IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 28 January 2023, accepted 11 February 2023, date of publication 14 February 2023, date of current version 20 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3245333

RESEARCH ARTICLE

The AILA Methodology for Automated and Intelligent Likelihood Assignment in Risk Assessment

GIAMPAOLO BELLA[1], **CRISTIAN DANIELE**[2], **AND MARIO RACITI**[1,3]
[1]Dipartimento di Matematica e Informatica, Università di Catania, 95131 Catania, Italy
[2]Department of Digital Security, Radboud University, 6525 XZ Nijmegen, The Netherlands
[3]IMT School for Advanced Studies Lucca, 55100 Lucca, Italy

Corresponding author: Cristian Daniele (cristian.daniele@ru.nl)

This work was supported by the project MEGABIT—Piano di incentivi per la Ricerca di Ateneo 2020/2022 (PIACERI)—linea di intervento 2, DMI—University of Catania.

ABSTRACT This article recognises the widespread application of risk assessment in ICT and aims at reducing the influence of human subjectivity and distraction by means of a methodology for the Automated and Intelligent Likelihood Assignment (AILA). The AILA Methodology, with its various components, applies when risk assessment proceeds exclusively upon information stated in a policy coming as a text document. This scenario is extremely common through small to medium sized institutions. Among the main contributions of this article lies the AILA Entity Extractor, which facilitates the risk assessor in the identification of entities, then of assets, from a given policy. Then, the AILA Classifier automates the assignment of likelihood values to given threats for assets. Moreover, the synergy of AILA with an existing tool for risk assessment demonstrates how to achieve more objective likelihood assignments. AILA is general in support of any risk assessment and, for the sake of demonstration, is applied to assess the privacy risk induced over physical persons by three real-world manufacturers from the automotive domain, namely Toyota, Mercedes and Tesla. AILA is also validated against a risk assessment methodology by ENISA, thereby confirming effectiveness and efficiency of the new methodology (which is dramatically more automated than ENISA's). AILA combines and consolidates together several techniques in an unprecedented fashion, including Natural Language Processing by summarisation and entity recognition, dataset labelling by appeal to the ToS;DR service, and fully-supervised Machine Learning and regression analysis. Finally, to contribute to open knowledge, the general, executable components of AILA, the AILA Entity Extractor and the AILA Classifier are released open source along with the privacy-specific components, the AILA Privacy Dataset and the AILA Privacy Model.

INDEX TERMS Convolutional neural network, likelihood, machine learning, natural language processing, policy, risk assessment.

I. INTRODUCTION

Risk assessment is core to any organisation's evaluation of risk. An asset inventory is often unavailable, especially within small to medium sized organisations, hence the assessment frequently proceeds from information stated in a policy coming as a text document. Therefore, the risk assessor, or analyst in brief, is called to understand documentation that can be long, unclear or incomplete. In consequence, subjectivity and distraction may significantly influence the process, particularly for what concerns the identification of each relevant asset and the assignment of the likelihood value of a given threat to an identified asset. While it still seems impossible to zero the analyst's effort entirely, this article seeks out to automate the analyst's perception of a policy and to reduce the analyst's subjectivity through what perhaps is the hardest step in a risk assessment process: the determination of the likelihood values.

The associate editor coordinating the review of this manuscript and approving it for publication was Li He.

A. RESEARCH QUESTIONS

Following ISO/IEC 27005 [1], the risk assessment process rests on the identification of assets and on the definition of potential threats related to each specific asset. For each asset-threat pair, the analyst is called to determine the "chance of something happening" [1], namely a likelihood value, typically on a scale of 1 to 5. The analyst also has to decide the impact of the occurrence of the threat over the given asset. Risk is calculated as a proportional combination of likelihood and impact for the given asset-threat pair. The overarching motivation for our work is that likelihood determination often implies an approximate estimation that may be biassed by subjectivity. One element of subjectivity is the understanding of available policies, documents that may provide useful information, particularly indicating the relevant assets and threats, in support of the entire risk assessment process, also to inform about likelihood and impact of the given threat on the given asset. Also, policies are often verbose or incomplete and, in any case, long to read, hence their interpretation may be subject to the reader's distraction. On such bases, we introduce the following research questions:

RQ1 *Can we define a computer-supported methodology to assist the human analyst through the extraction of an asset list from a given policy?*

The related work shall demonstrate below that this question is currently open. It is looking at a very common scenario where the analyst has a list of typical threats and is then asked to evaluate them over a target system, say an infrastructure or a software, described through a policy. It would be very useful if a computer-supported methodology could tell the analyst what entities, which include the relevant assets, arise from the policy and also provide a likelihood indication on whether each asset may be affected by any of the given threats.

