Universiteit Leiden
The Netherlands

Bachelor Computer Science & Datascience and Artificial Intelligence

Finding and Visualising Patterns and Knowledge Gaps in the Field of AutoML for Earth Observation

Kenji Opdam

Supervisors:
Mitra Baratchi & Julia Wąsala

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

01/06/2025

Abstract

Automated Machine Learning (AutoML) for Earth Observation (EO) is rapidly evolving, making it difficult for researchers to grasp the current state of the field. This thesis addresses the challenges in understanding trends and knowledge gaps by developing intuitive visualisations of the patterns found in the substantive information, such as frameworks and datasets used, and tasks performed. We created a dataset by retrieving the metadata of papers on AutoML for EO using a scraper that searches for papers on Google Scholar and Crossref. Additionally, we compare with a curated list of papers from a survey paper [WMS+26]. Using these datasets, we created data visualisations that can be used by researchers to answer questions about trends in the field of AutoML for EO. Answering these questions helps researchers understand the state of the field and make decisions based on this understanding. Additionally, these visualisations can be a companion to a review paper by giving the readers visualisations alongside a textual explanation, which makes it easier to grasp the information. The visualisations highlight patterns, such as the prevalence of specific AutoML frameworks for various EO tasks, dominant publishers, and the most and least researched tasks, which help both seasoned researchers and newcomers in gaining a deep understanding of the field making it easy to decide what to work on and deciding what technologies to use.

Contents

1 Introduction

Automated Machine Learning (AutoML) focuses on automating the creation of Machine Learning (ML) models. The field of AutoML for Earth Observation (EO) applies AutoML to various EO tasks to make it easier for researchers with little knowledge about ML to use ML models for these tasks. AutoML for EO is a relatively new and quickly growing field with a great amount of research being done in the past few years on various ways to incorporate AutoML into EO tasks [WMA+24, GKDF23, PSBvRV21, ZZY+25, dSBB+22]. These tasks are tasks like object detection, land use classification, estimating grass height, etc., where a large amount of data from a remote sensor is analysed using ML [GKDF23, AMCC20, dSBB+22].

It can be difficult for people new to this field to get an overview and define new problem statements because of the large number of AutoML frameworks, EO tasks, and EO datasets. Visualisations are needed to summarise paper metadata to quickly and efficiently show patterns and trends in a format easily understandable to humans [Ebe23, vW05]. Multiple tools create visualisations based on the citation network, for instance, Litmaps [Lit], Connected Papers [Con], and ResearchRabbit [Res]. These tools assist in finding papers relevant to a given topic [KGS+22, KSM+22, BJK23], and make researching a field more efficient by helping with the difficult task of navigating scientific literature [KGS+22, KSM+22, BJK23]. However, these tools focus on paper metadata such as authors and keywords, and include little or no information about the substantive information, information about the contents, such as methods or datasets. This information is crucial to understand current developments in the field.

This thesis addresses the lack of visualisations about study details by finding patterns in AutoML4EO papers to detect knowledge gaps that can inspire future research in the field. We design simple visualisations to improve ease of interpretation. Wąsala et al. provide an overview of AutoML4EO using a survey [WMS+26]. The visualisations made in this thesis will serve as a companion to that survey by displaying the information in a new way. Visualisations can show certain trends clearly that are otherwise difficult to extract from a textual survey. Making the visualisations is done by first gathering papers to make a dataset that can be used to create the visualisations. Then the visualisation libraries Bokeh [Bok18] and networkx [HSSC08] are used to explore the dataset. Finally, D3.js [BOH11] is used to create the final visualisations and to showcase them on a site.

We present the following contributions:

1. We propose a workflow for identifying and extracting information from AutoML4EO papers, and make the code available at https://github.com/konjiii/scraper.

2. We collect a dataset of AutoML4EO papers and metadata, which can help beginning researchers in the field to define problem statements and select target venues.

3. We create a visualisation tool that shows trends and patterns in frameworks, datasets, and tasks.

1.1 Thesis overview

the rest of this thesis is structured as follows. In Section 2 we will talk about Earth Observation, Automated Machine Learning and Visualisations; Afterwards, we discuss how the scraper works and how the visualisations were made in Section 3; The next Section 4 contains the research questions and gives examples on how this product can be used; Further, In Section 5 we look at some visualisations and answer the research questions stated in Section 4; Next, we conclude the thesis in Section 6; And finally, Section 7 shows where the code and visualisations can be found.

2 Related Work

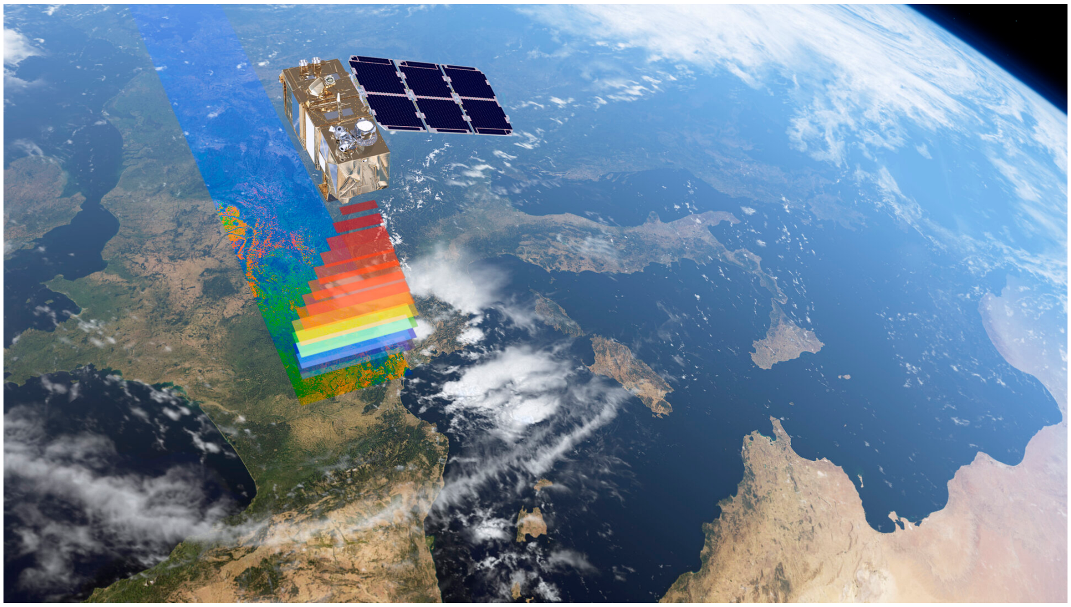

Figure 1: Source: www.esa.int.

This section discusses background on EO and ML, AutoML, and Visualisation. In the first subsection, we will cover EO and the use of ML in EO. The second subsection will discuss AutoML and how it is used in EO to ease the use of ML in EO tasks. Next, in the third subsection, we talk about visualisation. The final section covers the contributions of this thesis.

2.1 EO and ML

Earth Observation is important for monitoring Earth's climate systems, urban development, agriculture, and other applications [ARS+17, SW21, ZYD+22, PRKG20, PCS+18]. The volume of publicly available EO data has been increasing in recent years with the increase in satellite and other remote sensing technology [GQF20, LZL+18]. A reason for this is the increase in the number of satellites being launched, which is supported by the decrease in costs in launching satellites [LZL+18]. Additionally, the cost of data storage keeps decreasing, making it possible to capture large datasets [LZL+18]. It has become practically infeasible to manually analyse all available data due to the increase in data. Consequently, Scientists are increasingly using ML to process the raw data and perform tasks like image recognition or classification [TSD+25, LZL+18].

A trained ML model is able to very quickly parse through the data, making it possible to work with these big datasets. This makes ML an effective tool for analysis of remote sensing data [ZYD+22]. ML is applied to many different EO tasks like climate change, agriculture, and urban development problems as a result [RDK+22, PCS+18, FDMS19]. Anderson et al. showed the potential of using EO in achieving the Sustainable Development Goals (SDGs) by the UN, stating that EO introduces the ability to visualise data of e.g. temperature, ice extent, sea level, Land Cover change, etc. by providing consistent and detailed imagery of earths surface making decision making using this data easier [ARS+17]. EO has also been used for sustainable urban development [PRKG20]. Prakash et al. mention that medium and high resolution satellite imagery can be used to create mappings of the current land usage of a city at a very low cost compared to traditional survey methods [PRKG20]. Some other successful uses of EO include land use/cover classification [PCS+18] with an overall accuracy of 94% in land classification, mapping climate change-related disease [KBL+19], hydrology [MRA+17], and estimating tropical deforestation using EO [ASE+10].

2.2 AutoML for EO

One challenge with the use of ML models is the difficulty of making these models. Creating an ML model requires significant time and expertise. It is difficult for researchers with limited ML expertise to navigate the huge space of possible design choices. AutoML aims to address these problems. AutoML automates the creation of ML models by automating the tasks performed when creating an ML model that would normally be done manually [HKV19]. AutoML can make ML more accessible to domain experts, including in the field of EO. Furthermore, AutoML reduces the hands-on time designing ML models and can create ML models with results that are close and often better than state-of-the-art ML models that were created manually [PSBvRV21, GKDF23, TAR+19]. Salinas et al. make use of Auto-Keras [PSBvRV21], an AutoML framework that automates the process of building and tuning deep learning models, using techniques like neural architecture search (NAS), a subfield of AutoML, specifically targeting the automation of neural network architecture design [EMH19], to find the best model architecture and hyperparameters [JSH19]. Three models based on the Auto-Keras framework are compared against non-AutoML architectures on 7 datasets [PSBvRV21]. The results were promising, with at least one of the AutoML approaches outperforming the non-AutoML architectures on 5 out of the 7 datasets [PSBvRV21]. Another example of a successful AutoML implementation is [GKDF23], where Gudzius et al. used a MACU network, a CNN that is used for semantic segmentation, as a base. A MACU network needs extensive manual tuning when applied across different scenarios [GKDF23]. To overcome these limitations, the researchers perform NAS, which showed a better performance compared to a manually designed MACU, in addition to the NAS-MACU being faster in training itself compared to manually designing a network with limited training data [GKDF23].

2.3 Visualisation

Visualisation is defined as a visual representation of patterns found in some group of information designed to effectively communicate the content of the information and improve understanding [Ebe23] and is a widely used tool to communicate information to people [Ebe23, vW05]. Because of many different technologies like satellites or stock trading, the amount of produced information is increasing greatly [vW05]. Because of this, trying to get a grasp of all that information is not possible without the proper tools and visualisation is an effective tool to give insight into this large amount of data [vW05].

Visualisations can come in different forms. Quantitative graphs show trends or comparisons in numerical data. Qualitative diagrams illustrate non-numerical data, abstract visual metaphors, or artistic imagery: used to make certain concepts visual and memorable [Ebe23], and are often used to enhance comprehension and decision-making [Ebe23]. This is because visualisations can be easier to interpret when compared to textual explanations, even though visualisations are less precise than textual explanations. Smercnik et al. showed that visualised risk data requires less cognitive effort to interpret than the textual counterpart [SMK+10, Ebe23]. And Cassenti et al. data visualised in a scatterplot improved the accuracy of decision making by law enforcement compared to raw data [CRK19, Ebe23]. We created quantitative graphs, mostly in the form of bar plots, for this thesis.

2.3.1 Effects of visualisations

Visualisations can impact various aspects of judgement and decision-making, namely: Comprehension/decision accuracy, Decision speed and efficiency, Decision confidence, and Attitude change [Ebe23]:

Comprehension/decision accuracy: Numerous studies show that well-designed visualisations can improve the accuracy on comprehension, judgement, or decision [Ebe23].
Chandler et al. and Okan et al. show in experiments that correctly designed visualisations can greatly improve the comprehension of the information by the reader [Ca91, OSBdB18]. Okan et al. additionally show that foreground+background graphs and graphs incorporating simple numerical labels, increase the understanding of risk [OSBdB18]. "foreground+background" graphs are graphs that depict the number of people affected by a risk and those at risk of harm [OSBdB18] in contrast to "foreground-only" graphs that depict only the number of people affected by a risk.
Pfaff et al. and Semmler et al. show that visualisation tools can increase the comprehension when making decisions [PKD+13, Sa02]. Pfaff et al. uses Decision Space Visualisation (DSV), a method using visualisations to explore and understand the available decisions [PKD+13], to support Option Awareness (OA), which is when the individual is aware of the desirability of available options [PKD+13], and showed that it helped decision makers identify good decisions in a wide range of possible decisions [PKD+13]. Semmler and Brewer indicate that using a flow chart as a visualisation tool can increase jurors' comprehension accuracy of the judge's instructions compared to the judge's instructions delivered verbally [Sa02]. However, despite the improvements in comprehension, the accuracy of the understanding did not translate into a significant improvement in decision accuracy where their final judgement matched the expert's assessment [Sa02].

Decision Speed and Efficiency: Visualisations also help with an increase in the speed at which a decision is made [Ebe23].
Géryk and Moore showed that visualisations can improve the speed of information processing, which results in faster decision-making times [Gé17, Moo17]. Géryk used their analytics framework, which employed animated visualisations for educational data, which resulted in faster performance [Gé17].

