhl, M.A., Baum, K., Langer, M., Oster, D., Speith, T. and Bohlender, D., 2019. Explainability as a non-functional requirement. In 2019 IEEE 27th International Requirements Engineering Conference (RE), pp. 363-368.

Kohwalter, T., Oliveira, T., Freire, J., Clua, E. and Murta, L., 2016. Prov viewer: A graph-based visualization tool for interactive exploration of provenance data. In: Proceedings of the International Provenance and Annotation Workshop. McLean, VA, USA. pp. 71-82.

Krening, S., Harrison, B., Feigh, K.M., Isbell, C.L., Riedl, M. and Thomaz, A., 2016. Learning from explanations using sentiment and advice in RL. IEEE Transactions on Cognitive and Developmental Systems, 9(1), pp.44-55.

Krieger, L., Nijzink, R., Thakur, G., Ramakrishnan, C., Roskar, R. and Schymanski, S., 2021. Repeatable and reproducible workflows using the RENKU open science platform. In: Proceedings of 2021 EGU General Assembly, Virtual Meeting. Abstract EGU21-7655.

Kumar, A., McCann, R., Naughton, J. and Patel, J.M., 2016. Model selection management systems: The next frontier of advanced analytics. ACM SIGMOD Record, 44(4), pp.17-22.

Lanter, D.P., 1991. Design of a Lineage-Based Meta-Data Base for GIS. Cartography and Geographic Information Systems, 18(4), pp.255–261.

Lanter, D.P., 1993. A Lineage Meta-Database Approach Toward Spatial Analytic Database Optimization. Cartography and Geographic Information Systems, 20(2), pp.112–121.

Lecue, F., 2020. On the role of knowledge graphs in explainable AI. Semantic Web, 11(1), pp.41-51.

Lebo, T., Sahoo, S., McGuinness, D., Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S., and Zhao, J., 2013. PROV-O: The PROV Ontology. (W3C Recommendation). World Wide Web Consortium. http://www.w3.org/TR/2013/REC-prov-o-20130430/. Accessed on: September 21, 2021.

Lipton, Z.C., 2018. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. Queue, 16(3), pp.31-57.

Liu, S., Wang, X., Liu, M. and Zhu, J., 2017. Towards better analysis of machine learning models: A visual analytics perspective. Visual Informatics, 1(1), pp. 48-56.

Lucero, C., Coronado, B., Hui, O. and Lange, D.S., 2018. Exploring explainable artificial intelligence and autonomy through provenance. In: Proceedings of the 2nd Workshop on Explainable Artificial Intellgence. Stockholm, Sweden. pp. 85 -89.

Lundberg, S.M. and Lee, S.I., 2017, December. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, pp. 4768-4777.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.I., 2020. From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), pp.56-67.

Ma, X., 2018. Metadata. In: Schintler, L.A., McNeely, C.L. (eds.) Encyclopedia of Big Data. Springer, Cham, Switzerland. 5pp. doi:10.1007/978-3-319-32001-4_135-1.

Ma, X., 2022. Metadata. In: Schintler, L.A., McNeely, C.L. (eds) Encyclopedia of Big Data. Springer, Cham. pp. 650–654. doi:10.1007/978-3-319-32010-6_135.

Ma, X., Beaulieu, S.E., Fu, L., Fox, P., Di Stefano, M., West, P., 2017. Documenting Provenance for Reproducible Marine Ecosystem Assessment in Open Science. In: Diviacoo, P., Glaves, H.M., Leadbetter, A. (eds.) Oceanographic and Marine Cross-Domain Data Management for Sustainable Development. IGI Global, Hershey, PA, USA. pp. 100-126.

Ma, X., Fox, P., Tilmes, C., Jacobs, K. and Waple, A., 2014a. Capturing provenance of global change information. Nature Climate Change, 4(6), pp.409-413.

Ma, X., Zheng, J.G., Goldstein, J., Zednik, S., Fu, L., Duggan, B., Aulenbach, S., West, P., Tilmes, C., Fox, P., 2014. Ontology engineering in provenance enablement for the National Climate Assessment. Environmental Modelling & Software 61, pp. 191-205.

Mahendran, A. and Vedaldi, A., 2015. Understanding deep image representations by inverting them. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. pp.5188-5196.

Mayernik, M.S., Callaghan, S., Leigh, R., Tedds, J. and Worley, S., 2015. Peer review of datasets: When, why, and how. Bulletin of the American Meteorological Society, 96(2), pp.191-201.