RQ2 *Can the methodology mentioned in the previous question also assist the human analyst through the assignment of a likelihood to each given threat for any of the assets extracted from the given policy?*

This is most challenging. It would be extremely helpful for the analyst if a computer-supported methodology could provide a likelihood indication on the extent to which each asset may be affected by any of the given threats.

RQ3 *Can the methodology mentioned in the previous questions be integrated with or automated through a tool, namely a software application, if this exists, supporting the overall risk assessment process?*

As we shall see below, also this question is currently open. A few tools exist, and noteworthy is the one that we choose as our main software application due to its maturity and European Commission endorsement, PILAR [2]. However, the arguments unfolded below as well as our corresponding solutions are general and applicable to any tool devoted to supporting risk assessment, e.g. MONARC [3]. For example, existing tools usually come with builtin likelihood values associated with each threat. These values obviously are ignorant of the features and niceties of the target system. For this reason, the tools may also challenge the analyst with the task of determining appropriate likelihood values by hand or, if necessary, of modifying some predefined ones. PILAR, in particular, allows the analyst to enter modifier values to the likelihood values that the tool predefines. Arguably, the modifiers account for target-system specific details, but the challenge for the analyst remains the same.

B. CONTRIBUTIONS

The overarching contribution of this work is a novel methodology for the Automated and Intelligent Likelihood Assignment in a risk assessment process based upon information from a given policy and a given list of threats. Termed AILA Methodology, or AILA in brief, the methodology is released open source in all its components, thereby contributing to open knowledge and fostering further, independent research. The name of the methodology can be understood from the answers it provides to the research questions, as explained here.

1) ANSWERING THE RESEARCH QUESTIONS

The first challenge that the AILA Methodology takes up, following RQ1, is the automated recognition of the relevant assets from the policy by means of Natural Language Processing (NLP) techniques: The software component that we contribute to address this challenge is the AILA Entity Extractor, which is released open source [4] . The biggest challenge for our methodology, following RQ2, is the automated classification of (each statement of) the policy. This is inherently prone to errors [5] but AILA takes an intelligent approach, namely it prescribes the use of a Convolutional Neural Network to train a model in a semi or fully supervised fashion, depending on the available dataset. Next, AILA adopts the trained model to classify the target policies and, following the identification of the relevant assets, facilitate the analyst's task of assigning likelihood values to threats, thereby thwarting subjectivity and distraction. The AILA Classifier is the software component that addresses this second challenge, and is released open source [4]. AILA is general and applicable to different inputs concerning relevant properties such as privacy, cybersecurity and safety. This article demonstrates it over privacy policies, also to overcome the limitations of previous work whose authors state that "*we do not have any publicly available large dataset in the legal domain that has explicitly tagged privacy policies*" [6]. By contrast, we obtain that dataset from outputs of the "Terms of Service; Didn't Read" [7] (ToS;DR), which labels a sentence as fair when it will "respect your rights and will not abuse your data" [7]. Therefore, a fairness label signifies the extent to which a policy statement respects natural persons' privacy. The service covers the privacy policies of popular services such as Amazon, Facebook and Wikipedia. Assuming that dataset, termed AILA Privacy Dataset, to only suffer negligible bias, AILA trains a model, termed AILA Privacy Model, through fully supervised Machine Learning. Both the AILA Privacy Dataset and the AILA Privacy Model are released open source.

AILA also takes the challenge, following RQ3, of enabling the analyst to combine the automatedly assigned likelihood values with those derived from PILAR. We gained by regression analysis an in-depth understanding of PILAR's calculation of risk so as to meaningfully combine it with the AILA Likelihood with the ultimate aim of increasing realism and reliability. AILA's required inputs are one or more policy documents aimed at regulating assets to achieve an overarching property (and the documents are meant to source a risk assessment exercise over the property), a labelled dataset of policies with the same aim and a list of threats to the property.