Decision confidence: Many papers also show a difference in decision confidence when using visualisations [Ebe23].
Butavicius and Lee show that structured visualisation techniques give people increased confidence in decision making compared to an unstructured random list [BL07], however, the increase in confidence is less significant than the increase in accuracy of judgment and

decisions and efficiency [BL07]. Batavicius and Lee also show that 2D visualisations are more effective than 1D displays in increasing confidence [BL07].
Creating such visualisations should, however, be done very carefully so as not to include any biases, since biases are translated into the decision-making of the people using the visualisations. Andrade states that people tend to have overconfidence in estimates derived from visual information [And11]. This overconfidence can lead them to rely on inaccurate visual inputs, resulting in biased decisions [And11]. Sun et al. and Hevia et al. show that the decision-making of individuals can be influenced by changing the physical properties of graphs [SLBL16, dHGBV08]. Altering the graph scale and relative physical distance of the datapoints might change how viewers perceive the graph [SLBL16, dHGBV08]. Since humans take into account the physical distance of datapoints, even if this distance does not accurately represent the real datapoints, people's bias can be influenced by only changing the physical attributes of the graph [SLBL16, dHGBV08]. These visualisations should thus be made to contain as few biases and inaccuracies as possible.

Attitude change: Visualisations can also influence the attitude of viewers, meaning a change in how they behave around certain data [Ebe23].
Okan et al. and Dambacher et al. show that visualisation design can influence the willingness to engage in risk-avoidant behavior [OSBdB18, DHGH16]. By using foreground-only displays, the number of affected individuals becomes more noticeable, resulting in viewers having an increased willingness to take a drug for heart attack prevention [OSBdB18]. Probabilities presented graphically instead of numerically also increased risk aversion in people [DHGH16]. This is due to the graphs making the probability of winning more noticeable [DHGH16], making people realise that they will probably lose money. Visualising the potential outcomes, however, decreases the risk aversion in people if visualised in the right way [DHGH16]. Dambacher et al. state that the use of icon arrays can decrease the risk aversion by making the potential gains more noticeable [DHGH16]. Pie charts, however, increase risk aversion in people [DHGH16]. The authors suggest this is probably because pie charts are an uncommon representation for monetary values, leading to difficulties in comprehension [DHGH16].

Visualisations can thus be an important tool for people to get a grasp of a great amount of information, and can be used in scientific fields so people can quickly and efficiently get insights into the state of the field [Ebe23]. It is, however, important to state that there is a great risk of misinterpreting information with visualisations, so researchers should be careful not to add bias to their visualisations. Comprehension/decision accuracy is particularly relevant for understanding research fields because the main goal when researching a field is to comprehend the information as accurately as possible and make decisions based on that information. We make use of visualisations to help users achieve an increase in comprehension/decision accuracy.

2.3.2 Existing tools for visualising papers

Visualisations are used to show the connections between papers in a field. This helps researchers quickly find related papers and speed up the process of researching the current state of the field. There are multiple tools that do this, like Litmaps [Lit], Connected Papers [Con], and ResearchRabbit [Res]. These tools give a visualisation in the form of a graph of related papers.

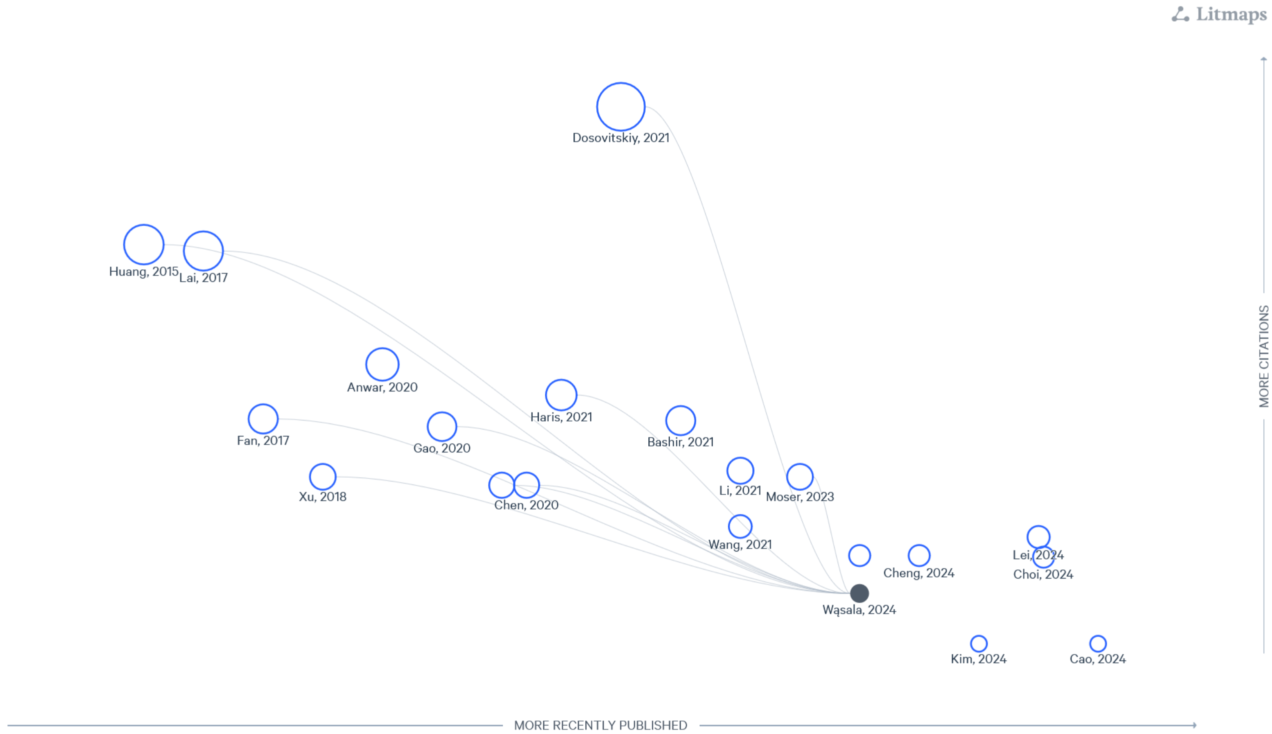

Figure 2: Litmaps graph showing the citation network found when searching for the paper: "AutoSR4EO: An AutoML Approach to Super-Resolution for Earth Observation Images" [WMA+24]

This graph is created by looking at the citations. Examples of Litmaps' and Connected Papers' graphs can be found in Figure 2 and Figure 3.

Litmaps: Litmaps was created to solve the problem of having to find related papers when researching a topic [KGS+22] by creating citation networks, a network of papers that cite each other. The papers on the visualisations are ordered by publication date from left to right, with earlier-published papers on the left and later-published papers on the right, and by increasing citations from down to up. These citation networks are created by following the citations of papers to create a big network of papers [Lit]. Navigating the visualisation created by litmaps makes navigating a field fast and efficient. Kaur et al. reported Litmaps as a very effective tool that significantly improves the process of exploring scientific literature [KGS+22]. With interactive visualisations, robust search features, a recommendation system, and dynamic mapping functionalities, this tool helps researchers discover relevant work [KGS+22]. By visualising the citation network of relevant papers, grasping all of the available literature in the field becomes much easier [KGS+22].

Litmaps does, however, lack in providing visualisations about the contents of the papers. It only shows papers that cite each other, but reading and researching the papers still have to be done by the user. Since Litmaps does not look at the contents, it also means that there will be many papers that are partially about the topic you are researching, e.g. there will be many EO papers cited in an AutoML for EO paper, even though you might only want other AutoML for EO papers. You will thus still have to filter the papers to find the ones that you want. It also has a freemium model, meaning you have a limited number of searches if you do not pay. This can be a limiting factor for

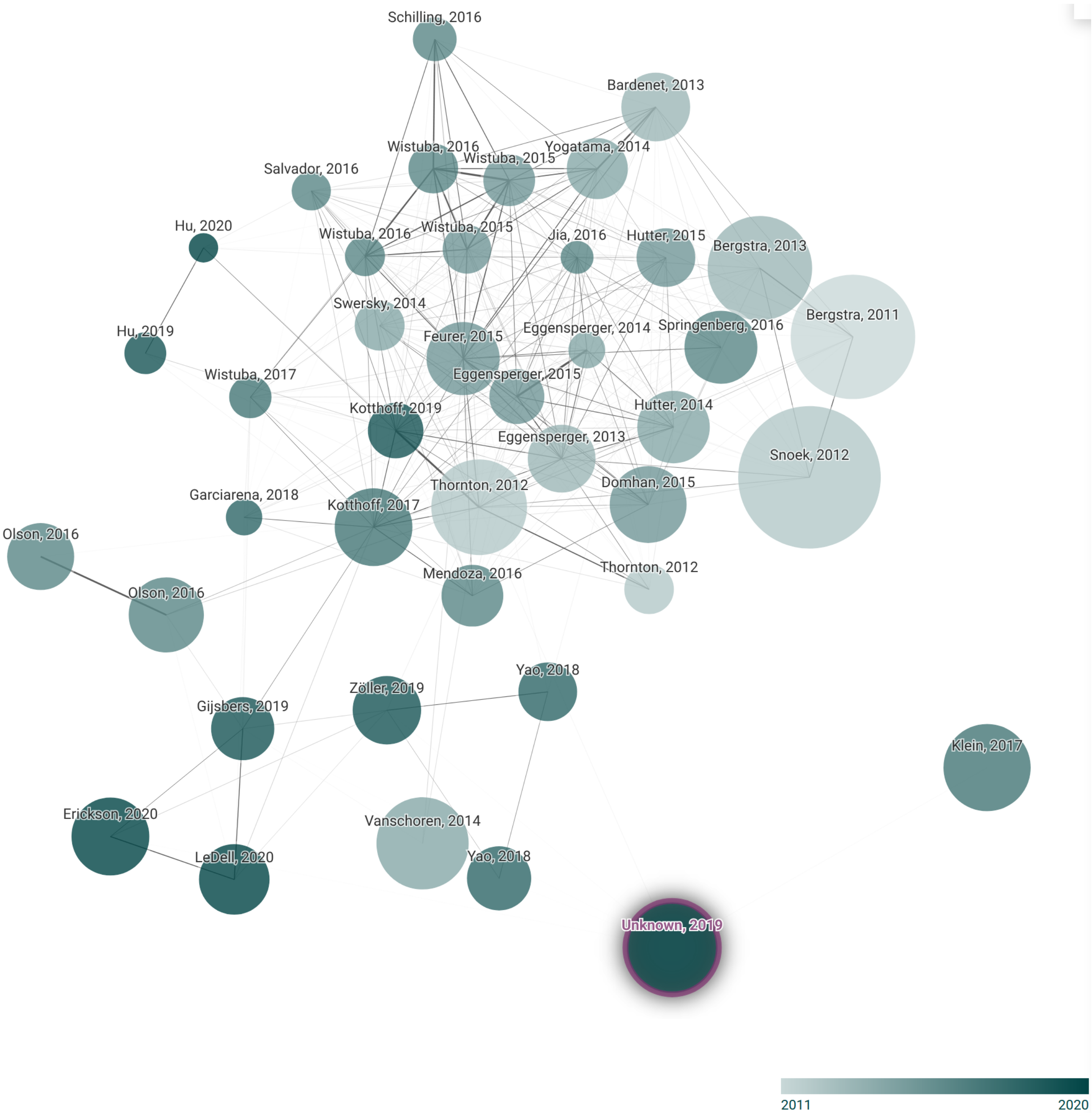

Figure 3: Connect Papers graph showing the citation network found when searching for the paper: "Automated Machine Learning - Methods, Systems, Challenges" [HKV19]

people who do not have enough money.

Connected Papers: Connected Papers was created for the same reason as Litmaps, making finding related papers easier [KSM+22]. Just like Litmaps, Connected Papers creates a network visualisation based on the citations of the papers. However, unlike Litmaps, which just makes a graph by recording the citations as edges from one paper to another, Connected Papers uses a similarity metric between papers to decide which papers will be connected in the visualisation. Two papers with highly overlapping citations and references have a higher chance of being related [Con]. The papers on the visualisation are not ordered, but change in colour based on the publication year and change in size of the vertex based on the number of citations. Behera et al. and Kauer et al. report similar thoughts about Connected Papers [Con], stating it provides an easy-to-navigate graph showing relevant literature in the field, making navigation and understanding of the current state of the field much easier [BJK23, KSM+22].

Connected Papers suffers from similar drawbacks to Litmaps, where it does not provide visualisations about the contents of the papers, and it has a freemium model. Connected Papers will, however, probably provide more relevant suggestions since it uses a similarity metric between the papers. A paper about AutoML for EO will probably have citations about both AutoML and EO, and will thus have a higher similarity metric with other papers about AutoML for EO.

Another drawback is that the visualisations are less intuitive than those of Litmaps. The nodes are not ordered in any way, and with more papers in one search result, it can look a bit chaotic. However, more papers in one search result can also be seen as a good thing, as that means that there are more relevant papers you can explore.

ResearchRabbit: ResearchRabbit is an AI-powered tool for researching literature [Res]. Similar to the previous two tools, ResearchRabbit was also created to simplify the process of researching the current state of the research field by finding relevant papers for you [Res]. Just like the previous tools, you can create networks of similar papers based on the citations using ResearchRabbit. Additionally, describing itself as the Spotify for papers, ResearchRabbit comes with a recommendation system that learns what you love based on the papers you have added to your collections [Res].

ResearchRabbit has the drawback that its UI is unintuitive. When trying out this tool, we found it difficult to navigate the app and figure out how to make visualisations.

