

# **PLANNING**

## **ESCUELA POLITÉCNICA NACIONAL**



David Fabián Cevallos Salas

José Antonio Estrada Jiménez

Danny Santiago Guamán Loachamín

# CONTENT INDEX

<b>1</b>	<b>GOAL, RESEARCH QUESTIONS AND SEARCH STRING .....</b>	<b>4</b>
1.1	Topic .....	4
1.2	Goal.....	4
1.3	Research Questions .....	4
1.4	Search string.....	5
<b>2</b>	<b>CRITERIA FOR AUTOMATED INCLUSION – EXCLUSION PROCEDURE .....</b>	<b>7</b>
2.1	Criteria for Inclusion.....	7
2.2	Criteria for Exclusion.....	8
<b>3</b>	<b>CRITERIA AND PROCEDURE FOR MANUAL INCLUSION - EXCLUSION .....</b>	<b>8</b>
3.1	General.....	8
3.1.1	Goal.....	8
3.1.2	General instructions .....	9
3.1.3	Inclusion criteria .....	9
3.1.4	Exclusion criteria .....	9
3.1.5	Possible mark options .....	10
3.2	Screening stage .....	10
3.2.1	Pilot Phase.....	10
3.2.2	Main Phase.....	10
3.3	Procedure for screening each paper .....	11
3.4	Decision trees.....	12
<b>4</b>	<b>CODEBOOK FOR INFORMATION EXTRACTION AND CODING .....</b>	<b>15</b>
4.1	Research Questions .....	15
4.2	Data to be extracted .....	15
4.3	Classification for coding.....	16
4.3.1	Heuristic Techniques.....	18
4.3.2	Levels of Assessment .....	19
4.3.3	Levels of Automation.....	19
4.3.4	Types of Protection Mechanisms .....	20
4.3.5	Broad Domains .....	21
4.4	Coding procedure.....	21
4.4.1	Setting up necessary materials .....	22
4.4.2	Coding execution .....	22
4.4.3	Meeting of the coders to review results .....	22
4.5	General Coding validation .....	23

## FIGURES INDEX

<b>Figure 1</b> First search scope.....	6
<b>Figure 2</b> Second search scope .....	7
<b>Figure 3</b> Third search scope .....	7
<b>Figure 4</b> Decision tree Pilot Phase .....	13
<b>Figure 5</b> Decision tree Main Phase .....	14
<b>Figure 6</b> Classification Scheme.....	17

## TABLES INDEX

<b>Table 1</b> Main keywords .....	5
<b>Table 2</b> Number of results applied by keyword combinations .....	6
<b>Table 3</b> Automatic Information to be extracted.....	15
<b>Table 4</b> Fields for extracting manual information .....	16
<b>Table 5</b> Generalities of Heuristic Techniques categories.....	18
<b>Table 6</b> Descriptions of Heuristic Techniques categories.....	18
<b>Table 7</b> Generalities of Levels of Assessment categories.....	19
<b>Table 8</b> Descriptions of Levels of Assessment categories.....	19
<b>Table 9</b> Generalities of Levels of Automation categories.....	19
<b>Table 10</b> Descriptions of Levels of Automation categories.....	20
<b>Table 11</b> Generalities of Types of Protection Mechanisms categories .....	20
<b>Table 12</b> Descriptions of Types of Protection Mechanisms categories .....	20
<b>Table 13</b> Generalities of Broad Domains categories .....	21
<b>Table 14</b> Descriptions of Broad Domains categories .....	21

# PLANNING

## 1 GOAL, RESEARCH QUESTIONS AND SEARCH STRING

### 1.1 Topic

HEURISTIC TECHNIQUES FOR ASSESSING INTERNET PRIVACY: A  
COMPREHENSIVE REVIEW AND ANALYSIS

### 1.2 Goal

To identify the heuristic techniques used for assessing Internet privacy and its related protection mechanisms.

### 1.3 Research Questions

The following research questions have been established:

**RQ1:** What are the main heuristic techniques and their associated metrics used for assessing Internet privacy?

#### **Motivation**

Internet Privacy is a subjective term open to various interpretations. To assess Internet privacy effectively, each heuristic technique employs a metric that supports subsequent analysis, interpretation, and linkage to a level of Internet privacy. Consequently, a variety of heuristic techniques relying on diverse metrics can be applied.

**RQ2:** At what level is Internet privacy assessed through the main heuristic techniques?

#### **Motivation**

The various heuristic techniques and their associated metrics used for assessing Internet privacy have limitations when applied in practical scenarios. Some metrics offer a criterion of order and allow performing comparisons, while others require additional criteria to establish a connection with Internet privacy.

**RQ3:** At what level of automation is Internet privacy assessed through the main heuristic techniques?

## Motivation

Assessing Internet privacy involves collecting and manipulating information using various tools and solutions. Depending on the technologies employed and their level of automation, each heuristic technique could be executed either automatically, semi-automatically, or manually.

**RQ4:** What types of protection mechanisms have been applied in order to address the lack of privacy when using services over the Internet?

## Motivation

Numerous services are provided to users over the Internet, and the lack of privacy is a significant concern. Therefore, it is crucial to understand the types of protection mechanisms that have been implemented to address the absence of privacy when using Internet services.