McCausland, P., 2019. Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk, NBC News, 9 November. https://www.nbcnews.com/tech/tech-news/self-driving-uber-car-hit-killed-woman-did-not-recognize-n1079281. Accessed on: September 21, 2021.

McFarland, M., 2018. Uber shuts down self-driving operations in Arizona. https://money.cnn.com/2018/05/23/technology/uber-arizona-self-driving/index.html. Accessed on: September 21, 2021.

MEAEF (Ministry of Economic Affairs and Employment of Finland), 2019. Leading the Way into the Era of Artificial Intelligence: Final Report of Finland's Artificial Intelligence Program 2019. Ministry of Economic Affairs and Employment of Finland. 133pp. Available: http://urn.fi/URN:ISBN:978-952-327-437-2.

Microsoft's Earth AI program: https://www.microsoft.com/en-us/ai/ai-for-earth-tech-resources#primaryR16. Accessed on: December 29, 2021

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S. and Dean, J., 2013. Distributed representations of words and phrases and their compositionality. In Advances in Neural Information Processing Systems, pp. 3111-3119.

Miles, S., Wong, S.C., Fang, W., Groth, P., Zauner, K.P. and Moreau, L., 2007. Provenance-based validation of e-science experiments. Journal of Web Semantics, 5(1), pp. 28-38.

Miller, T., 2019. Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, pp. 1-38.

Missier, P., Belhajjame, K. and Cheney, J., 2013. The W3C PROV family of specifications for modelling provenance metadata. In: Proceedings of the 16th International Conference on Extending Database Technology, Genoa, Italy. pp. 773-776.

Montavon, G., Lapuschkin, S., Binder, A., Samek, W. and Müller, K.R., 2017. Explaining nonlinear classification decisions with deep taylor decomposition. Pattern Recognition, 65, pp. 211-222.

Moreau, L. and Groth, P., 2013. Provenance: An introduction to PROV. Synthesis Lectures on the Semantic Web: Theory and Technology, 3(4), pp. 1-129.

Moreau, L. and Missier, P., 2013. PROV DM: The PROV Data Model, W3C. https://www.w3.org/TR/2013/REC-prov-dm-20130430/. Accessed on: September 21, 2021.

Moreau, L., Groth, P., Miles, S., Vazquez-Salceda, J., Ibbotson, J., Jiang, S., Munroe, S., Rana, O., Schreiber, A., Tan, V. and Varga, L., 2008. The provenance of electronic data. Communications of the ACM, 51(4), pp.52-58.

Moreau, L., Huynh, T.D. and Michaelides, D., 2014. An online validator for provenance: Algorithmic design, testing, and API. In: Proceedings of the International Conference on Fundamental Approaches to Software Engineering, Grenoble, France. pp. 291-305.

Moreau, L., Ludäscher, B., Altintas, I., Barga, R. S., Bowers, S., Callahan, S., Chin, G., Clifford, B., Cohen, S., Cohen-Boulakia, S., Davidson, S., Deelman, E., Digiampietri, L., Foster, I., Freire, J., Frew, J., Futrelle, J., Gibson, T., Gil, Y., Goble, C., Golbeck, J., Groth, P., Holland, D. A., Jiang, S., Kim, J., Koop, D., Krenek, A., McPhillips, T., Mehta, G., Miles, S., Metzger, D., Munroe, S., Myers, J., Plale, B., Podhorszki, N., Ratnakar, V., Santos, E., Scheidegger, C., Schuchardt, K., Seltzer, M., Simmhan, Y. L., Silva, C., Slaughter, P., Stephan, E., Stevens, R., Turi, D., Vo, H., Wilde, M., Zhao, J. and Zhao, Y., 2008. Special Issue: The First Provenance Challenge. Concurrency and Computation: Practice and Experience, 20(5), pp. 409–418.

Murdoch, W.J., Singh, C., Kumbier, K., Abbasi-Asl, R. and Yu, B., 2019. Interpretable machine learning: definitions, methods, and applications. pp. 1-11. Available: https://arxiv.org/abs/1901.04592.

NAS (National Academies of Sciences), 2018. The Frontiers of Machine Learning: 2017 Raymond and Beverly Sackler U.S.-U.K. Scientific Forum. The National Academies Press, Washington, DC, 32pp. DOI: 10.17226/25021.

Organization for the Advancement of Structured Information Standard (OASIS), 1998. Available on: https://www.oasis-open.org/standards . Accessed on: January 9, 2023.

Osoba, O.A. and Welser IV, W., 2017. An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence. RAND Corporation, Santa Monica, CA, USA, 34pp.