2.4 Thesis Contribution

We propose a tool that can be used to quickly gain insight into the field of AutoML for EO by providing visualisations of the papers in this field. In contrast to previous work, we use both the metadata of the papers and the substantive information of the papers to create these visualisations. Currently, that substantive information consists of the AutoML frameworks used, the datasets used, and the EO tasks performed. While these existing tools do help with speeding up the process of navigating and understanding the state of a field and the relationships between papers, they do not help with understanding the methodologies and trends specific to the field, in this case, AutoML for EO. Visualisations of the substantive information are needed to be able to thoroughly understand the state of the field, including the technologies, methodologies, trends, etc.. Our contribution addresses this gap by making visualisations for the field of AutoML for EO. This includes information about the different frameworks and databases used, and the tasks performed using these frameworks and databases. With these visualisations, it becomes easier to get answers to questions about what kind of technologies are popular and what subtopics still need to be researched.

3 Methods

We propose a set of visualisations giving insights into AutoML for EO papers, which will be a companion to a survey paper by Wąsala et al. [WMS+26] by displaying the information in a new way, making it easy to see certain trends in the field. This tool makes the survey data simpler to interpret and understand by highlighting trends and patterns, which is difficult with only a textual survey. To create these visualisations, we first need a dataset of papers. Since a big dataset is not readily available in this field, we create a scraper to make that dataset ourselves. There are two versions of the scraper used to make this dataset, one that makes use of both the Google Scholar[1] and Crossref[2] APIs, and one that makes use of only the Crossref API. Finally, there is a small number of papers that were manually found and added to this dataset. Since the scraper was not able to find a large number of papers, and retrieving the substantive information is difficult, we additionally use a list of papers from a survey paper on AutoML for EO provided by the authors [WMS+26]. This dataset contains per-paper information about methods, datasets, and tasks. The pre-processing of this dataset is discussed in the last part of this subsection. Using these datasets, we create visualisations for researchers to answer questions about what technologies are used and what EO tasks can be further researched. We will talk in more detail about the custom-built scraper and the creation of the visualisations in the next two sections.

3.1 Scraper

There are two versions of the scraper, one that uses both Google Scholar and Crossref, and another that only uses Crossref. The first version collects papers in three steps: gathering the titles, retrieving the metadata, and filtering papers. The second version combines the first two steps into one step. The processing of the existing dataset is also addressed.
A flowchart of the general process of retrieving papers can be found in Figure 4.

3.1.1 Gathering titles

The scraper makes use of the Google Scholar API to first collect titles of papers that are returned in the search result. The query used to search for papers is as follows: “("AutoML" OR "Automated Machine Learning" OR "Neural Architecture Search") AND "Earth Observation"”. We cannot use the abbreviations NAS and EO, for Neural Architecture Search and Earth Observation, respectively, to search for papers. This is because nas and eo are words in the Portuguese language meaning “in the” and “and the” respectively. Because of this, when those abbreviations were used in the query, a great number of Portuguese papers were returned in the search results, which lowers the number of papers found that are about AutoML for EO.

The scraper searches for papers that are about Automated Machine Learning and Earth Observation or Neural Architecture Search and Earth Observation on Google Scholar using the mentioned query. This is done in the form of a Python script that uses the scholarly [CISK23] package (unofficial) to search for papers on Google Scholar. The query is given to the search_pubs() function, which then returns an iterable containing dictionaries that have some of the metadata of the paper. We later use the crossref_commons API to search for the metadata of the papers since

[1] https://scholar.google.com/
[2] https://www.crossref.org/

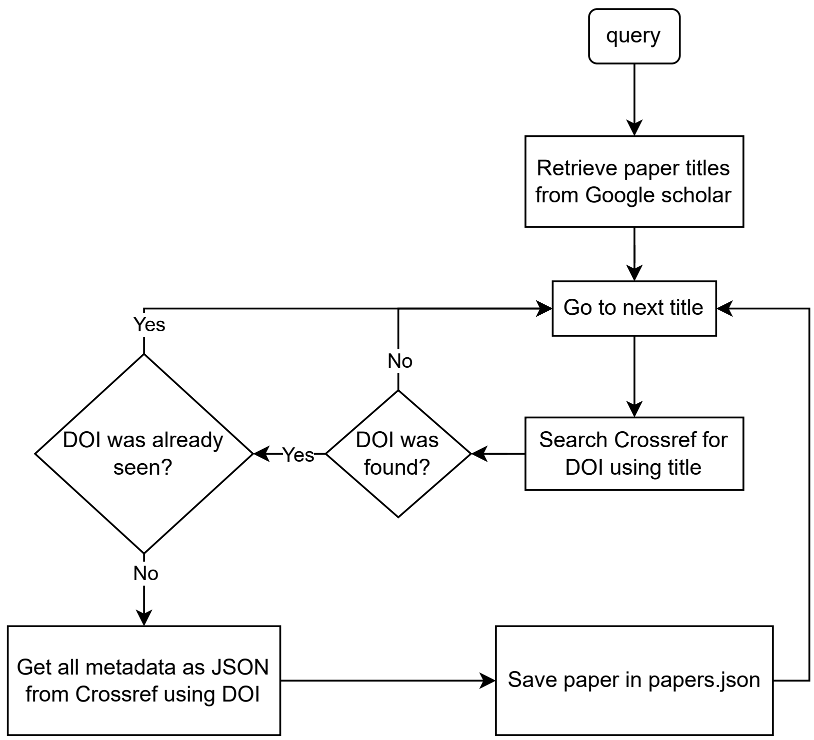

Figure 4: Flowchart showing the process of the scraper

Table 1: The metadata of the papers that is collected.

1.	title
2.	authors
3.	date
4.	type (e.g. paper article, proceedings article, book chapter, etc.)
5.	DOI
6.	publisher
7.	funders
8.	link
9.	abstract
10.	journal (name of the journal(s) this paper was published to)
11.	short-journal (shorter name of the journal(s))

most of the metadata, like DOI or abstract, is missing in these results. We thus only need the titles of these papers in one big list, which is then used to search for the same paper in the Crossref API. Using this method, around 700 titles are gathered.

3.1.2 Retrieving metadata

We use the Crossref API to obtain the remaining metadata we want to use in the visualisations. This is done using the official crossref_commons Python package [Cro20]. This package has a function iterate_publications_as_json, that searches for publications that most closely match the given title. We just take the first found publication for our case. We sometimes have a mismatch since Crossref does not have all existing papers in its database, and the iterate_publications_as_json function returns the paper with the best matching title. This is not a big problem, however, since a paper that has a similar title is probably also about a similar subject and could thus also be used in the visualisations if they are about AutoML for EO. The DOI is checked against all previously found papers after getting the DOI of the closest matching paper to prevent duplicate papers in our dataset. The DOI is used with the get_publication_as_json function from crossref_commons [Cro20] if there is no duplicate to retrieve all metadata Crossref has on that paper in the form of a JSON. We then take the metadata we are interested in (Table 1). The title is updated to the title in the Crossref database since we do not know if it is the same paper.

The other version of the scraper, which only uses the Crossref API is used to retrieve the metadata of 300 papers using the query: “("AutoML" OR "Automated Machine Learning" OR "NAS" OR "Neural Architecture Search") AND ("EO" OR "Earth Observation")” resulting in 1000 papers in total from both scrapers. The abbreviations NAS and EO were used in this query since this version did not have the problem of returning many Portuguese papers. The papers found with these scrapers are then moved to the next section, where they are filtered.

3.1.3 Filtering papers

Data that is scraped automatically needs to be cleaned up to remove duplicates and to filter out papers that do not meet certain criteria, since there are often papers that match the query, but are not about this topic. This can happen because Google Scholar returns results that best match the query given, but since the amount of papers on this topic is limited, a great number of papers that are only somewhat related, e.g. the paper "FPGA-Enabled Machine Learning Applications in Earth Observation: A Systematic Review" [LSS25] which contains ML4EO but not AutoML4EO, are given. Without cleaning up the data, you would thus get a messy dataset that might not represent the field very well, which would lead to visualisations that are incorrect. The criteria for a paper to be accepted are: they have to be about a topic in the field of AutoML for EO, and they have to be in English. The dataset is filtered to filter out all the unrelated/unsuitable papers, meaning they are not about the topic AutoML for EO or they do not apply AutoML to EO tasks, to make sure all the papers in the dataset meet the criterion. This is done by a script that first tries to check if the paper is about AutoML for EO and in English automatically. The script first checks if the language attribute of the paper is English, and then it checks if automl/automated machine learning and earth observation/neural architecture Search are both in either the lowercase title or the lowercase abstract. We check with the lowercase title/abstract to prevent any mismatches because of a difference in the case.

Since there are many papers that do not explicitly state in the title or abstract that it uses AutoML, or does a task in EO, if a paper was not accepted automatically, the user has to manually check if it is about AutoML for EO to make sure it is not about this topic and tell the script yes or no.

After filtering, out of the 700 papers of the Google Scholar and Crossref versions, 66 remained. And out of the 300 of the Crossref only version, 18 remained. With 3 of the papers overlapping, this resulted in a total of 81 accepted papers. We also added another 10 papers that were manually discovered. And with 4 overlapping papers, this dataset contains 87 papers in total.

3.1.4 Survey dataset

Another source of papers we use is a dataset of 81 papers that was made for the survey by Wąsala et al. [WMS+26] since our dataset is limited in size and does not contain any substantive information. The dataset created using our scraper will be referred to as the scraper dataset, and the dataset from the survey by Wąsala et al. [WMS+26] will be referred to as the survey dataset. The survey dataset contains three fields with substantive information, which are: task, dataset, and framework, which is a good source to create visualisations. The dataset field contains any collection of data that is used to train an ML model to do a task, or the source from which the dataset was made (e.g. which satellite the data was gathered from) when the dataset itself is not available publicly. The framework field contains any big library that is used for training ML models. This includes mostly AutoML libraries, but also often used algorithms like DARTS, an Architecture Search method [LSY19], and non-AutoML libraries like YOLO, a real-time object detection model [RDGF16]. Lastly, a task is defined as any problem that can be learnt and solved by ML models. This includes general tasks like Image classification and Land use, but also more specific tasks like Groundwater potential mapping, Soil moisture estimation, and grass height estimation.

The survey dataset does, however, miss most of the metadata used for the visualisations, namely: DOI, type, publisher, funders, and abstract. The previous method of retrieving the DOI using the title cannot be used since these papers have important information already, and often the paper found is not the same as the one used to search for it. This would be a problem in this case since the task, dataset, and framework information would then be incorrect for the paper found. So, to find the DOI for every paper, a scraper was made for every publisher site that appears. It uses the link attribute that every paper in the dataset has and finds the DOI on the site of the link. For every site, it uses a simple regex unique to that site to find the DOI inside the HTML code or the link itself, since often the DOI is used as part of the URL. The found DOIs are then used to get the rest of the metadata, just like in the scraper.

This dataset is used most often for the charts since only the survey dataset contains the substantive information, and the charts are mostly focused on the substantive information. The scraper dataset is used together with the survey dataset for visualisations, where substantive information is not needed, to compare and gain additional insight into the data from these datasets.

3.2 Visualisations

We made visualisations by first exploring the datasets and identifying interesting patterns, and then designing visualisations to answer questions we had thought of about the field of AutoML for EO. We identify these interesting patterns and relevant questions by exploring the dataset and finding these patterns by using Bokeh [Bok18] to quickly create simple visualisations of arbitrarily chosen fields of the datasets, and networkx [HSSC08] to quickly create graph visualisations. When a visualisation could be helpful in informing users about the state of this field, we think of a question that can be answered with them, and we try to create a final version using D3.js [BOH11], which can be showcased on a website, and think of more visualisations that may help explain the question better or other questions that come up when exploring the first visualisation.

The goal of this exploration is to identify patterns in the metadata of the papers and think of questions that could be answered with these patterns. The exploration is done by making simple bar charts of two of the fields of the metadata, or thinking of a possible network that can be made and implementing that in networkx. Using these visualisations, we thought of questions that could be answered with these visualisations and implemented the relevant visualisations in D3.js.

The visualisations used in this thesis are the following:

- Figure 5: Shows what frameworks are used a lot for certain EO tasks.
- Figure 6: Shows what the most popular frameworks are.
- Figure 7: Shows what frameworks were popular in what year.
- Figure 8: Shows what EO tasks were popular in what year
- Figure 9: Shows what journal is used to publish papers with certain EO tasks.
- Figure 10: Shows what the most popular tasks are.
- Figure 11: Shows what the most popular publishers are in the survey dataset [WMS+26].
- Figure 12: Shows what the most popular publishers are in the scraper dataset.
- Figure 13: Shows what the least popular tasks are.
- Figure 14: Shows what the most popular journals are in the survey dataset.
- Figure 15: Shows what the most popular journals are in the scraper dataset.

- Figure 16: Shows when the most papers were found in the survey dataset.
- Figure 16: Shows when the most papers were found in the scraper dataset.

These visualisations are showcased on a website that can be used as a companion to the survey by Wąsala et al. [WMS+26].

These visualisations can, as stated in the related work 2, help with increased efficiency and speed in gaining insights into the state of the field. These visualisations can also help researchers with making decisions by helping them get a better understanding of the field.