**RQ5:** In what broad domains are heuristic techniques employed for the assessment of Internet privacy?

## Motivation

The Internet encompasses a wide range of applications and services, not solely limited to the web scenario. It also extends to various broad domains, including Mobile and Industrial Control Systems. Therefore, this review has been carried out in a domain-agnostic manner in order to link each contribution to its general broad domain.

## 1.4 Search string

In order to establish the search string, four main keywords have been identified:





- Assessment
- Internet Privacy
- Heuristic technique
- Protection mechanism

Keyword	Related terms
Assessment	Measure
	Metric
Internet Privacy	Online privacy
Heuristic Technique	Method
	Methodology
	Procedure
	Strategy
Protection Mechanism	Control
	Countermeasure
	Defense mechanism
	Privacy-enhancing technologies (PETs)

**Table 1** Main keywords

The following search string has been defined:

(**assess\*** OR **measure\*** OR **metric**)  
**AND** ((**Internet** W/3 **privacy**) OR (**online** **privacy**))  
**AND** ((**heuristic** W/3 **technique**) OR (**method\***) OR (**proced\***) OR (**strateg\***)  
OR (**protection** W/3 **mechanism**) OR (**control**) OR (**countermeasure**) OR (**defen\*** W/3  
**mechanism**) OR (**PET\***))

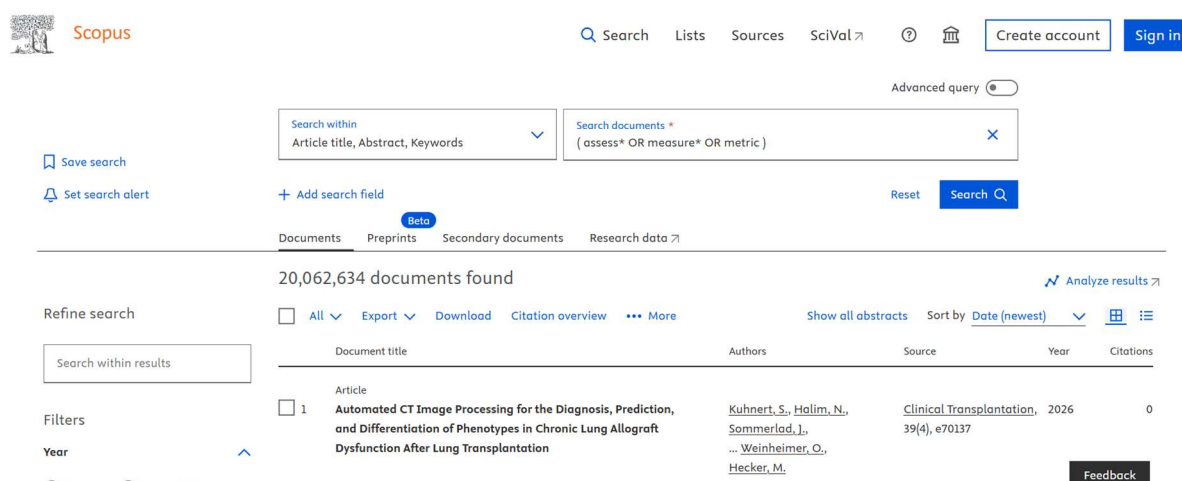
	Assessment
	Internet privacy
	Heuristic technique
	Protection mechanism

The following results were obtained, for example, on June 1<sup>st</sup> 2025 from the Scopus database, using the search string defined:

Keywords	Number of results
Assessment	20.062.634
Assessment + Internet privacy	5.512
Assessment + Internet privacy + Heuristic Technique + Protection Mechanism	3.121


**Table 2** Number of results applied by keyword combinations

Figure 1, Figure 2, and Figure 3 exposes the number of results obtained after the execution of each search.



The screenshot displays the Scopus search interface. At the top, the Scopus logo is on the left, and navigation links (Search, Lists, Sources, SciVal) and user options (Create account, Sign in) are on the right. The search bar contains the query '(assess\* OR measure\* OR metric)'. Below the search bar, there are options to 'Save search', 'Set search alert', and 'Add search field'. The search results section shows '20,062,634 documents found'. On the left, there are filters for 'Refine search' and 'Filters' (Year, Range, Individual). The main results table has columns for Document title, Authors, Source, Year, and Citations. The first result is an article titled 'Automated CT Image Processing for the Diagnosis, Prediction, and Differentiation of Phenotypes in Chronic Lung Allograft Dysfunction After Lung Transplantation' by Kuhnert, S., Halim, N., Sommerlad, J., Weinheimer, O., and Hecker, M., published in Clinical Transplantation, 2026, 39(4), e70137.

**Figure 1** First search scope



Scopus

Search Lists Sources SciVal ? Create account Sign in

Advanced query

Search within: Article title, Abstract, Keywords

Search documents \*: ( assess\* OR measure\* OR metric ) AND ( ( Internet W/ 3 privacy ) OR ( online priv

+ Add search field

Reset Search

Documents Preprints Secondary documents Research data

5,512 documents found

Analyze results

Refine search

Search within results

Filters

Year


Document title Authors Source Year Citations

1 Sentimental analysis based federated learning privacy detection in fake web recommendations using blockchain model Samriya, J.K., Kumar, A., Bhansali, A., ... Alsulami, B.S., Gupta, B.B. Scientific Reports, 15(1), 13551 2025 0