2) REAL-WORLD DEMONSTRATION

This article continues by demonstrating AILA for the privacy property over three real-world privacy policies, using the dataset derived from ToS;DR and considering a list of threats derived from PILAR. Our target case studies come from the automotive industry and are large car manufacturers' privacy policies. Our choices are Toyota and Mercedes, the first two car manufacturers in Interbrand's 2020 Best Global Brands (BGB) Report [8] (7th and 8th places, respectively, in the overall classification, which also accounts for other areas than automotive). We also add Tesla as a third case study due to its pioneer role on electric cars. We intentionally choose policies that are not currently available from (ToS;DR) to ensure that the particular sentence structure of the dataset would not influence our case studies. In a, Tesla's average risk likelihood per asset (and, arguably, corresponding risk level) turns out lower than Mercedes's, which in turn is found lower than Toyota's.

3) EFFECTIVENESS AND EFFICIENCY

AILA's benefits to risk assessment are validated by comparison with a methodology that was published by ENISA [9]. ENISA's methodology guides the analyst through a number of questions aimed at distilling out the salient features of the target system. The methodology considers the provided answers and uses such information to produce likelihood values that are specific to the target system. Therefore, the ENISA's methodology structures the traditional analyst's step of understanding the technicalities of the system to assign the relevant values. Not surprisingly, we found that going through an execution of the methodology may take up to several dozen minutes. Moreover, it is arguable that it should be followed by a team of analysts to address bias, as is customary. By using AILA as well as ENISA's methodology over the same target, we were pleased to observe that the outcomes are remarkably similar, that is, they bear very high correlation. However, AILA is automated and may run in the lapse of a few minutes, a result that makes it comparatively as effective as ENISA's methodology but dramatically more efficient.

C. ARTICLE SUMMARY

While the gist of AILA was introduced in a short conference paper [10], the present article provides the full details and is approximately three times longer. In particular, the complete methodological technicalities, the integration with PILAR and the final validation step are unpublished. The organisation of the manuscript follows a simple waterfall style.

Section II outlines the related work. Section III defines AILA and demonstrates it over a simple running example. The details of the various steps of the methodology are then given in order: Section IV describes the first step; Section V the second step and Section VI the third step. Section VII presents the application of AILA to the case studies, section VIII provides a validation for the proposed methodology and Section IX concludes.

II. RELATED WORK

The state of the art is conveniently partitioned depending on the related topic.

A. RISK ASSESSMENT TOOLS

A few risk assessment frameworks and tools exist and are applicable to various scenarios in different ways. The main supporting tools developed so far are commercial products and services offered by leading companies in the field. Therefore, it is inherently daunting to interrelate all state-of-the-art software. The European Union Agency for Cybersecurity (ENISA) listed an inventory of the most popular RM / RA approaches [11], between December 2005 and March 2006. Among these, those that are currently maintained are: PILAR [2], offered by the Spanish Ministry for Public Administrations to support the MAGERIT methodology [12]; and TRICK [13], provided by the private company trust and compliant with ISO/IEC 27005. Although both are commercial tools, only PILAR can be used without explicit request to the vendor — the user is allowed to download the software as well as to generate a 30-days evaluation licence locally. Hence our choice to adopt PILAR as the supporting tools for our methodology and experiments, which remain general and may be applied with other tools.