4 Experimental Design

For this section, we define two users, Dr. A and Dr. B. Dr. A is an EO researcher who is working on a large satellite dataset, trying to classify different biomes in Africa. He finds the dataset to be too large to analyse manually and wants to use ML to analyse the data efficiently, even though he does not know much about ML. With the lack of knowledge about ML, coupled with the many AutoML frameworks that are available, he is struggling to choose which framework to use for their use case. Dr. B is also an EO researcher and very new to this field. She is researching object detection from satellite data. Since she is very new to this field, she lacks the knowledge about what topics are common for which journal. So she does not know which journal she should approach for her research subject. These users are used to examine how these visualisations can be useful to them in answering two of the research questions.

Our empirical design aims to answer the following questions:

RQ1 What (AutoML) frameworks should you use for what task? Dr. A can use this research question to decide which frameworks they should work with.

RQ2 In what journal do I publish my paper? Dr. B can use this research question to decide which journal they should publish their paper in.

RQ3 What task should I research?

RQ4 What are the differences between the used datasets?

4.1 What (AutoML) frameworks should you use for what task?

Answering this question can be beneficial in many ways. This can help researchers choose what framework is popular, and thus probably a good option, for a given task. A researcher may also want to try new combinations of frameworks and tasks that have not been used that much to see if they give good results.

The visualisations of this thesis can be used to solve this problem by seeing if there are similar tasks that were researched, and what frameworks were used for those tasks. Figure 5 was made with the dataset curated for the survey on AutoML for EO [WMS+26], which contains information about the frameworks and tasks, to plot the top three frameworks for a given task. Dr. A could search for similar tasks to what they are trying to achieve and see what frameworks are often used in those situations by looking at this chart. Dr. A could decide which frameworks they should use based on this information.

4.2 In what journal do I publish my paper?

Answering this question can help researchers with choosing the correct journal to publish their paper in. Figure 9 can be used to solve this problem. In this chart, similar tasks can be looked at to see which journal is most popular for similar tasks. Dr. B could search for similar tasks to what they are doing and see which journals are used to publish those papers by looking at this chart. Dr. B could decide which journal to publish their paper in based on this information.

4.3 What task should I research?

The answer to this question can be found by looking at the most and least popular tasks in the field of AutoML for EO. You can see what has been extensively researched and what still needs more research by looking at the most and least popular tasks. This information can also help researchers find popular tasks that are similar to a task that is not that popular, from which they may get insight or ideas for their task.

4.4 What are the differences between the datasets used?

This answer to this research question provides insights into the differences in the datasets and the pros and cons of the datasets that are used in this thesis.

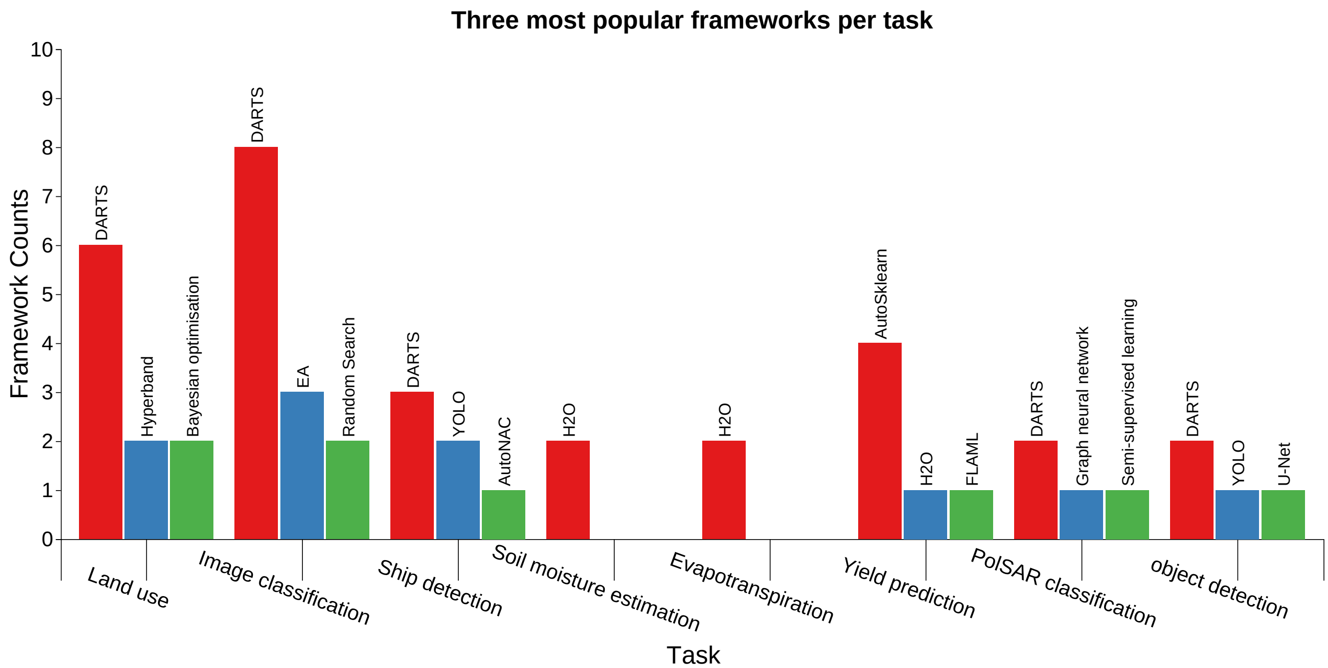

Figure 5: The 3 most used AutoML frameworks per Earth Observation task (not all tasks are shown). Showing DARTS as a very popular framework among the classification/detection tasks. whereas H2O/AutoSklearn is more popular among the prediction tasks. Only the top three showing the rest makes the visualisation cluttered, and is not needed to understand this visualisation. This visualisation is created with the survey dataset since information about the task and framework is needed. The usage of the survey dataset applies to all other visualisations except if stated otherwise.

5 Results

In this section, we explore the results obtained to answer the questions in Experimental Design 4.

5.1 (RQ1) The most/least popular framework for a given task

We will look at a few of the visualisations that were made to answer the first research question 4.1, "What frameworks should you use for what task?". Figure 5 shows the most used AutoML frameworks per EO task. This visualisation can be used to answer RQ1 by comparing the popularity of the frameworks for tasks similar to yours. Dr. A could look at this visualisation and see that DARTS (Differentiable Architecture Search) [LSY19] is a popular framework for classification tasks like Land use, Image classification, and PolSAR classification, and decide to research more about this framework. If Dr. A then finds that DARTS excels at image classification tasks, as mentioned by the creators of DARTS [LSY19], he could also choose to use that framework. He could also, however, choose another framework that is not used much for classification tasks and research if they give good results to fill that knowledge gap.

Ship detection and Object detection, two detection tasks, also mainly make use of DARTS with YOLO (You Only Look Once) [RDGF16] as a close second. YOLO is a pre-designed end-to-end neural network, a type of neural network that is trained from raw input to final output [BYC+17], that specialises in object detection and generalises well in new domains [RDGF16]. The reason

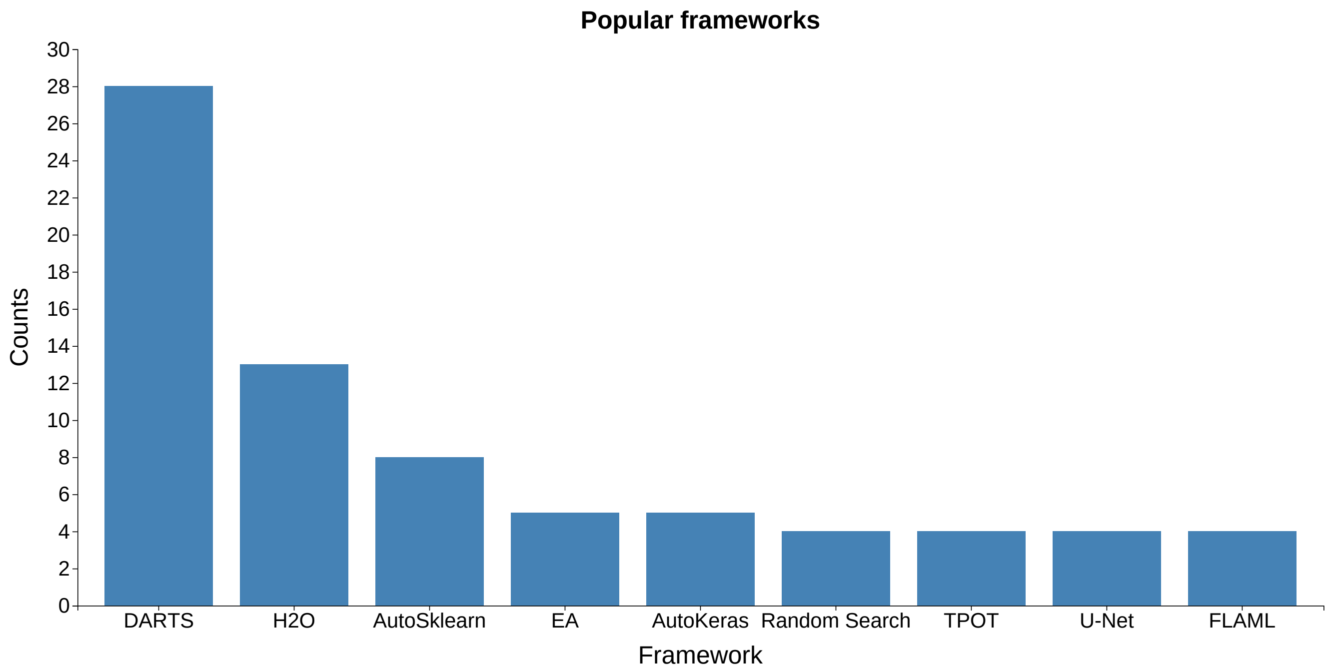

Figure 6: Top-most used frameworks in AutoML for EO papers (not all frameworks are shown). Shows DARTS as the most used framework with 28 appearances in 81 papers. The second and third most used frameworks are H2O and Auto-sklearn with 13 and 8 appearances respectively.

YOLO is used frequently in object detection tasks, even though it is not an AutoML framework, is that it is a commonly used backbone in NAS frameworks, like DARTS and AutoNAC, so they show up in the dataset together.

The soil moisture estimation, evapotranspiration (evaporation + transpiration where water moves from the surface into the air [KAL+14]), and yield prediction, which are all regression tasks, mainly make use of H2O [H2O22] and Auto-sklearn [FKE+15, FEF+20]. A possible explanation for H2O and Auto-sklearn being the most used for those tasks can be seen in Figure 6. H2O and Auto-sklearn are the most popular frameworks behind DARTS. And since DARTS is more focused on image classification and language modelling, H2O and Auto-sklearn naturally become the top used. This corresponds with the survey paper of Wąsala et al., which states that H2O is the most popular framework for EO tasks. Truong et al. performed a benchmark on a few commonly used AutoML frameworks [TWG+19]. They showed that besides Auto-keras, H2O, and Auto-sklearn performed better in the regression task compared to the other frameworks, which could also be a reason they are used predominantly in these tasks that require regression.

Another interesting thing to note, as can be seen in Figure 7, is the sudden disappearance of the DARTS framework out of the top five in 2023 and not appearing again in the following years, even though it was by far the most dominating and fast growing framework in the years 2019-2022. Looking at Figure 8, we can see another possible explanation for the decrease in DARTS use. We can see that in that same year, the tasks Image classification and Land use classification both also decrease significantly. Since those two tasks were the heaviest users of the DARTS framework, as can be seen in Figure 5, the decrease in this task would naturally mean a decrease in the use of the DARTS framework. Another possibility of the sudden decrease in DARTS use could be a

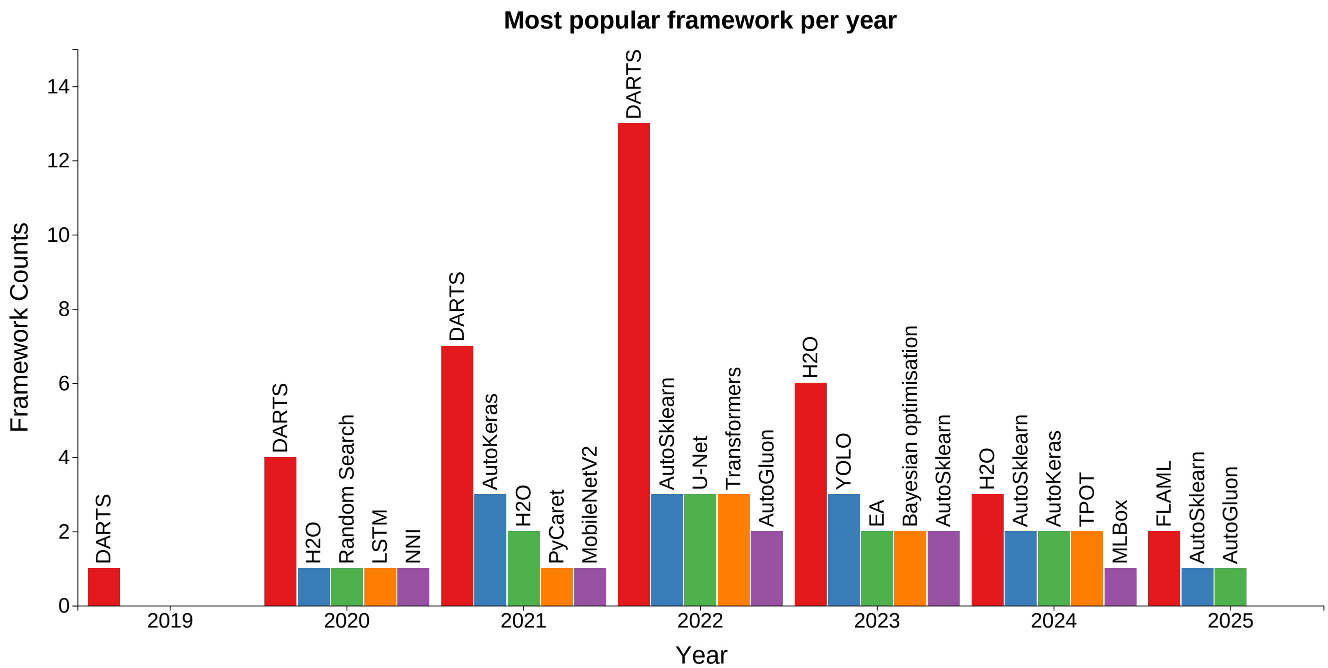

Figure 7: Top five most used frameworks in AutoML for EO papers per year. Shows DARTS as the most used and growing framework in the years 2019-2022, with a sudden disappearance out of the top five in the years 2023-2025. Other popular frameworks that are often at the top are Auto-sklearn, H2O, and YOLO.

coincidence because of the limitation that our datasets are not very large. Because of the limited number of papers in our dataset, especially in early and later years, it could be a coincidence that there are no/very limited DARTS papers in the dataset.