Show abstract View at Publisher Related documents

Feedback

Figure 2 Second search scope



Scopus

Search Lists Sources SciVal ? Create account Sign in

Advanced query

Search within: Article title, Abstract, Keywords

Search documents \*: ( assess\* OR measure\* OR metric ) AND ( ( Internet W/ 3 privacy ) OR ( online priv

+ Add search field

Reset Search

Documents Preprints Secondary documents Research data

3,121 documents found

Analyze results

Refine search

Search within results

Filters

Year

Document title Authors Source Year Citations

1 Sentimental analysis based federated learning privacy detection in fake web recommendations using blockchain model Samriya, J.K., Kumar, A., Bhansali, A., ... Alsulami, B.S., Gupta, B.B. Scientific Reports, 15(1), 13551 2025 0

Show abstract View at Publisher Related documents

Feedback

Figure 3 Third search scope

## 2 CRITERIA FOR AUTOMATED INCLUSION – EXCLUSION PROCEDURE

The next points expose the automated inclusion and exclusion criteria that will be applied using Scopus, Web of Science and IEEE Xplore search engines.

### 2.1 Criteria for Inclusion

Due to the nature of the research in an engineering field, the next inclusion criteria have been established:

- **Language:** English.

Several papers and academic work written in English have been demonstrated to be of relevance in the fields of Information Security and Privacy.

- **Document type:** Conference Paper or Journal.

Conference paper will allow for determining the new pragmatic solutions in the fields, as long as a journal paper will contribute with means for analyzing pragmatic and theoretical solutions.

- **Research field:** Computer Science.

Research topics in a strong relationship with the field of Computer Science will be considered.

- **Publication stage:** Final

Consider just manuscripts at the final publication stage, that is, after passing at least initial quality reviews.

## 2.2 Criteria for Exclusion

The exclusion criteria are deduced directly from the inclusion criteria.

- **Language:** Papers not written in English nor Spanish will be discarded.
- **Document type:** If the material falls under a category different of Conference paper or Journal will be discarded. The book chapters found will be kept in order to acquire knowledge.
- **Research field:** Overall, if the field of knowledge is not Computer Science the paper must be discarded.

## 3 CRITERIA AND PROCEDURE FOR MANUAL INCLUSION - EXCLUSION

### 3.1 General

#### 3.1.1 Goal



To discard the papers obtained from the search phase that are not suitable for the purpose of the research.

### **3.1.2 General instructions**

- The research group will be composed of the main researcher and the director. Possibly, the research group will be completed with one codirector.
- A pilot with a limit of 1% of the papers resulted from the search phased will be constituted. The inclusion and exclusion criteria will be applied over the pilot and the following iterations.
- The Krippendorff's Alpha index [2] will be used in order to establish the level of consensus between the research group members. An index of at least 0,85 has been establish as a minimum threshold for consensus between the opinions.

### **3.1.3 Inclusion criteria**

The papers must comply with the following criteria:

- The paper is a primary contribution. Secondary contributions must be kept for analysis and acquire new ideas, but will be discarded.
  - Keep secondary contributions in order to acquire general knowledge that could be helpful in the process.
- The paper presents a solution related directly to Internet privacy or that could be adapted to Internet privacy.
  - Work in other contexts of privacy will be ruled out.
- The paper contribution is related with a technique and/or a protection mechanism able to assess and/or protect Internet privacy.
- A proof that allows to corroborate the technique or protection mechanism used by the solution presented must be clearly identified in the paper.

**Note:** At this stage it will not be considered as an inclusion/exclusion criterion the percentile rank of the Thomson Reuters Indicators due to the possible reduced number of papers to obtain. However, if in the execution phase is considered necessary, this criterion also will be applied according to the results obtained.

### **3.1.4 Exclusion criteria**

Based on the inclusion criteria, the following exclusion criteria have been defined:

- The paper is a secondary or tertiary contribution. But remember to keep and analyze the secondary contributions.
- The paper does not present a solution related to Internet privacy and cannot be adapted to this context.
- The paper is not related with a technique or a protection mechanism able to assess and/or protect Internet privacy.

- The paper does not present a proof of validation of the technique or protection mechanism used.

### 3.1.5 Possible mark options

**[I]** The paper fulfills the inclusion criteria and none of the exclusion criteria.

**[S]** The paper fulfills the inclusion criteria but it is a secondary contribution. The paper is kept but is discarded.

**[E]** The paper does not fulfill with at least one inclusion criteria or fulfills with at least one exclusion criteria. So, the paper is discarded.

**[U]** A dilemma has been created of the paper analysis. So, it must be treated and discussed in depth later by the research group.

## 3.2 Screening stage

In order to get the main papers for the research and achieve consensus between the researcher group members, the screening stage has been divided in two phases: a pilot phase and the main phase.

### 3.2.1 Pilot Phase

In the pilot phase a sample of 5% the papers will be chosen randomly. Each of the papers will be analyzed, through the inclusion and exclusion criteria, foreach of the research group members.

For this phase, just the **Title, Abstract and Full Paper** will be read.

After an iteration, a Krippendorff's Alpha index of at least **0,85** must be achieved to establish a line of consensus between the members.

If the index is not achieved in the first iteration, the process will be repeated through several discussions and pilots between the research group member until achieve the desired index.

The pilot phase main goal is to unify the research criteria between the member of the research group.

### 3.2.2 Main Phase

The remained papers that not conformed the pilot phase will be considered in the main phase. This phase has as goal to obtain the final set of the papers for the research.