B. NLP AND ML APPLICATIONS TO PRIVACY POLICIES

The importance of fully understanding policies and unveiling hidden risks through their analysis has been observed in several papers. The use of Natural Language Processing techniques has helped researchers to develop tools like the Completeness Analyzer by Costante et al. [14], which assigns a degree of completeness to a policy, as the level of completeness is an important aspect to evaluate in the analysis process. Similarly, other studies aim to estimate the extraction of salient features from a policy, such as the automatic categorisation proposed by Ammar et al. [15], or the possibility to quickly identify and understand relevant privacy statements, using text categorisation, as in the contributions by Liu et al. [16] or Story et al. [17], who framed the problem of identification of practice statements as a classification problem. Furthermore, Ghosh et a. [18] outlined the need for automation to extract requirements specification from documents in a formal fashion. This has subsequently involved the combination of Natural Language Processing with Machine Learning techniques, proven to be a successful duo to fulfil common requirements for policies analysis, namely text summarisation and classification. In fact, while Sathyendra et al. [19] described approaches to automatically extract choice instances from privacy policy documents involving pure NLP, Zaeem et al. [20] advanced a data mining methodology, along with a tool named PrivacyCheck, leveraging both NLP and ML to automatically extract summaries of online privacy policies. Similarly, Tesfay et al. [21] proposed a method to summarise privacy policies into short and condensed notes following a risk-based approach, under the EU GDPR aspects as assessment criteria. Speaking of GDPR, Ou et al. [22] advanced a privacy policy annotation scheme along with an automated Machine Learning method to detect GDPR suspected compliance violations in Websites. In addiction, more complex projects and frameworks have been developed, such as Polisis by Harkous et al. [23], which enables queries on natural language privacy policies, predicting a set of classes for each part of the corpus. This work is useful in understanding the nature of the policy as well as in an automatic annotation of the policy with labels from a prespecified taxonomy. Zaeem et al. [24] used exactly Polisis and PrivacyCheck to compare privacy policies of government agencies and companies. This abundance of contributions led Del Alamo et al. [25] to present the first overview of the different techniques used to analyse privacy policy texts automatically, obtained through a systematic mapping study. However, none of these works suggest relevant information for the purposes of risk assessment, especially for the determination of the likelihood to a certain asset-threat pair.

C. FAIRNESS IN PRIVACY POLICIES

Nagpal et al. [6] were among the first to conduct studies related to the fairness of a policy — the fairness level indicates how fair, proper and clean a text is, regarding the users' privacy concerns. They proposed a methodology to automatically extract a fairness value from public law documents leveraging semantic relatedness, namely the identification of some form of lexical or functional association between two words or concepts, based on the contextual or semantic similarity of those two words, regardless of their syntactical differences. An inherent limitation is the necessity of manually creating a seed set of WordNet [26] senses, which have to be used as a reference for the similarity. Also, the word vector model, namely a type of word representation that allows words with similar meaning to have a similar representation, may turn out unable to represent the various shades of meaning of the same word. Other challenges arise from those sentences bringing hidden implied meaning, as well from those that are meaningful in a specific domain.

III. OVERVIEW OF THE AILA METHODOLOGY

AILA addresses the research questions stated above by supporting the analyst during risk assessment through the automation of the following three steps:

Step 1. ***Entity extraction.*** Automatic entity extraction for the manual identification of the relevant assets from the given policy;

Step 2. ***AILA Likelihood determination.*** Determination of the AILA Likelihood for the automatic assignment of likelihood values to each threat affecting one of the identified assets, in consideration of (specific details of the target system gathered from) the given policy;

Step 3. ***Combined Likelihood determination.*** Determination of the Combined Likelihood for the automatic assignment of likelihood values to each threat affecting one of the identified assets, in consideration of how the AILA Likelihood modifies the likelihood assigned via state-of-the-art tools for risk assessment.

At this point in the presentation, it is useful to demonstrate the outcomes of AILA on a simple running example, while the details of the methodology will be presented in depth later (Section IV for Step 1, Section V for Step 2 and Section VI for Step 3).

A. DEMONSTRATING AILA ON A RUNNING EXAMPLE

Let us consider a fragment of a file management policy as a running example.

> *File management policy, North America. To ensure data privacy, users with different privileges can be created. Any agent can be a Normal User or a Super User. Any agent playing the role of Normal User is Permitted to read the public files. All the public files are stored in the root folder, to be accessible to all the users. Any agent playing the role of Normal User is Permitted to write his own public folders. Each folders can contain only text files. Any agent playing the role of Super User is Obliged to change his password weekly. Any agent playing the role of Normal User is Obliged to change his password monthly. If any agent lost his password, it is possible to request a temporary password which will be valid for 15 minutes from the moment of the request. A Super User can create different Normal Users. Any agent playing the role of Super User is Permitted to read the all files. Any agent playing the role Super User is Permitted to write his own secret file. All the agent names are stored in a special file.*

1) ENTITY EXTRACTION

This step removes the parts of the text that are irrelevant to the extraction of the assets, producing a list of entities, from which the analyst — thanks to his experience — may choose the relevant assets. Additionally, each asset is bundled with the sentences that mention the very asset name or a synonym. First, the policy undergoes text summarisation in terms of N-Grams. The bigrams extracted from our running example are:

[Super, User]
[Normal, User]