Peng, R.D., 2011. Reproducible research in computational science. Science, 334(6060), pp.1226-1227.
Pimentel, J.F., Freire, J., Braganholo, V. and Murta, L., 2016. Tracking and analyzing the evolution of provenance from scripts. In: Proceedings of the 6th International Provenance and Annotation Workshop, McLean, VA, USA. pp. 16-28.

Preece, A., Harborne, D., Braines, D., Tomsett, R. and Chakraborty, S., 2018. Stakeholders in Explainable AI. Available: https://arxiv.org/pdf/1810.00184.pdf.

Rakova, B., Chowdhury, R. and Yang, J., 2020. Assessing the intersection of organizational structure and FAT* efforts within industry: implications tutorial. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain. pp. 697-697.

Rasp, S., Pritchard, M.S. and Gentine, P., 2018. Deep learning to represent subgrid processes in climate models. In: Proceedings of the National Academy of Sciences, 115(39). Princton, NJ, USA. pp.9684-9689.

Ribeiro, M.T., Singh, S. and Guestrin, C., 2016. 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, San Francisco, CA, USA. pp. 1135-1144.

Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V. and Floridi, L., 2021. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. AI & Society, 36(1), pp. 59-77.

Ross, C. and Swetlitz, I., 2018. IBM's Watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. https://www.statnews.com/wp-content/uploads/2018/09/IBMs-Watson-recommended-unsafe-and-incorrect-cancer-treatments-STAT.pdf Accessed on: September 21, 2021.

Rudin, C. and Radin, J., 2019. Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From An Explainable AI Competition. Harvard Data Science Review, 1(2), doi:10.1162/99608f92.5a8a3a3d.

Rudin, C., 2018. Please stop explaining black box models for high stakes decisions. In: Proceedings of the 32nd Conference of Neural Information Processing Systems (NIPS), Workshop on Critiquing and Correcting Trends Machine Learning, Montreal, Canada, 20pp. Available: https://arxiv.org/abs/1811.10154.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), pp. 206-215.

Samuel, S., 2019. A provenance-based semantic approach to support understandability, reproducibility, and reuse of scientific experiments. Doctoral dissertation, Friedrich-Schiller-Universität Jena, 241pp.

Samuel, S., Löffler, F. and König-Ries, B., 2020. Machine learning pipelines: provenance, reproducibility and FAIR data principles. In: Proceedings of the 2021 International Provenance and Annotation Workshop, Charlotte, NC. Available at: https://arxiv.org/abs/2006.12117.

Sarpatwar, K., Vaculin, R., Min, H., Su, G., Heath, T., Ganapavarapu, G. and Dillenberger, D., 2019. Towards Enabling Trusted Artificial Intelligence via Blockchain. In: Calo, S., Bertino, E., Verma, D. (eds.) Policy-Based Autonomic Data Governance. Springer, Cham, pp. 137-153.

Schelter, S., Boese, J.H., Kirschnick, J., Klein, T. and Seufert, S., 2017. Automatically tracking metadata and provenance of machine learning experiments. In: Proceedings of Machine Learning Systems Workshop at the 31st Conference on Neural Information Processing Systems, Long Beach, CA, USA, pp. 27-29.

Shaw, J., Rudzicz, F., Jamieson, T. and Goldfarb, A., 2019. Artificial intelligence and the implementation challenge. Journal of Medical Internet Research, 21(7), e13659.

Showstack, R., 1998. NASA selects Earth science information partners. Eos, Transactions American Geophysical Union Transactions, 79(5), pp.58-58.

Simmhan, Y.L., Plale, B. and Gannon, D., 2005. A survey of data provenance in e-science. ACM SIGMOD Record, 34(3), pp. 31-36.

Singh, J., Cobbe, J. and Norval, C., 2018. Decision provenance: Harnessing Data Flow for Accountable Systems. IEEE Access, 7, pp.6562-6574.

Spinner, T., Schlegel, U., Schäfer, H. and El-Assady, M., 2019. explAIner: A visual analytics framework for interactive and explainable machine learning. IEEE Transactions on Visualization and Computer Graphics, 26(1), pp. 1064-1074.

Sun, Z., Di, L., Burgess, A., Tullis, J.A. and Magill, A.B., 2020. Geoweaver: Advanced cyberinfrastructure for managing hybrid geoscientific AI workflows. ISPRS International Journal of Geo-Information, 9(2), 119. doi: 10.3390/ijgi9020119.

Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mousavi, S.M., Wang, J., Lin, C., Cristea, N., Tong, D., Carande, W.H., Ma, X. and Rao, Y., 2022. A review of Earth Artificial Intelligence. Computers & Geosciences, 159,105034. doi:10.1016/j.cageo.2022.105034.