The main phase establishes the following procedure:

- **Step 1:** Firstly, each member of the research group will read just the **Title** of a group of papers individually. Under his or her understanding will classify each paper as **Included**, **Excluded** or **Unclear**. In this step the Krippendorff's Alpha index does not have to be calculated.

If the paper is a secondary locate it as included in this first step.

- **Step 2:** Secondly, each member of the research group will read individually the **Title and Abstract** of the paper classified by himself or herself as Included or Unclear in the first step. Each of these analyzed papers will be then classified by the researcher as **Included**, **Excluded** or **Unclear** category. Again, in this step the Krippendorff's Alpha index does not have to be calculated.

If the paper is a secondary locate it as included in this second step.

- **Step 3:** Finally, each member of the research group will read individually the **Full Paper** that will be classified as Included or Unclear in the second step.

Under his or her understanding, the researcher will classify each paper in the **Included**, **Excluded** category.

If the paper presents a secondary contribution fulfilling all the other inclusion criteria, it will be located in the **Secondary** group for review in order to acquire knowledge but it will be discarded.

In this step is important that the researcher takes a final decision due to the lack of subsequent iterations. Any discrepancies will be discussed and analyzed in work meetings.

### 3.3 Procedure for screening each paper

Depending of the phase (pilot or main) each defined section of the paper must be read carefully.

**Step 1:** While reading is important to take into consideration the following terms:

- **Assessment** and its related terms: Measure, and metric.
- **Internet Privacy** and its related term: Online privacy.
- **Heuristic technique** and its related terms: Method, methodology, procedure, strategy.
- **Protection mechanism** and its related terms: Control, countermeasure, defense mechanism.

These terms have been taken from the IEEE Thesaurus [1].

## Step 2: For Title, Abstract and Full Paper based screening

Answer according to your criteria each question:

**Question 1:** Is the paper a primary contribution?

- If **no**, mark the paper as **Excluded** and go to the next paper.
- If **yes** or **unclear**, being a primary contribution, continue to the following question.
- If the paper is a **Secondary** work mark it as **no** and continue with the following question.

**Question 2:** Does the paper present a solution related to Internet privacy or one that could be adapted to Internet privacy?

- If **no**, mark the paper as **Excluded** and go to the next paper.
- If **yes** or **unclear**, continue to the following question.

**Question 3:** Is the paper contribution related to a technique and/or a protection mechanism able to assess or protect Internet privacy?

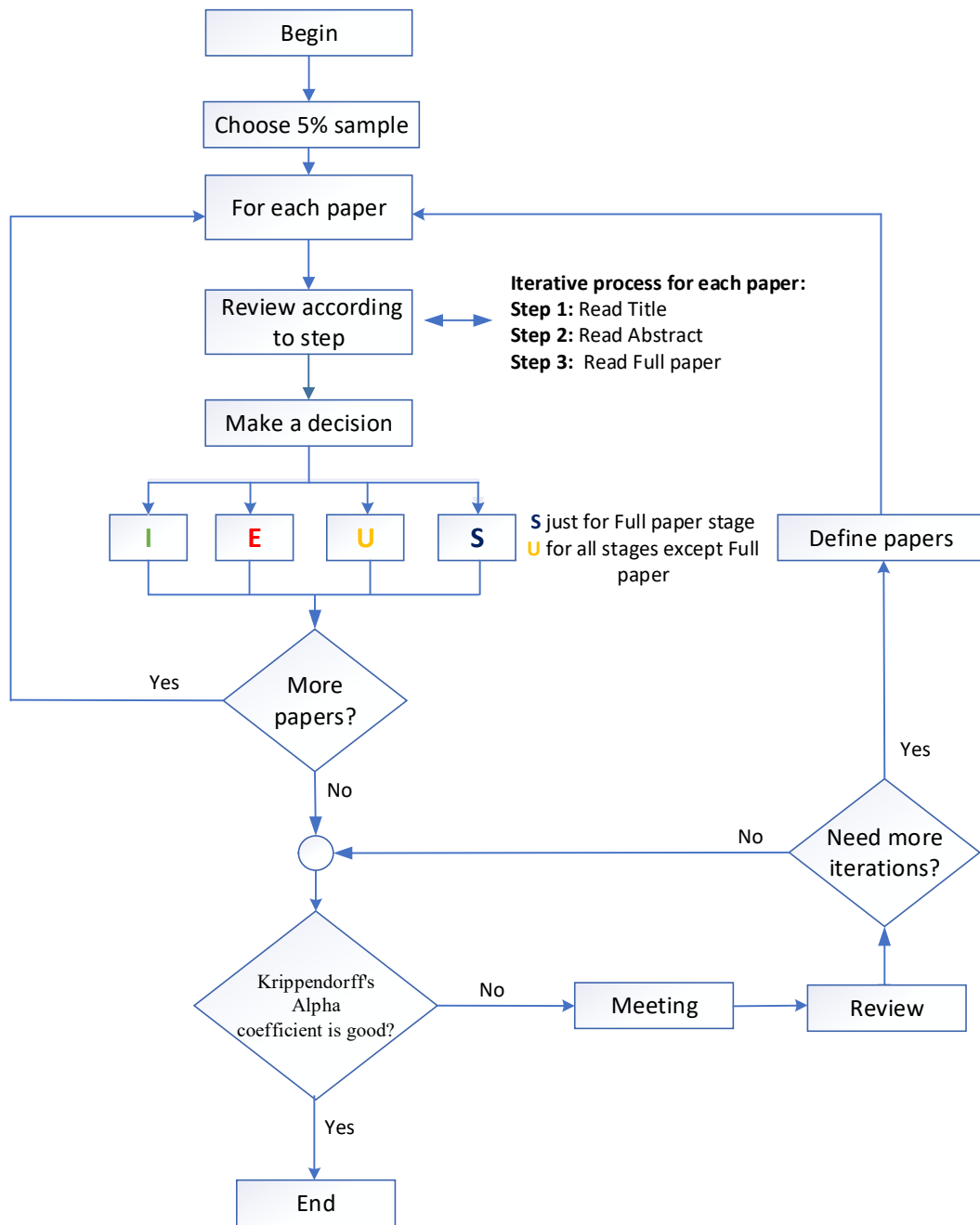
- If **no**, mark the paper as **Excluded** and go to the next paper.
- If **yes** or **unclear**, continue to the following question.