[his, password]
[are, stored]

These enter an entity recognition algorithm, which produces the following outcome over our example:

1) *Super user*
2) *Normal user*
3) *Password*

The tool also produces the policy sentences pertaining to the identified entity. In our example, these are:

(A_1) Super user:
- ($S_{1;}1$) *Any agent playing the role of Super User is Obliged to change his password monthly.*
- ($S_{1;}2$) *Any agent playing the role Super User is Permitted to read the all files.*
- ($S_{1;}3$) *Any agent playing the role Super User is Permitted to write his own secret file.*
- ($S_{1;}4$) *A Super User can create different Normal Users.*

(A_2) Normal user:
- a) *Any agent playing the role of Normal User is Permitted to read the public files.*
- b) *Any agent playing the role of Normal User is Permitted to write his own public folders.*
- c) *Any agent playing the role of Normal User is Obliged to change his password monthly.*
- d) *A Super User can create different Normal Users.*

(A_3) Password:
- a) *Any agent playing the role of Super User is Obliged to change his password weekly.*
- b) *Any agent playing the role of Normal User is Obliged to change his password monthly.*
- c) *If any agent lost his password, it is possible to request a temporary password which will be valid for 15 minutes from the moment of the request.*

It is now the analyst's manual work to select the assets among the entities that the tool produced automatically. In this example, it turns out that all entities are valuable assets.

The technical details leading to what was demonstrated here are presented below (IV).

2) AILA LIKELIHOOD DETERMINATION

This step leverages ML for the computer-based assignment of a likelihood value. Of course, having a properly labelled dataset to use to train a model is an essential prerequisite. One of the overarching properties that our example policy aims at is user privacy, which lies among the most discussed ones worldwide at present. Therefore, a dataset that is labelled according to privacy is necessary. We found that this can be formed out from a corpus of privacy policies where each sentence is labelled by the ToS;DR community to signify the sentence's *fairness*, which the community itself interprets as the extent to which the sentence respects natural persons' privacy. Precisely, each sentence is labelled either with a 0 or with a 1 fairness value. Therefore, a privacy-labelled dataset is reached, which we term the AILA Privacy Dataset, and signifies a tangible and general result that can be used to train ML models in various ways. In particular, we used it through a Convolutional Neural Network to train what we term the AILA Privacy Model. Also the AILA Privacy Model is a tangible and general result that can be applied to target privacy policies. The findings obtained by applying the AILA Privacy Model to our running example are reported as "Fairness per sentence" in Table 1. It is then natural to average the fairness values of the sentences related to an asset to derive the fairness value of the asset, which the third column in the table shows as values between 0 and 1 with two decimals. All this information is conveniently displayed to the analyst, as detailed later, through a navigable HTML page, also supporting manual adjustments that the analyst may want to make to the fairness values depending on specific environmental conditions.

The relevant pieces of information are now available to define a likelihood function because the given threats are assumed to pertain to the same property, privacy here, that the given policy wants to establish on the assets. One simple way to define the AILA Likelihood is as the opposite of fairness, but of course finer relations between likelihood and fairness could be encoded.

$$Likelihood = 1 - Fairness$$

The underlying assumption for this choice is that fairness strictly correlates with privacy, coherently with the official arguments by the ToS;DR community. Table 2 maps ranges for the parameters of the equation into the classical likelihood values on range from 1 through to 5 or, equivalently, from very low to very high. This explains the AILA Likelihood values in Table 1, whose remaining columns are discussed below.

The technical details leading to what was demonstrated here are presented below (V).

3) COMBINED LIKELIHOOD DETERMINATION

Once the AILA Likelihood is calculated, it could be used, for example, to continue the risk assessment exercise on a spreadsheet as customary. Moreover, it can be leveraged in various ways to better inform the exercise as carried out on any existing tool for risk assessment. One such tool is PILAR and AILA can perfection the Pilar Likelihood by means of the AILA Likelihood, for example, by the floor of the average of the two. This is demonstrated in the last two columns of Table 1.

Moreover, we must understand how Pilar works to evaluate whether and how the risk levels that the tool calculates are influenced by likelihood variations such as those induced through the AILA Likelihood. It turns out that for any non-irrelevant risk level, namely above 2, likelihood linearly influences risk levels, hence we may conclude that likelihood variations are consistently reflected on risk levels. Not only is this useful to appreciate how AILA ultimately perfections the risk levels produced by the tool, but it is also valuable towards the future goal of implementing an open-source risk assessment tool that natively rests on AILA.