5.2 (RQ2) The most popular journal for similar tasks

To answer the question of what journal to approach 4.2, we will look at the visualisation shown in Figure 9. Figure 9 shows the number of times a journal was used to publish a paper about a specific EO task. Dr. B is searching for tasks related to object detection, which in Figure 9 are the tasks Object detection and Ship detection. We can see that Object detection has one paper published in "IEEE Transactions on Geoscience and Remote Sensing" and one paper published in "Remote Sensing". And Ship detection also has one paper published in "IEEE Transactions on Geoscience and Remote Sensing" and the other one in "IET Conference Proceedings". Based on this information, Dr. B could conclude that they should publish their paper in the IEEE journal, which has the most papers on a related task.

There are only 4 papers published for a task related to object detection, as can be seen in Figure 9. This is not a sufficient amount of data for people to be able to make well-informed decisions. One reason is the small size of the dataset, with 81 entries and 34 tasks; some of the tasks will have a small number of papers, which makes searching for patterns of these tasks difficult. Another reason is the lack of journal title data in the Crossref database. We can see that Ship detection has 4 papers from Figure 10, but Figure 9 shows only 2 papers for Ship detection. This is due to

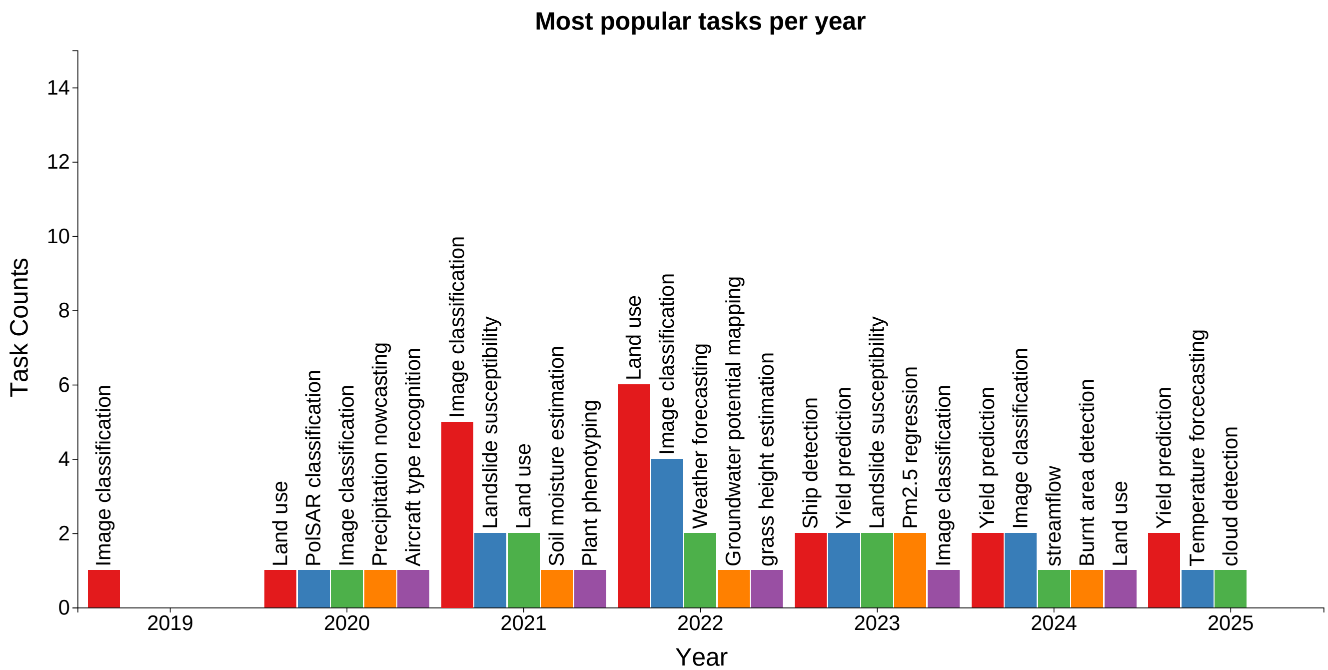

Figure 8: Top five most performed tasks in AutoML for EO papers per year. Shows Image classification and Land among the most popular tasks from 2019-2022. Yield prediction is one of the most popular tasks in the years 2023-2025.

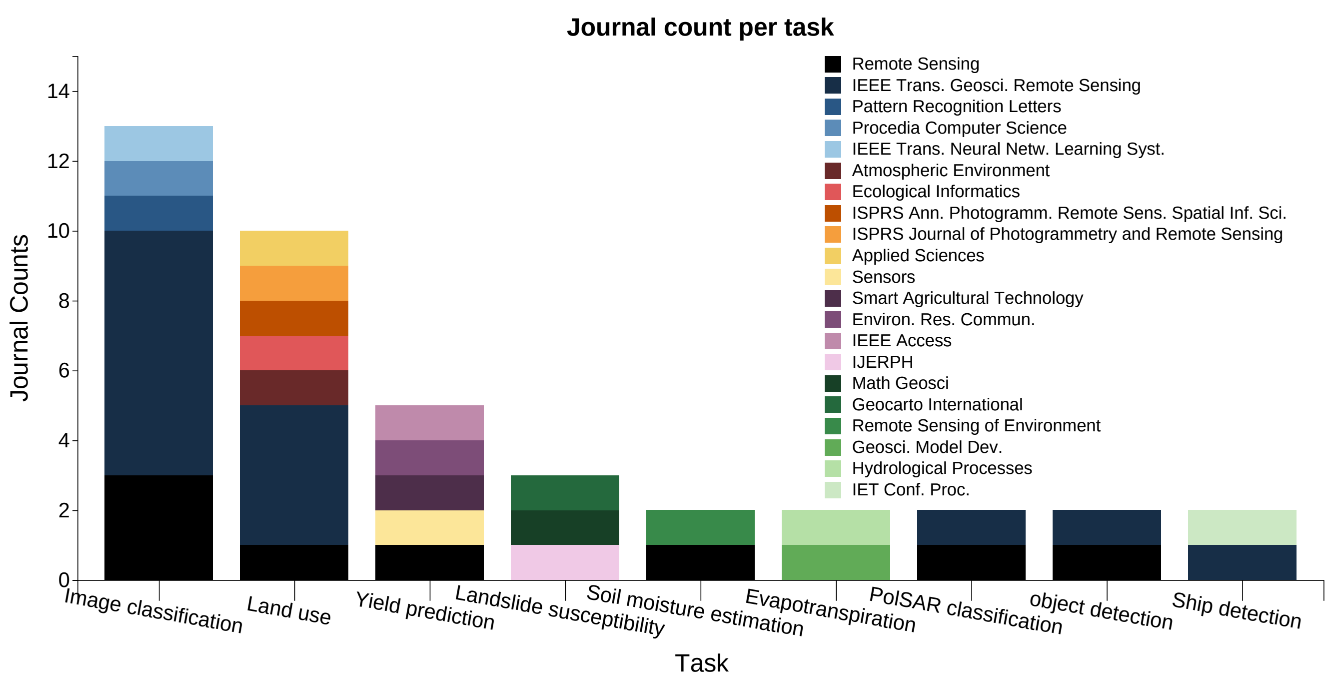

Figure 9: Journal Counts for AutoML for EO papers (not all tasks are shown). Most of the papers from Image classification or Land use are published in “IEEE Transactions on Geoscience and Remote Sensing” from IEEE or “Remote Sensing” from MDPI. In other tasks, journals from these two publishers also show up frequently. A stacked bar chart is used since there are many different journals which can’t be shown horizontally like the earlier charts.

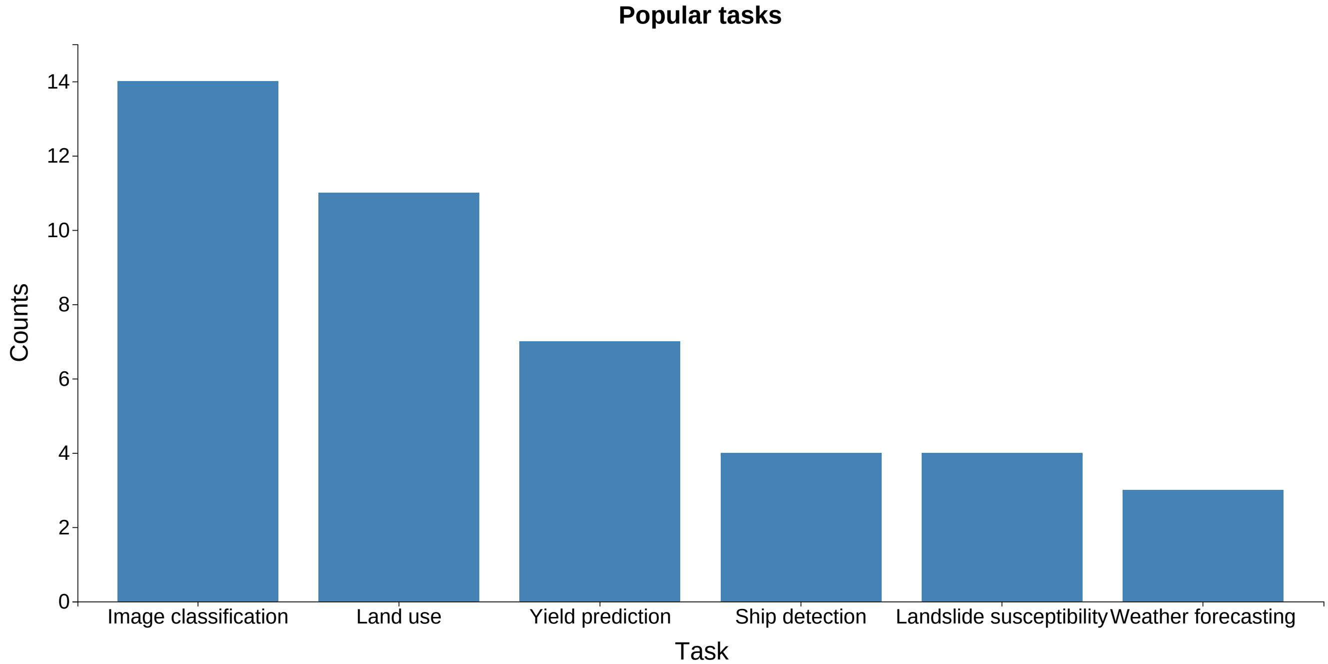

Figure 10: Number of times a task appeared in the dataset (tasks with less than 3 counts are not shown). Image classification is the most popular task with 14 appearances, followed by Land use and Yield prediction with 11 and 7 appearances, respectively. Ship detection and Landslide susceptibility both have 4 appearances, and Weather forecasting has 3.

some of the entries missing the short journal name. And since some of the long journal names are too long to display properly, these cannot be used. The decision of what journal to approach may be made easier if there is more data to compare. However, the amount of substantive information about the papers is limited in our current dataset.

Figure 11 shows why IEEE and MDPI journals might show up frequently. In Figure 11, we can see that IEEE and MDPI are by far the most dominant publishers in this dataset, with them having published more than half of the papers in the dataset. This may be because of a bias in collecting the papers. With IEEE and MDPI both being among the biggest platforms, other platforms with papers about this topic could be overlooked.

Figure 12 confirms that the dataset from the survey [WMS+26] might be biased towards IEEE. While the dataset from the survey contains 29 entries of IEEE papers, the dataset from our scraper contains only 14 entries of IEEE and more entries from other publishers like Springer and Elsevier.

5.3 (RQ3) Understudied tasks

The question of what task should be researched 4.3, can be answered by looking at Figure 10 and Figure 13. Looking at Figure 10, we can see that Image classification, Land use, and Yield prediction have all been extensively researched, which may make it difficult to find new problems to work on in this field, or make it easier because of the many examples available.