**Question 4:** Does the paper provide a proof that helps corroborate the proposed solution?

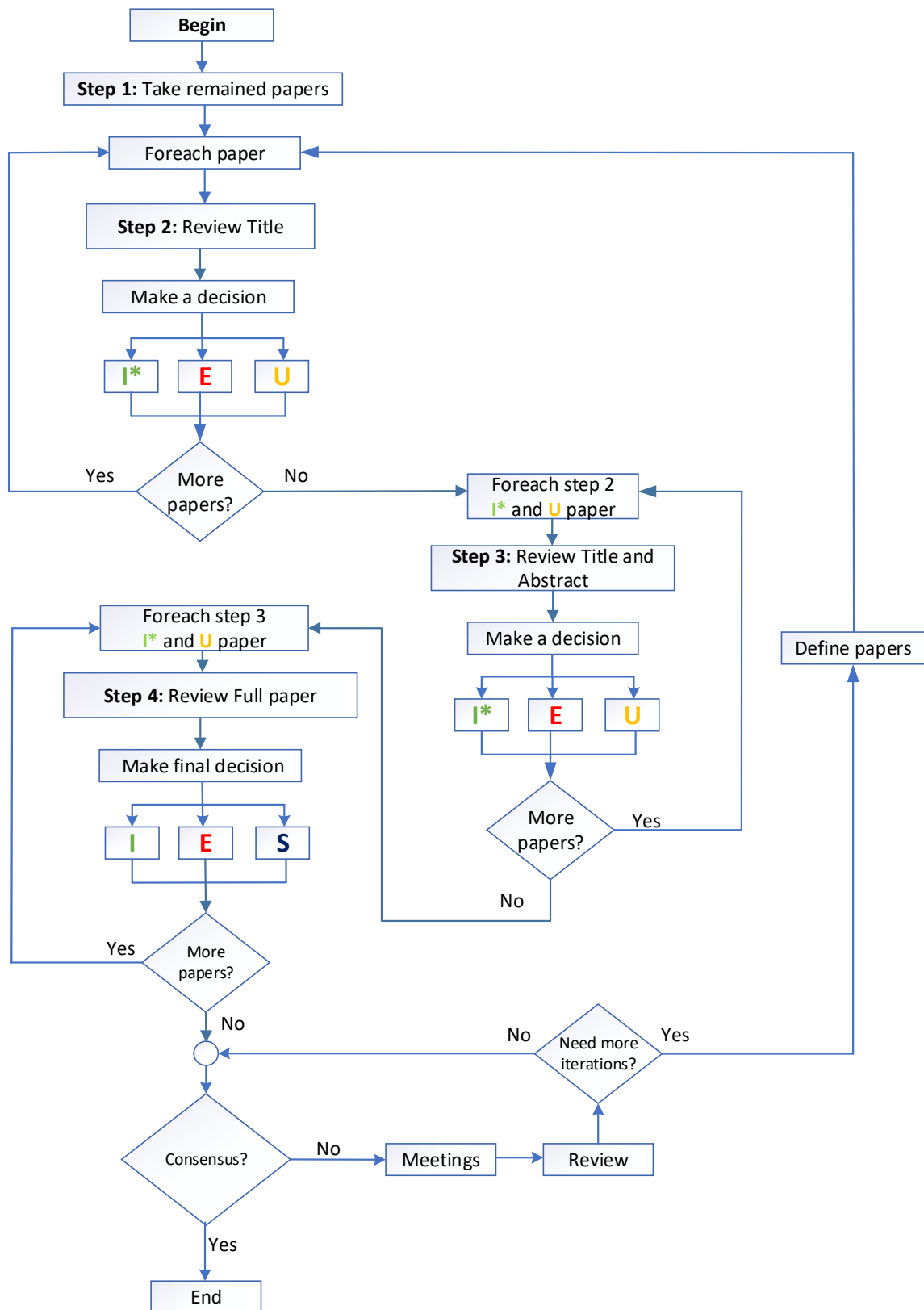
- If **yes** but the contribution is secondary and you are in the Full paper step, mark it as **Secondary** in order to review it and acquire new knowledge, but the paper is discarded for the extraction and coding process. If this is not the full paper stage, choose **yes** and continue with the rest of stages.
- If **yes** and the contribution is primary, there are no unclear answers to previous questions and this is the final Full paper stage, mark the paper as **Included**.
- If **yes** and the contribution is primary but there are unclear answers to previous questions mark it as **Unclear** if you are in the Title or Abstract step. If this is the final Full paper step organize meetings to resolve the unclear status of each question and take a final decision to include or discard the paper.
- If **no**, mark the paper as **Excluded** and go to the next paper.