Sundararajan, M., Taly, A. and Yan, Q., 2017. Axiomatic Attribution for Deep Networks. In: Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, pp. 3319-3328.

Tan, S., Caruana, R., Hooker, G. and Lou, Y., 2017. Detecting bias in black-box models using transparent model distillation. In: Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (AIES '18), New Orleans, LA, USA, pp. 303–310.

Talia, D., 2013. Workflow systems for science: Concepts and tools. ISRN Software Engineering, 2013

Tennery, A. and Cherelus, G., 2016. Microsoft's AI Twitter bot goes dark after racist, sexist tweets, Reuters, 24 March. https://www.reuters.com/article/us-microsoft-twitter-bot/microsofts-ai-twitter-bot-goes-dark-after-racist-sexist-tweets-idUSKCN0WQ2LA. Accessed on: September 21, 2021.

Tenopir, C., Rice, N.M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R. and Sandusky, R.J., 2020. Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. PloS one, 15(3), p.e0229003. doi: https://doi.org/10.1371/journal.pone.0229003.

Thiebes, S., Lins, S. and Sunyaev, A., 2020. Trustworthy artificial intelligence. Electronic Markets, pp.447–464.

Tilmes, C., Fox, P., Ma, X., McGuinness, D., Privette, A.P., Smith, A., Waple, A., Zednik, S., Zheng, J., 2013. Provenance representation for the National Climate Assessment in the Global Change Information System. IEEE Transactions on Geoscience and Remote Sensing 51 (11), pp. 5160-5168.

Tseng, G., Kerner, H., Nakalembe, C., Becker-Reshef, I., 2020. Annual and in-season mapping of cropland at field scale with sparse labels. Tackling Climate Change with Machine Learning Workshop at NeurIPS ’20, Virtual Conference. Available at: https://www.climatechange.ai/papers/neurips2020/29. Accessed: 17 November 2022.

Van der Aalst, W.M., 1998. The application of Petri nets to workflow management. Journal of Circuits, Systems, and Computers, 8(01), 21-66.

Van Eck, N.J. and Waltman, L., 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2), pp. 523-538.

Van Lent, M., Fisher, W. and Mancuso, M., 2004. An explainable artificial intelligence system for small-unit tactical behavior. In: Proceedings of the national conference on artificial intelligence. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999. pp.900-907.

Vanschoren, J., Van Rijn, J.N., Bischl, B. and Torgo, L., 2014. OpenML: networked science in machine learning. ACM SIGKDD Explorations Newsletter, 15(2), pp. 49-60.

Vartak, M., Subramanyam, H., Lee, W.E., Viswanathan, S., Husnoo, S., Madden, S. and Zaharia, M., 2016, June. ModelDB: a system for machine learning model management. In: Proceedings of the Workshop on Human-In-the-Loop Data Analytics, San Francisco, CA, USA. pp. 1-3.

Vincent, J., 2016. Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day, The Everge, 24 March. https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist. Accessed on: September 21, 2021.

Werder, K. and Ramesh, B., 2021. Establishing Data Provenance for Responsible Artificial Intelligence Systems. ACM Transactions on Management Information Systems, In Press.

White House, 2020. Executive Order on Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. https://trumpwhitehouse.archives.gov/articles/promoting-use-trustworthy-artificial-intelligence-government/. Accessed on: September 23, 2021.

Wing, J.M., 2020. Ten Research Challenge Areas in Data Science. Harvard Data Science Review, 2(3), doi:10.1162/99608f92.c6577b1f.

Wing, J.M., 2021. Trustworthy AI. Communications of the ACM, 64(10), pp. 1-12.

Wohlin, C., 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In: Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering, Karlskrona, Sweden. pp. 1-10.

Workflow Management Coalition (WfMC), W.P.D.I.X., 2001. Process Definition Language (XPDL), WfMC Standards. WFMC-TC-1025.

Zeng, Y., Su, Z., Barmpadimos, I., Perrels, A., Poli, P., Boersma, K.F., Frey, A., Ma, X., de Bruin, K., Goosen, H. and John, V.O., 2019. Towards a traceable climate service: Assessment of quality and usability of essential climate variables. Remote sensing, 11(10), 1186. doi:10.3390/rs11101186.

Zhu, J., Liapis, A., Risi, S., Bidarra, R. and Youngblood, G.M., 2018. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. In 2018 IEEE Conference on Computational Intelligence and Games (CIG). Maastricht, Netherlands. pp.1-8.