Looking at Figure 13, we can see all the tasks that have not been researched much and could use more contributions. A researcher could, based on these visualisations, decide to work on one of

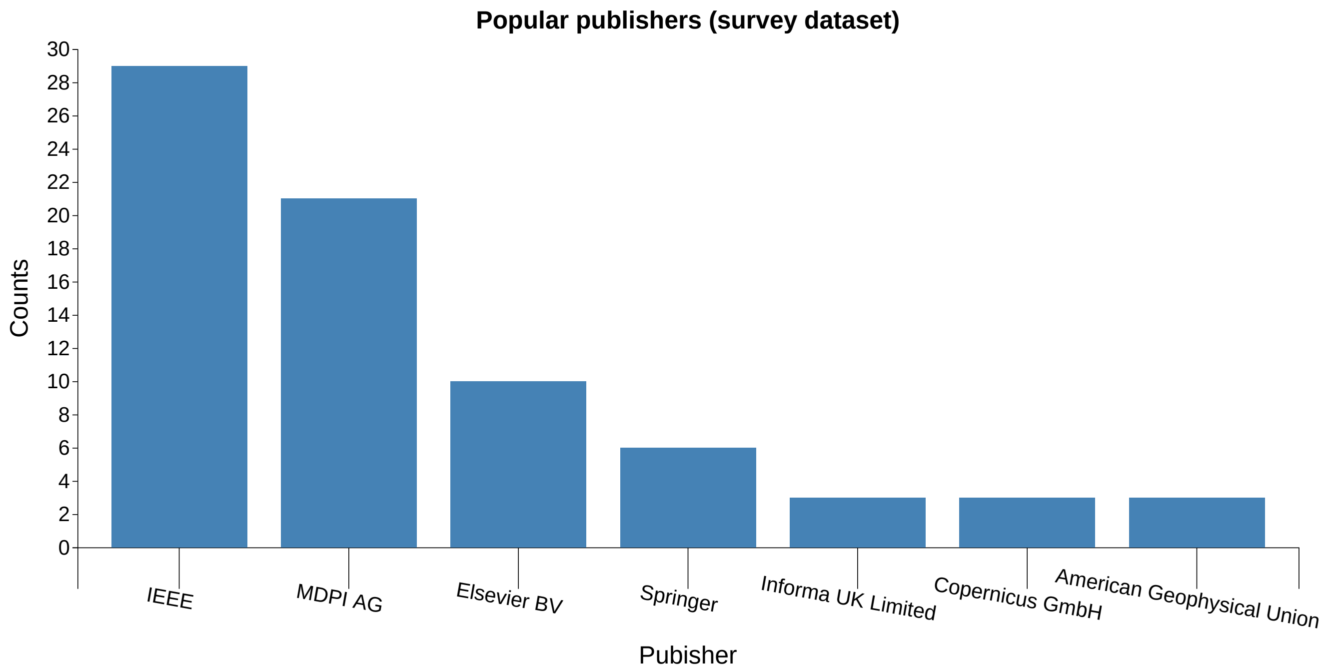

Figure 11: Number of times a paper has been published by a publisher (not all publishers are shown). IEEE is by far the most used publisher, with 29 out of 81 papers being published by IEEE. MDPI is also pretty dominant; together, IEEE and MDPI published 50 out of 81 papers. This visualisation uses the dataset from the survey [WMS+26] since it is used to explain the visualisation in Figure 9, which is also made with this dataset.

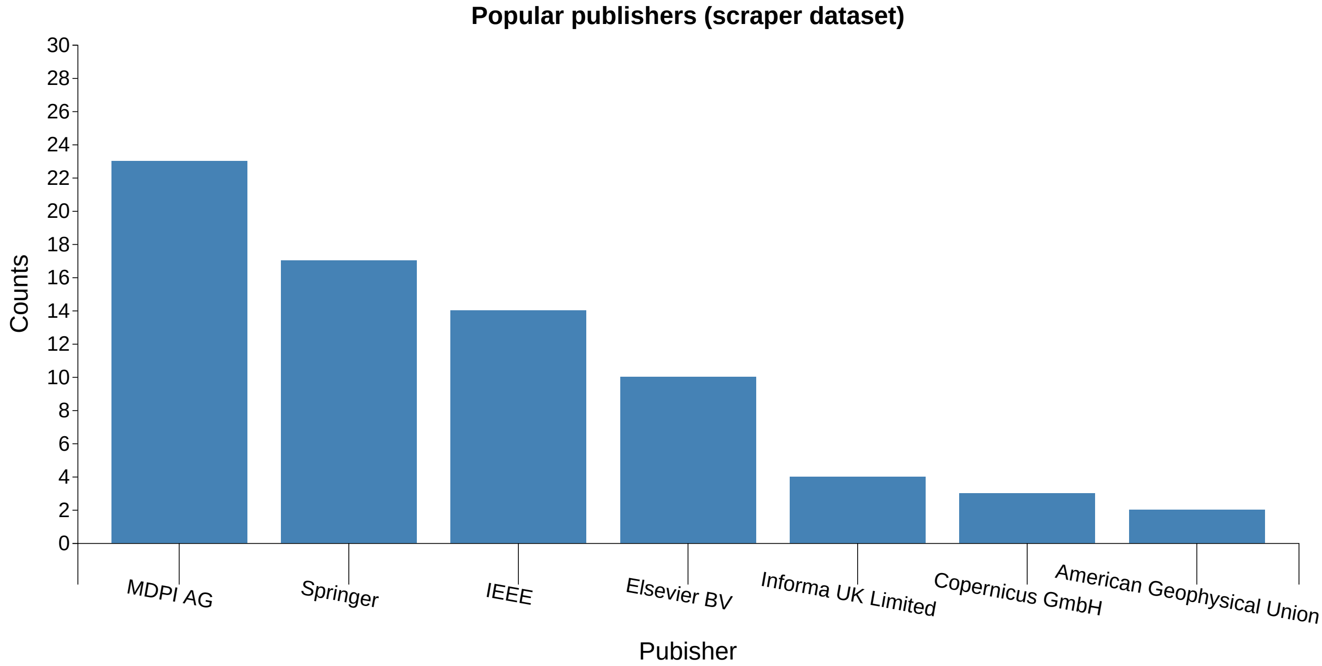

Figure 12: Number of times a paper has been published by a publisher (not all publishers are shown). MDPI is the most used publisher, with 23 out of 81 papers being published by MDPI. SPRINGER and IEEE are also pretty dominant; together, MDPI, SPRINGER, and IEEE published 54 out of 81 papers. This visualisation is made with our dataset from the scraper since we use this visualisation to compare to the same visualisation made with the dataset from the survey [WMS+26].

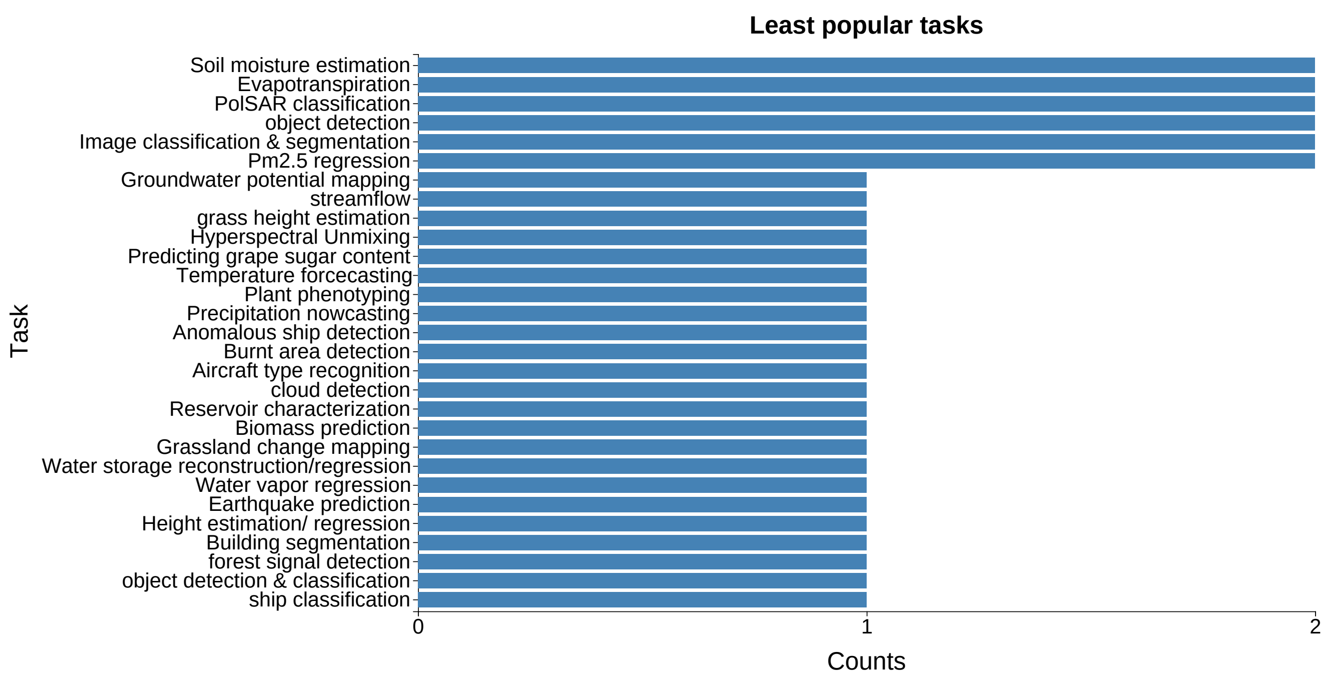

Figure 13: Number of times a task appeared in the dataset (tasks with more than 3 counts are not shown). The tasks from Soil moisture estimation to PM2.5 regression all have 2 appearances. The remaining tasks all have 1 appearance.

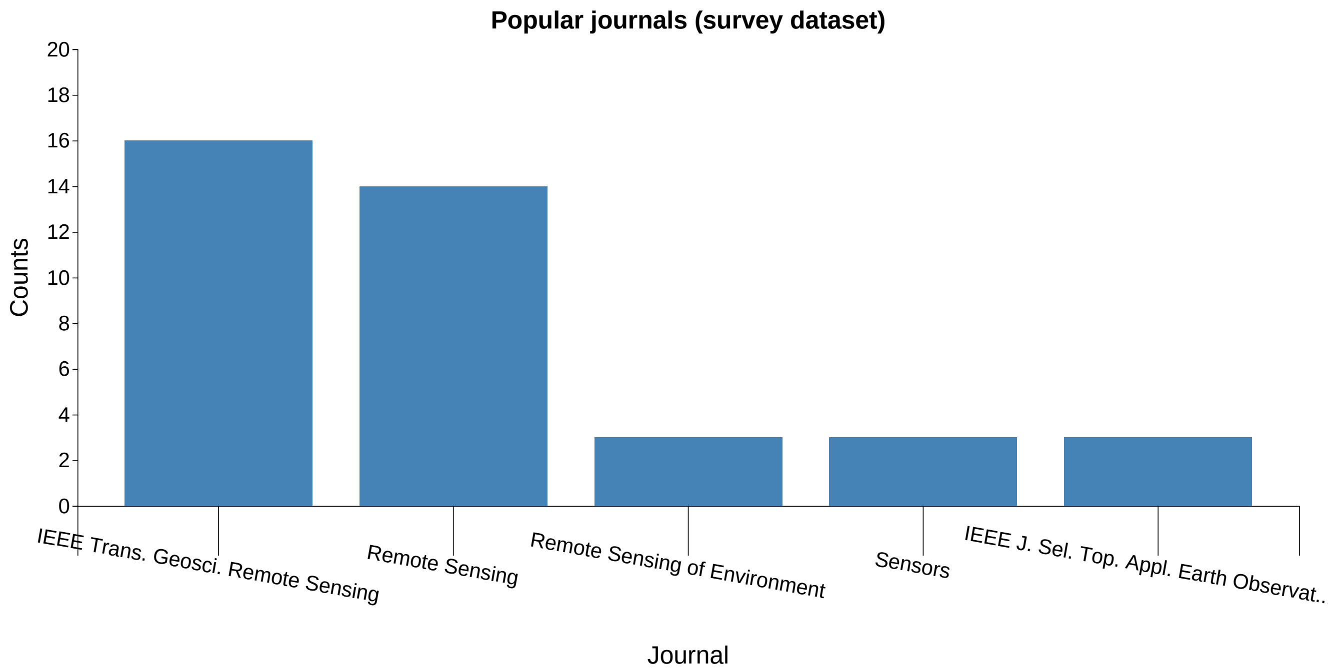

Figure 14: Number of times a paper has been published by a journal (not all journals are shown). "IEEE Trans. Geosci. Remote Sensing" from IEEE and "Remote Sensing" from MDPI are by far the most used journals, with 16 out of 81 papers being published in "IEEE Trans. Geosci. Remote Sensing" and 14 out of 81 papers being published in "Remote Sensing". The remaining journals have 3 or fewer papers in this dataset. This was made using the datset from the survey [WMS+26].

these tasks with only one or two appearances in the dataset, leading to contributions to more niche branches of AutoML for EO.