## 3.4 Decision trees

The following pictures depicts the decision trees procedures.



**Figure 4** Decision tree Pilot Phase



I\* Include secondary contributions

**Figure 5** Decision tree Main Phase

## 4 CODEBOOK FOR INFORMATION EXTRACTION AND CODING

This document has as goal to present the information to be extracted from the papers after the application of the inclusion and exclusion criteria, the procedure for extracting that information and the classification scheme to be used to achieved this goal.

### 4.1 Research Questions

The following research questions, which guide the process, were defined:

- **RQ1 – Heuristic Techniques:** What are the main heuristic techniques and their associated metrics used for assessing Internet privacy?
- **RQ2 – Assessment Level:** At what level is Internet privacy assessed through the main heuristic techniques?
- **RQ3 – Automatization Level:** At what level of automation is Internet privacy assessed through the main heuristic techniques?
- **RQ4 – Types of Protection Mechanisms:** What types of protection mechanisms have been applied in order to address the lack of privacy when using services over the Internet?
- **RQ5 – Broad Domains:** In what broad domains are heuristic techniques employed for the assessment of Internet privacy?

### 4.2 Data to be extracted

In order to respond to the established research questions, automatic and manual information will be extracted from the papers. Automatic information will be extracted according to the fields explained in Table 3.

**Table 3** Automatic Information to be extracted

	Field	Detail
	Authors	Authors names
<b>General Information</b>	ID	Identifier
	Year	Year of publication
	Citation key	Key for citing the paper
	Source	Publisher(s) of the paper
	Journal or Conference name	Name of the journal or conference if it applies

Information will be extracted manually according to the details shown in Table 4.

**Table 4** Fields for extracting manual information

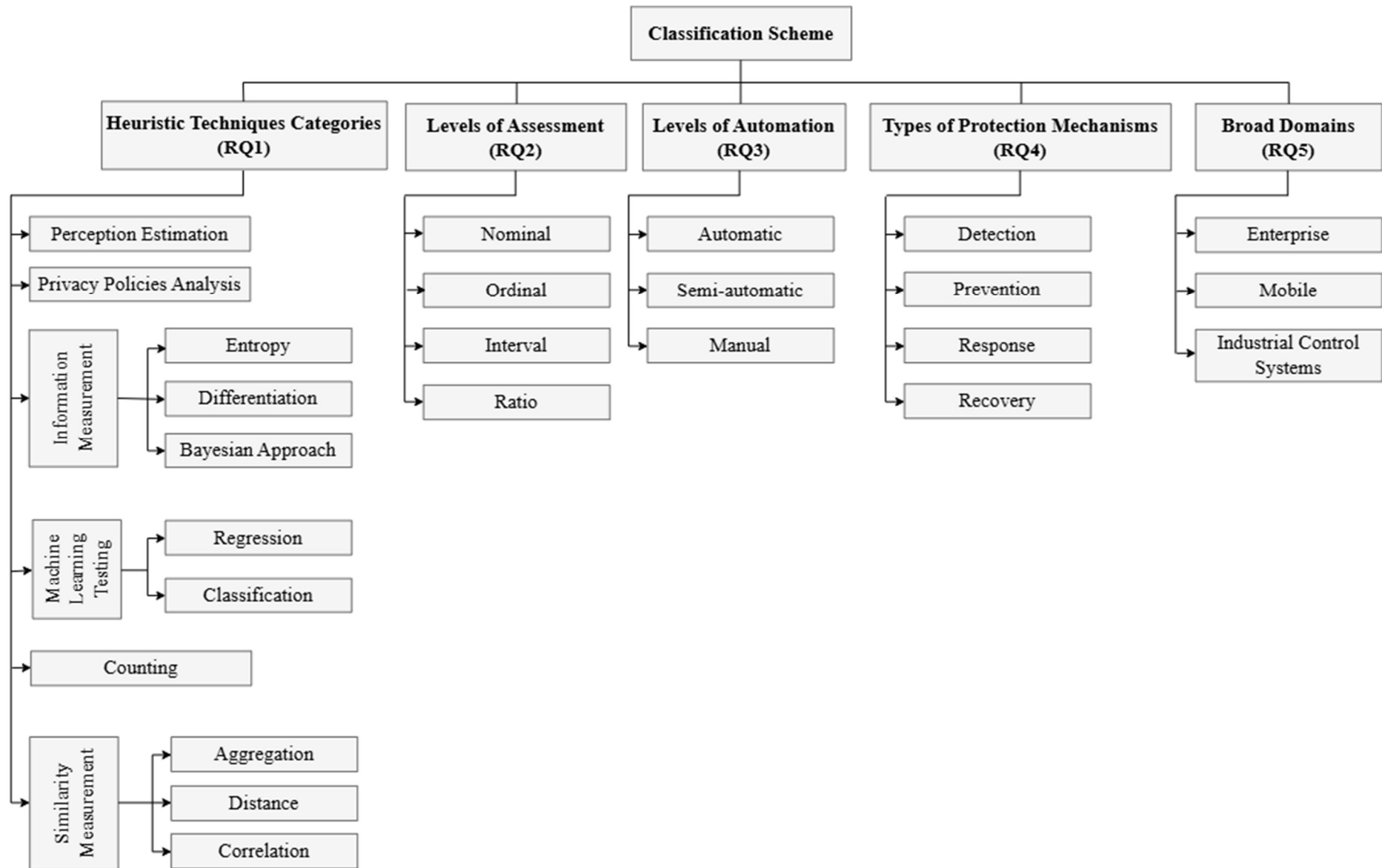
	<b>Title</b>	<b>Detail</b>
<b>Compulsory</b>	Abstract	Resume of the paper. Identify the goal and the problem.
	Introduction	Corroborates what was stated in the introduction and broadens the concepts
	Results	Results demonstrate if there is a proof of the technique or protection mechanism proposed.
	Conclusion	Identify the conclusion achieved by the researcher(s).
<b>Optional</b>	Methodology	The procedure used in order to make the contribution.
	Discussion and Analysis	Review the figures and the results obtained identifying the technique used.

The extracted information will be used for coding the paper according to the coding procedure explained in this document.

### 4.3 Classification for coding

In order to mark and classify and mark the paper the following classification schema has been defined.





**Figure 6** Classification Scheme

### 4.3.1 Heuristic Techniques

Under this category will be identified and located the different heuristic techniques found in the papers.

#### 4.3.1.1 Generalities

**Table 5** Generalities of Heuristic Techniques categories

Descriptor	Detail
Related Research Question	RQ1
Baseline Classification	The proposed taxonomy is a contribution of the research.
Total categories	11

#### 4.3.1.2 Descriptions

**Table 6** Descriptions of Heuristic Techniques categories

Category	Description
Perception Estimation	This technique is based on assessing what people think and feel about their privacy.
Privacy Policies Analysis	The analysis of privacy policies can help to understand how well is privacy carry out in Internet.
Information Measurement-Entropy	This technique measures the grade of uncertainty for a system.
Information Measurement-Differentiation	Measures the difference of uncertainty especially after the application of a protection mechanism.
Information Measurement-Bayesian Approach	Exposes techniques based on the Bayes Theorem.
Machine Learning Testing-Regression	Use metrics for testing machine learning regression models.
Machine Learning Testing-Classification	Use metric for testing machine learning classification models.
Counting	Count the success or failures cases.
Similarity Measurement-Aggregation	Establish how well is specific data aggregated in order to avoid induction.
Similarity Measurement-Distance	Establish the distance from a general data to a specific data.
Similarity Measurement-Correlation	Establish the correlation between general data and specific data.

### 4.3.2 Levels of Assessment

#### 4.3.2.1 Generalities

**Table 7** Generalities of Levels of Assessment categories

Descriptor	Detail
Related Research Question	RQ2
Baseline Classification	Statistical classification [3]
Total categories	4

#### 4.3.2.2 Descriptions

**Table 8** Descriptions of Levels of Assessment categories

Category	Description
<b>Nominal</b>	Metrics that do not expose a sense of order.
<b>Ordinal</b>	Metrics that expose a sense of order.
<b>Interval</b>	Metrics that expose a sense of order and a same interval difference determine the same difference in the assessed property along the scale; but the zero value does not involve the lack of the property.
<b>Radio</b>	Metrics that expose a sense of order and a same interval difference determine the same difference in the assessed property along the scale; but the zero value does not involve the lack of the property.

### 4.3.3 Levels of Automation

This category depicts the levels of automation of the solution found in the literature.

#### 4.3.3.1 Generalities

**Table 9** Generalities of Levels of Automation categories

Descriptor	Detail
Related Research Question	RQ3
Baseline Classification	NIST SP 800-55 Rev. 1 [4]
Total categories	3

#### 4.3.3.2 Descriptions

**Table 10** Descriptions of Levels of Automation categories

Category	Description
<b>Automatic</b>	Automatic tools such as algorithms, bots, web crawlers, and software systems handle all aspects of the assessment process or implementation of the protection mechanism with minimal or no human intervention.
<b>Semi-automatic</b>	Combines both manual and automatic processes. While some aspects of data collection, processing, or analysis are automatic, others require human intervention or oversight. For instance, surveys and questionnaires that are executed manually but have their results analyzed using automatic statistics tools fall into this category.
<b>Manual</b>	Relies on human perception assessment and the interpretation of privacy policies includes, for instance, the execution of surveys, questionnaires, and telephone calls to gather opinions, as well as manual document collection and interpretation.

#### 4.3.4 Types of Protection Mechanisms

The subcategories of types of protection mechanisms are described within this category.

##### 4.3.4.1 Generalities

**Table 11** Generalities of Types of Protection Mechanisms categories

Descriptor	Detail
Related Research Question	RQ4
Baseline Classification	NIST Cybersecurity Framework [5] NIST Privacy Framework [6]
Total categories	4

##### 4.3.4.2 Descriptions

**Table 12** Descriptions of Types of Protection Mechanisms categories

Category	Description
<b>Detection</b>	If the protection mechanism is able to identify and generate alerts to other systems in order to apply other controls. For example, if the protection mechanism is able to send an alarm if certain malware is found in a server, the protection mechanism applies this type of control.
<b>Prevention</b>	This category is oriented to preventing techniques in privacy. For instance, if the protection mechanism includes

	a built-in ciphering protecting personal data from being stolen by a man-in-the-middle attack.
<b>Response</b>	This category includes all the protection mechanisms oriented to respond when a vulnerability has been exploited in order to mitigate its impact at early stages. For example, responding to privacy malware in execution.
<b>Recovery</b>	Once the damage has been done, the protection mechanism contributes to its resilience. For instance, a protection mechanism able to erase stolen personal data by a malware.