5.4 (RQ4)

In Section 3, it was mentioned that the scraper dataset contains 87 papers and the survey dataset contains 81 papers. Since the survey dataset was curated manually, the small number of papers is to be expected. The scraper, however, contained only 87 valid papers from the 1000 returned from either Google Scholar or Crossref, which is unexpected. This could be due to many reasons; the query could have been insufficient to find the correct papers, Google Scholar and Crossref may not contain a lot of papers about this topic, or there may have been a lot of false negatives when filtering the papers because of the difficulty in detecting the use of AutoML in some papers that do not state it explicitly. By looking at Figure 11 and 12 discussed in RQ2, we can see that the dataset from the survey [WMS+26] has way more papers from IEEE than the dataset from the scraper. This could be because of a bias in the papers that are curated for the dataset from the survey. Since the papers in the dataset from the survey were collected and read manually, there could be a bias papers with certain properties, like a specific kind of task, type, or, in the case of Figures 11 and 12, a publisher. Figure 14 and 15 also confirm this, where the dataset from the survey seems to have a bias towards papers from "IEEE Trans. Geosci. Remote Sensing" and "Remote Sensing", while the dataset from the scraper shows a more evenly distributed journal count, with

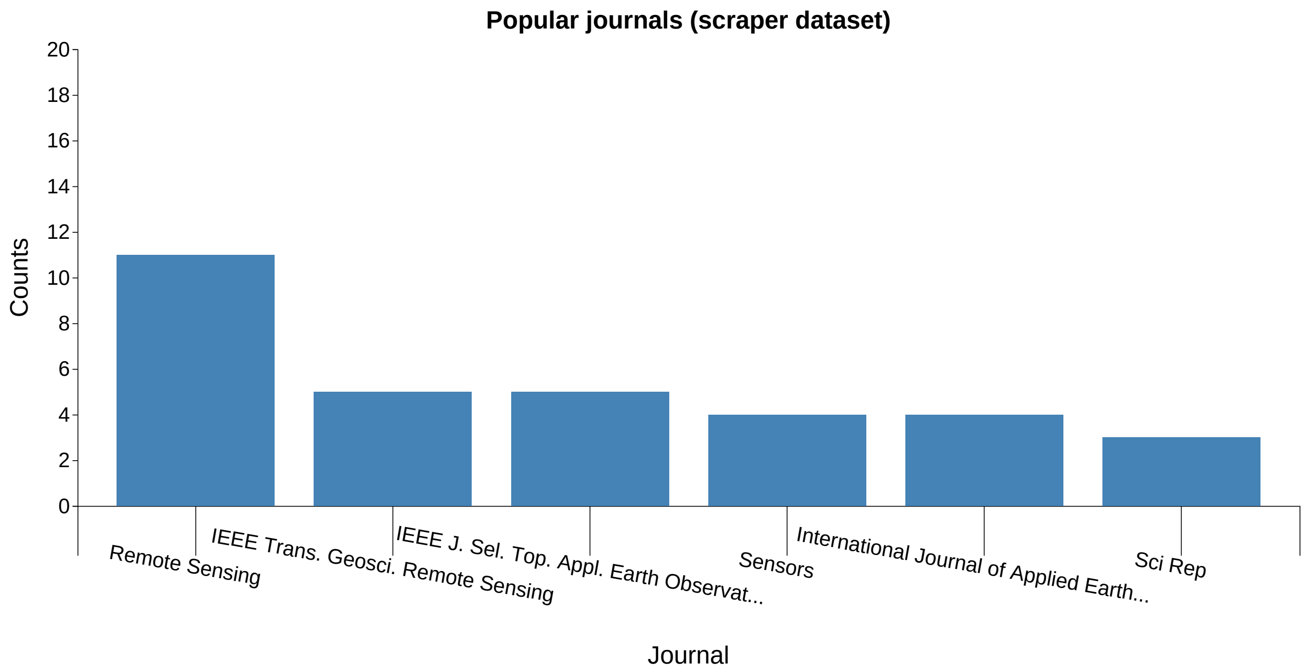

Figure 15: Number of times a paper has been published by a journal (not all journals are shown). “Remote Sensing” from MDPI is the most used journal, with 11 out of 81 papers being published in “Remote Sensing”. The remaining journals have 5 or fewer papers in this dataset. This was made using the dataset from the scraper.

only “Remote Sensing” having a high amount of papers compared to the other journals.

The dataset from the survey does, however, have the advantage of containing the substantive information, which the dataset from the scraper does not have. Because retrieving this information is difficult and takes a great amount of time, the dataset from the scraper does not contain this information and can thus not be used for most of the visualisations. The dataset from the survey does contain this information since every paper was manually read, and the information on the dataset, framework, and task was manually filled in.

Figure 16 and 17 show the difference in what years the papers of the datasets were published. We can see that the survey paper focuses on the year 2022 and decreases in paper amounts when going further in the past or future. The scraper dataset, however, focuses on the most recent papers from more recent years, containing more papers, except for 2025, since that is the current year, meaning the amount of papers published in 2025 is limited. This could explain why the scraper cannot find a lot of papers about this topic, since the papers may be difficult to find on Google Scholar or Crossref when they are older than 2023.

5.5 Limitations

As discussed a bit in the results already, this research encountered a few limitations that impacted the quality of the resulting visualisations. First, the datasets used to create these visualisations were relatively small since the scraper could not find many papers about this topic. From the 1000 papers found on Crossref, only 87 were about this topic. This could be due to the fact that the

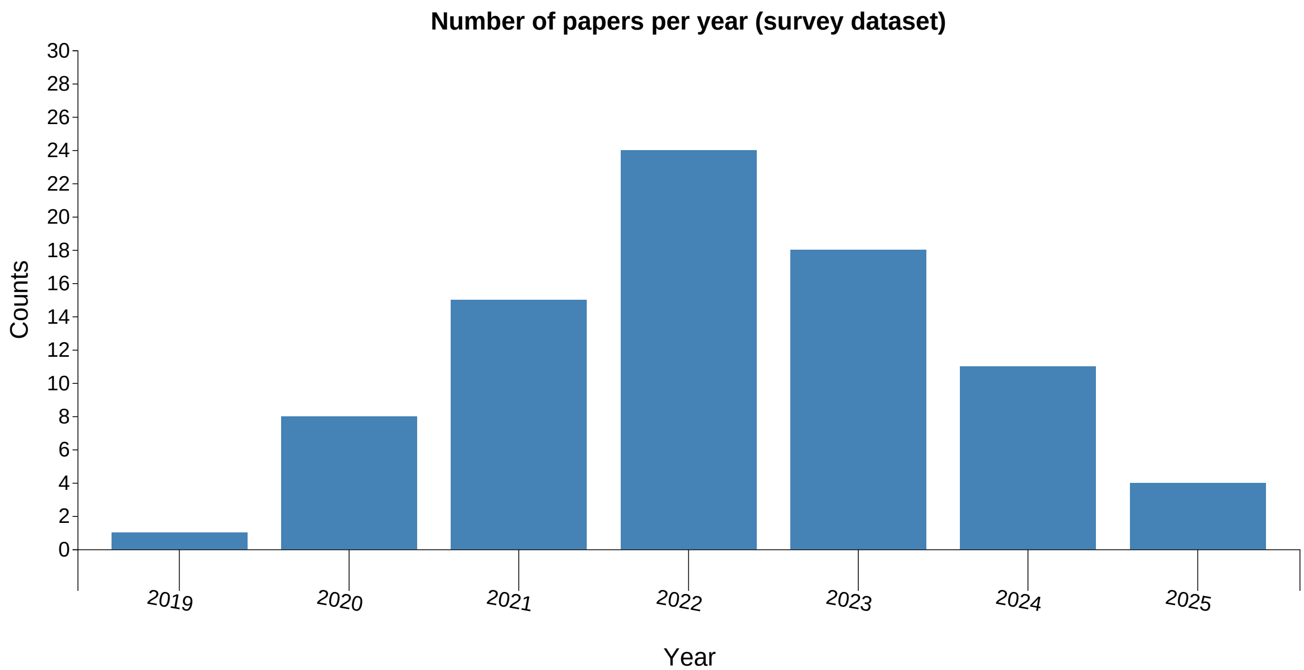

Figure 16: Number of papers from each year in the survey dataset. Most papers in this dataset were published in 2022. The farther away from 2022, the fewer papers are from that year.

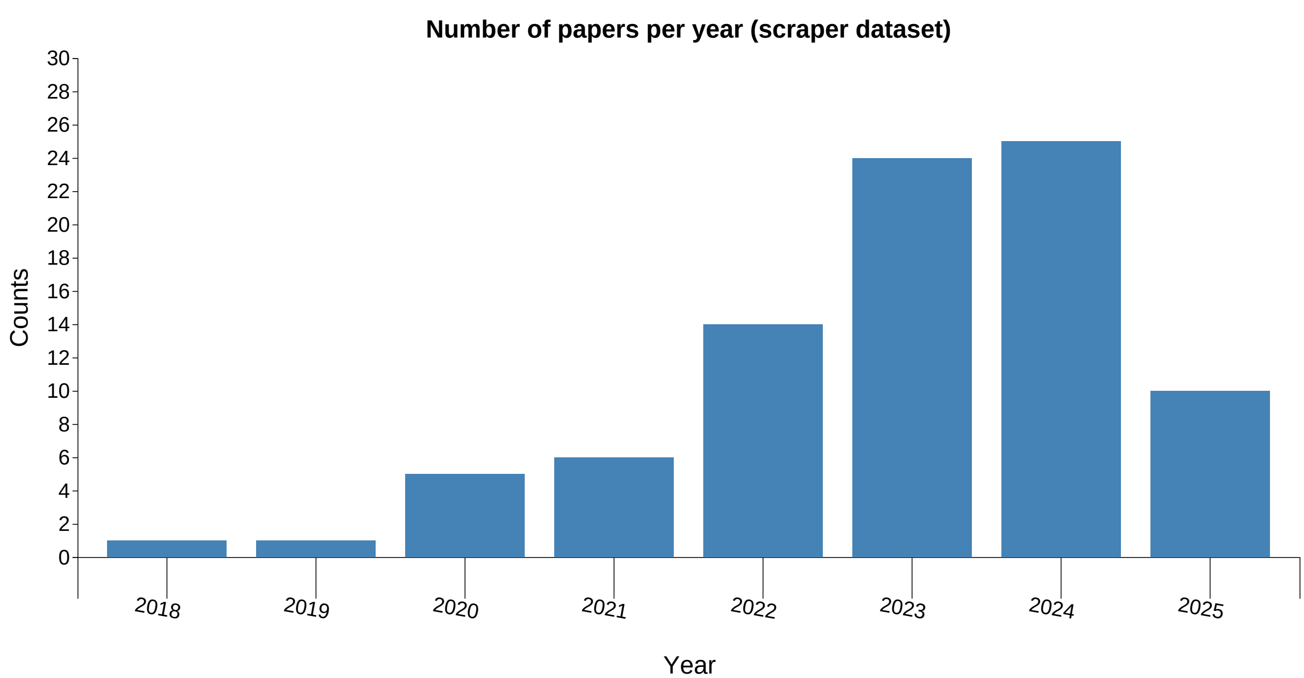

Figure 17: Number of papers from each year in the scraper dataset. Most papers in this dataset were from 2024. We can see that most of the papers the scraper finds are very recent.

scraper only searches on Google Scholar and Crossref, missing out on papers that may only show up on other sites. Because of the small datasets, it may be difficult to see trends.

Secondly, scraping certain information automatically is difficult. Because of this the amount of substantive information is limited. The papers have to be skimmed through manually to find that information, which is difficult to do with limited time and only a single person working on this thesis. This limited amount of substantive information causes a limited amount of possible useful visualisations to answer questions about this field.

6 Conclusion and Future Work

In this thesis, we presented visualisations made with metadata and substantive data of papers in the field of AutoML for EO. EO is a field that deals with large amounts of data. To analyse this, ML is used. AutoML automates the creation of ML models, and this, combined with EO, is a relatively new field that is quickly growing. When exploring the literature of a field, visualisations can be a great tool which can speed up and increase understanding with an increase in comprehension, decision speed, and confidence. Tools like Litmaps, Connected Papers, and ResearchRabbit do this by visualising the citation networks. However, they lack the more important, substantial information. Our visualisations fill this gap in the field of AutoML for EO. Questions about the use of technologies and what topics are researched can easily be answered by looking at a few of the visualisations. This makes it easy for people to see what tasks are understudied and thus still need to be researched, and what they could work on and with what frameworks and datasets. This would result in more time being spent researching and less time spent figuring out what can be researched and what technologies can be used. DARTS is the most popular framework, but since 2023, H2O has been promising. AutoSklearn has also been used a lot throughout the years 2021-2025 consistently. IEEE and MDPI are very popular journals for Image classification and Land use, while other tasks are more varied in journals. Image classification, Land use, and Yield prediction are the most popular tasks.

6.1 Future Work

The limitation of not being able to find many papers through Google Scholar and Crossref could be solved by making use of tools like Litmaps or Connected Papers, as the citations in papers in this field will probably contain quite a large number of other papers in this field. ResearchRabbit could also help greatly by using its recommendation system. The limitation of not having a very big dataset with a small amount of substantive information could be solved with the use of Large Language Models (LLMs). With the emergence of more and more capable LLMs, future work can research the use of such models to automate the finding of relevant papers, and potentially automate information retrieval from the contents of the papers. With this approach, it is important, however, to make sure the LLM does not hallucinate imaginary papers since we only want real papers in our datasets. We can see that many frameworks and tasks are underrepresented from our visualisations, so for future research, these could be explored more.

7 Code availability

All code used to retrieve papers and make visualisations can be found at:
https://github.com/konjiii/scraper (will be made public when the survey paper gets published [WMS+26], as a companion)
All visualisations made for this thesis can be found at:
https://scraper-autml-eo.vercel.app/

References

[AMCC20] Matthew D. Adams, Felix Massey, Karl Chastko, and Calvin Cupini. Spatial modelling of particulate matter air pollution sensor measurements collected by community scientists while cycling, land use regression with spatial cross-validation, and applications of machine learning for data correction. *Atmospheric Environment*, 230:117479, 2020.

[And11] Eduardo B. Andrade. Excessive confidence in visually-based estimates. *Organizational Behavior and Human Decision Processes*, 116(2):252–261, 2011.

[ARS+17] Katherine Anderson, Barbara Ryan, William Sonntag, Argyro Kavvada, and Lawrence Friedl and. Earth observation in service of the 2030 agenda for sustainable development. *Geo-spatial Information Science*, 20(2):77–96, 2017.

[ASE+10] Frédéric Achard, Hans-Jürgen Stibig, Hugh D Eva, Erik J Lindquist, Alexandre Bouvet, Olivier Arino, and Philippe Mayaux and. Estimating tropical deforestation from earth observation data. *Carbon Management*, 1(2):271–287, 2010.

[BJK23] Prashanta Kumar Behera, Sanmati Jinendran Jain, and Ashok Kumar. Visual exploration of literature using connected papers: A practical approach. *Issues in Science and Technology Librarianship*, (104), Oct. 2023.

[BL07] Marcus A. Butavicius and Michael D. Lee. An empirical evaluation of four data visualization techniques for displaying short news text similarities. *International Journal of Human-Computer Studies*, 65(11):931–944, 2007.

[BOH11] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.

[Bok18] Bokeh Development Team. *Bokeh: Python library for interactive visualization*, 2018.

[BYC+17] Mariusz Bojarski, Philip Yeres, Anna Choromanska, Krzysztof Choromanski, Bernhard Firner, Lawrence Jackel, and Urs Muller. Explaining how a deep neural network trained with end-to-end learning steers a car, 2017.

[Ca91] Paul Chandler and John Sweller and. Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4):293–332, 1991.

[CISK23] Steven A. Cholewiak, Panos Ipeirotis, Victor Silva, and Arun Kannawadi. Scholarly: Simple access to google scholar authors and citation using python, 2023.

[Con] Connectedpapers. (2024). [search tool]. https://www.connectedpapers.com/.

[CRK19] Daniel N. Cassenti, Heather Roy, and Sue E. Kase. Cognitive processing of visually presented data in decision making. *Human Factors*, 61(1):78–89, 2019. PMID: 30130125.

[Cro20] Crossref commons: High-level library for getting data from crossref apis, 2020.

[dHGBV08] Maria-Dolores de Hevia, Luisa Girelli, Emanuela Bricolo, and Giuseppe Vallar. The representational space of numerical magnitude: illusions of length. *Q J Exp Psychol (Hove)*, 61(10):1496–1514, oct 2008.

[DHGH16] Michael Dambacher, Peter Haffke, Daniel Groß, and Ronald Hübner. Graphs versus numbers: How information format affects risk aversion in gambling. *Judgment and Decision Making*, 11(3):223–242, 2016.

[dSBB+22] Nuno César de Sá, Mitra Baratchi, Vincent Buitenhuis, Perry Cornelissen, and Peter M. van Bodegom and. Automl for estimating grass height from etm+/oli data from field measurements at a nature reserve. *GIScience & Remote Sensing*, 59(1):2164–2183, 2022.

[Ebe23] Karin Eberhard. The effects of visualization on judgment and decision-making: a systematic literature review. *Management Review Quarterly*, 73(1):167–214, Feb 2023.

[EMH19] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. Neural architecture search: A survey. *Journal of Machine Learning Research*, 20(55):1–21, 2019.

[FDMS19] Yuning Feng, Shihong Du, Soe W. Myint, and Mi Shu. Do urban functional zones affect land surface temperature differently? a case study of beijing, china. *Remote Sensing*, 11(15), 2019.

[FEF+20] Matthias Feurer, Katharina Eggensperger, Stefan Falkner, Marius Lindauer, and Frank Hutter. Auto-sklearn 2.0: Hands-free automl via meta-learning. *arXiv:2007.04074 [cs.LG]*, 2020.

[FKE+15] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and robust automated machine learning. In *Advances in Neural Information Processing Systems 28 (2015)*, pages 2962–2970, 2015.

[GKDF23] Povilas Gudzius, Olga Kurasova, Vytenis Darulis, and Ernestas Filatovas. Automl-based neural architecture search for object recognition in satellite imagery. *Remote Sensing*, 15(1), 2023.

[GQF20] Vitor C. F. Gomes, Gilberto R. Queiroz, and Karine R. Ferreira. An overview of platforms for big earth observation data management and analysis. *Remote Sensing*, 12(8), 2020.

[Gé17] Jan Géryk. Visual analytics of educational time-dependent data using interactive dynamic visualization. *Expert Systems*, 34(1):e12175, 2017. e12175 EXSY-Nov-15-243.R1.

[H2O22] H2O.ai. *H2O: Scalable Machine Learning Platform*, 2022. version 3.42.0.2.

[HKV19] Frank Hutter, Lars Kotthoff, and Joaquin Vanschoren, editors. *Automatic machine learning*. The Springer Series on Challenges in Machine Learning. Springer International Publishing, Basel, Switzerland, 1 edition, March 2019.

[HSSC08] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[JSH19] Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 1946–1956, New York, NY, USA, 2019. Association for Computing Machinery.

[KAL+14] D. Kool, N. Agam, N. Lazarovitch, J.L. Heitman, T.J. Sauer, and A. Ben-Gal. A review of approaches for evapotranspiration partitioning. *Agricultural and Forest Meteorology*, 184:56–70, 2014.

[KBL+19] S O Kotchi, C Bouchard, A Ludwig, E E Rees, and S Brazeau. Using earth observation images to inform risk assessment and mapping of climate change-related infectious diseases. *Can. Commun. Dis. Rep.*, 45(5):133–142, may 2019.

[KGS+22] Amanpreet Kaur, Sarita Gulati, Ritu Sharma, Atasi Sinhababu, and Rupak Chakravarty. Visual citation navigation of open education resources using litmaps. *Library Hi Tech News*, 39(5):7–11, Jan 2022.

[KSM+22] Amanpreet Kaur, Ritu Sharma, Pooja Mishra, Atasi Sinhababu, and Rupak Chakravarty and. Visual research discovery using connected papers: A use case of blockchain in libraries. *The Serials Librarian*, 83(2):186–196, 2022.

[Lit] Litmaps. (2024). litmaps (version 2025-01-16) [search tool]. https://app.litmaps.com/.

[LSS25] Cédric Léonard, Dirk Stober, and Martin Schulz. Fpga-enabled machine learning applications in earth observation: A systematic review, 2025.

[LSY19] Hanxiao Liu, Karen Simonyan, and Yiming Yang. Darts: Differentiable architecture search, 2019.

[LZL+18] David J. Lary, Gebreab K. Zewdie, Xun Liu, Daji Wu, Estelle Levetin, Rebecca J. Allee, Nabin Malakar, Annette Walker, Hamse Mussa, Antonio Mannino, and Dirk Aurin. *Machine Learning Applications for Earth Observation*, pages 165–218. Springer International Publishing, Cham, 2018.

[Moo17] Jeanne Moore. *Visualisation of data to optimise strategic decision making*. PhD thesis, 2017.

[MRA+17] M. F. McCabe, M. Rodell, D. E. Alsdorf, D. G. Miralles, R. Uijlenhoet, W. Wagner, A. Lucieer, R. Houborg, N. E. C. Verhoest, T. E. Franz, J. Shi, H. Gao, and E. F. Wood. The future of earth observation in hydrology. *Hydrology and Earth System Sciences*, 21(7):3879–3914, 2017.

[OSBdB18] Yasmina Okan, Eric R. Stone, and Wändi Bruine de Bruin. Designing graphs that promote both risk understanding and behavior change. *Risk Analysis*, 38(5):929–946, 2018.

[PCS+18] Michelle Cristina Araujo Picoli, Gilberto Camara, Ieda Sanches, Rolf Sim oes, Alexandre Carvalho, Adeline Maciel, Alexandre Coutinho, Julio Esquerdo, Jo ao Antunes, Rodrigo Anzolin Begotti, Damien Arvor, and Claudio Almeida. Big earth observation time series analysis for monitoring brazilian agriculture. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145:328–339, 2018. SI: Latin America Issue.

[PKD+13] Mark S. Pfaff, Gary L. Klein, Jill L. Drury, Sung Pil Moon, Yikun Liu, and Steven O. Entezari. Supporting complex decision making through option awareness. *Journal of Cognitive Engineering and Decision Making*, 7(2):155–178, 2013.

[PRKG20] Mihir Prakash, Steven Ramage, Argyro Kavvada, and Seth Goodman. Open earth observations for sustainable urban development. *Remote Sensing*, 12(10), 2020.

[PSBvRV21] Nelly Rosaura Palacios Salinas, Mitra Baratchi, Jan N. van Rijn, and Andreas Vollrath. Automated machine learning for satellite data: Integrating remote sensing pre-trained models into automl systems. In Yuxiao Dong, Nicolas Kourtellis, Barbara Hammer, and Jose A. Lozano, editors, *Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track*, pages 447–462, Cham, 2021. Springer International Publishing.

[RDGF16] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016.

[RDK+22] David Rolnick, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, Alexandra Sasha Luccioni, Tegan Maharaj, Evan D. Sherwin, S. Karthik Mukkavilli, Konrad P. Kording, Carla P. Gomes, Andrew Y. Ng, Demis Hassabis, John C. Platt, Felix Creutzig, Jennifer Chayes, and Yoshua Bengio. Tackling climate change with machine learning. *ACM Comput. Surv.*, 55(2), February 2022.

[Res] Researchrabbit. (2025). [search tool]. https://www.researchrabbit.ai/.

[Sa02] Carolyn Semmler and Neil Brewer and. Using a flow-chart to improve comprehension of jury instructions. *Psychiatry, Psychology and Law*, 9(2):262–270, 2002.

[SLBL16] Yan Sun, Shu Li, Nicolao Bonini, and Yang Liu. Effect of graph scale on risky choice: Evidence from preference and process in decision-making. 2016.

[SMK+10] Chris M. R. Smerecnik, Ilse Mesters, Loes T. E. Kessels, Robert A. C. Ruiter, Nanne K. De Vries, and Hein De Vries. Understanding the positive effects of graphical risk information on comprehension: Measuring attention directed to written, tabular, and graphical risk information. *Risk Analysis*, 30(9):1387–1398, 2010.

[SW21] Yongze Song and Peng Wu. Earth observation for sustainable infrastructure: A review. *Remote Sensing*, 13(8), 2021.

[TAR+19] Lukas Tuggener, Mohammadreza Amirian, Katharina Rombach, Stefan Lörwald, Anastasia Varlet, Christian Westermann, and Thilo Stadelmann. Automated machine learning in practice: State of the art and recent results. In *2019 6th Swiss Conference on Data Science (SDS)*, pages 31–36, 2019.

[TSD+25] Devis Tuia, Konrad Schindler, Begüm Demir, Xiao Xiang Zhu, Mrinalini Kochupillai, Sašo Džeroski, Jan N. van Rijn, Holger H. Hoos, Fabio Del Frate, Mihai Datcu, Volker Markl, Bertrand Le Saux, Rochelle Schneider, and Gustau Camps-Valls. Artificial intelligence to advance earth observation: A review of models, recent trends, and pathways forward. *IEEE Geoscience and Remote Sensing Magazine*, page 2–25, 2025.

[TWG+19] Anh Truong, Austin Walters, Jeremy Goodsitt, Keegan Hines, C. Bayan Bruss, and Reza Farivar. Towards automated machine learning: Evaluation and comparison of automl approaches and tools. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1471–1479, 2019.

[vW05] J.J. van Wijk. The value of visualization. In *VIS 05. IEEE Visualization, 2005.*, pages 79–86, 2005.

[WMA+24] Julia Wąsala, S. Marselis, Laurens Arp, Holger H. Hoos, Nicolas Longépé, and Mitra Baratchi. Autosr4eo: An automl approach to super-resolution for earth observation images. *Remote Sensing*, 2024.

[WMS+26] Julia Wąsala, Joannes D. Maasakkers, Berend J. Schuit, Gijs Leguijt, Ilse Aben, Rochelle Schneider, Holger Hoos, and Mitra Baratchi. AutoML to advance earth observation. *Manuscript currently to be submitted for review.*, 2026.

[ZYD+22] Qiang Zhao, Le Yu, Zhenrong Du, Dailiang Peng, Pengyu Hao, Yongguang Zhang, and Peng Gong. An overview of the applications of earth observation satellite data: Impacts and future trends. *Remote Sensing*, 14(8), 2022.

[ZZY+25] Bo Zhang, Li Zhang, Min Yan, Jian Zuo, Yuqi Dong, and Bowei Chen. High-resolution mapping of forest parameters in tropical rainforests through automl integration of gedi with sentinel-1/2, landsat 8, and alos-2 data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18:9084–9118, 2025.