#### 4.3.5 Broad Domains

A broad domains classification is defined within this category.

##### 4.3.5.1 Generalities

**Table 13** Generalities of Broad Domains categories

Descriptor	Detail
Related Research Question	RQ5
Baseline Classification	MITRE domains [7]
Total categories	3

##### 4.3.5.2 Descriptions

**Table 14** Descriptions of Broad Domains categories

Category	Description
<b>Enterprise</b>	This broad domain comprises all communication environments that does not involve mobility. For instance, cloud computing and web communication are inside this domain.
<b>Mobile</b>	This broad domain is related to all the pervasive and wireless environments, independently of its technology. Smart Phones, Internet of Things, Ubiquitous systems are examples of this category.
<b>Industrial Control Systems</b>	This broad domain is related to all devices, systems, networks, and controls used to operate and/or automate industrial processes [8].

#### 4.4 Coding procedure

In general, the coding procedure will be carried out in three steps which must be followed in order to classify each paper.

The process will be carried out by each coder and the results will be review in meeting until reach the Alpha Krippendorff's coefficient [2].

#### **4.4.1 Setting up necessary materials**

First and foremost, the materials used for the coding phase must be prepared. These materials are:

- This codebook.
- Alpha Krippendorff's coefficient calculator.
- Software tools: Mendeley and Publish and Perish, or similar.
- Software Microsoft Excel.

#### **4.4.2 Coding execution**

The execution process must be applied to each one of papers and for each one of the coders.

The execution process for each paper is the following:

- Extract the automatic information of the paper in order to identify it.
- Read the compulsory fields in order to extract information to classify the paper.  
The proof of validation established in the Inclusion and Exclusion criteria must be considered as a baseline.  
If the paper does not present a Conclusion section, try to find out something similar in another section of the paper.
- If the compulsory fields are not sufficient to mark the paper in a category, read the optional fields. Try to find out as much information as you can in order to establish a criterion for information classification.
- Mark the paper according to the work carried out in the previous steps and register the result in a software tool or spreadsheet.

Repeat the process for each paper. The final result will be to have all the papers marked in the software tool or spreadsheet.

#### **4.4.3 Meeting of the coders to review results**

After each paper has been codified, a meeting for reviewing the results will be carried out and the following steps must be followed for each paper:

- If an Alpha Krippendorff's coefficient of 0,85 or greater is reached between the coders according to its coding result of the paper, mark the paper in the category established.
- If there is not consensus between the coders about the coding result of the paper, additional meetings in order to achieved that agreement must be carried out until reach an Alpha Krippendorff's coefficient of 0,85.

At the end of this procedure, each one of the papers must be marked with a mutual agreement between the coders.

#### 4.5 General Coding validation

In order to validate the coding procedure, the following general criteria must be applied:

- At least two coders must participate in the process.
- It is strongly encouraged the participation of an expert in Security if possible.
- An Alpha Krippendorff's coefficient of 0,85 has been considered as a measure of consensus.
- In order to mark the papers, the proof of validation established in the Inclusion and Exclusion criteria must be considered as a baseline.

#### REFERENCES

- [1] IEEE Thesaurus, "2023 IEEE Thesaurus", 2023, Available on: <https://www.ieee.org/content/dam/ieee-org/ieee/web/org/pubs/ieee-thesaurus.pdf> Last accessed: 2023-12-11
- [2] University of Pennsylvania, "Computing Krippendorff's Alpha-Reliability", 2011, Available on: [https://repository.upenn.edu/cgi/viewcontent.cgi?article=1043&context=asc\\_papers](https://repository.upenn.edu/cgi/viewcontent.cgi?article=1043&context=asc_papers), Last accessed on 2023-12-27.
- [3] S. L. Weinberg, D. Harel, and S. K. Abramowitz, Statistics Using R: An Integrative Approach. Cambridge University Press, 2023
- [4] Elizabeth Chew, Marianne Swanson, Kevin Stine, N Bartol, Anthony Brown, and W Robinson. 2008. Performance Measurement Guide for Information Security. [https://tsapps.nist.gov/publication/get\\_pdf.cfm?pub\\_id=152183](https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=152183)
- [5] NIST, "NIST Cybersecurity Framework", 2014, Available on: <https://www.nist.gov/system/files/documents/cyberframework/cybersecurity-framework-021214.pdf>. Last accessed on 2022-12-12.
- [6] NIST, "NIST Privacy Framework: A tool for improving privacy through enterprise risk management, version 1.0", 2020, Available on: <https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf>. Last accessed on 2023-12-12.
- [7] MITRE Organization, "MITRE ATT&CK", 2023, Available on: <https://attack.mitre.org/tactics/mobile/>. Last accessed on 2023-12-12.

[8] TrendMicro. “Industrial Control System”, 2023. Available on: <https://www.trendmicro.com/vinfo/us/security/definition/industrial-control-system>. Last accessed on 2023-12-12.