TABLE 1. Outcomes of AILA on our running example.

AILA Step 1				AILA Step 2	AILA Step 3	
Asset	Sententence	Fairness per sentence	Fairness per asset	AILA Likelihood	PILAR Likelihood	Combined Likelihood
A_1	$S_{1;1}$ $S_{1;2}$ $S_{1;3}$ $S_{1;4}$	1.0 0.20 1.0 0.50	0.67	2	3	2
A_2	$S_{2;1}$ $S_{2;2}$ $S_{2;3}$ $S_{2;4}$	0.0 1.0 1.0 0.50	0.62	2	1	1
A_3	$S_{3;1}$ $S_{3;2}$ $S_{3;3}$	1.0 0.80 0.70	0.83	1	3	2

TABLE 2. Discrete mapping of likelihood values.

Fairness per asset	Likelihood per asset	AILA Likelihood	
0 - 0.20	0.80 - 1	5	VH
0.21 - 0.40	0.60 - 0.79	4	H
0.41 - 0.60	0.40 - 0.59	3	M
0.61 - 0.80	0.20 - 0.39	2	L
0.81 - 1	0 - 0.19	1	VL

TABLE 3. Number of entities per preprocessor.

Preprocessor name	Entities extracted after NER		
	Toyota	Mercedes	Tesla
Tldrthis [29]	11	15	25
Autosummarizer [30]	21	25	46
Tools4noobs [31]	33	35	62
AILA preprocessor [4]	52	57	72
none	110	183	334

The technical details leading to what was demonstrated here are presented below (VI).

IV. AILA STEP 1: ENTITY EXTRACTION

This step can be taken using the AILA Entity Extractor (AILAEE), which implements some preprocessing task and then Named Entity Recognition (NER), as shown in Figure 1. Preprocessing influences both the number and the significance of the entities that are extracted later. Named Entity Recognition is now invoked to select the sentences pertaining to the entities produced by the AILA Preprocessor. AILAEE, with its two modules, the AILA Preprocessor and the NER Engine, is available open source [4].

Table 3 compares the number of extracted entities using various preprocessors, or no preprocessing at all, from the privacy policies of our case studies. It can be seen that the AILA Preprocessor leads to more entities than other existing preprocessors do but, at the same time, that number is much lower than the number originating from no preprocessing. However, entities that the AILA Preprocessor decided not to extract, respectively $58 = (110 - 52)$, $126 = (183 - 57)$ and $262 = (334 - 72)$, always contain assets that we deem insignificant for risk assessment, such as "California", "Mercedes" and "hand". Therefore, we are confident that the AILA Preprocessor yields the best tradeoff between cardinality and significance.

The analyst is now called in to manually select the relevant assets from the list of entities output by AILAEE, also with the specific context provided by the pertaining sentences. Even though the role of the analyst is still essential here, it is clearly facilitated because the relevant information comes from a list of entities (and related sentences) rather than from many lines of prose.

A. AILA PREPROCESSOR

The AILA Preprocessor is written in Python and takes advantage of one of the most popular library for Natural Language Processing, namely NLTK. It follows the steps discussed below.

1) BIGRAMS IDENTIFICATION

Bigrams identification is one of the most common techniques to summarise information from text through the most relevant components, namely nouns, verbs and adjectives. Other parts, e.g., prepositions, articles, adverbs, etc., play a lesser role in determining the meaning of sentences [27], [28], hence they are not considered when choosing significant *bigrams*. A bigram is a sequence of two adjacent words. Our preprocessor uses the nltk.word_tokenize function to extract tokens from character strings and then the BigramCollocationFinder function to obtain a list of bigrams. From this list, it removes all those bigrams containing a punctuation character, articles or matchmakers, such as [, and], [, a], [, secondly], [the reason], [an important], and [1)]. Then the processor removes duplicate bigrams.

2) SENTENCE EXTRACTION

For each of the resulting bigrams, the AILA Preprocessor selects parts of the original text near it, extracting the original sentence that contains it as well as the sentence before and after it. However, only the sentences containing a verb are stored, and this is achieved by POS tagging through the nltk.pos_tag

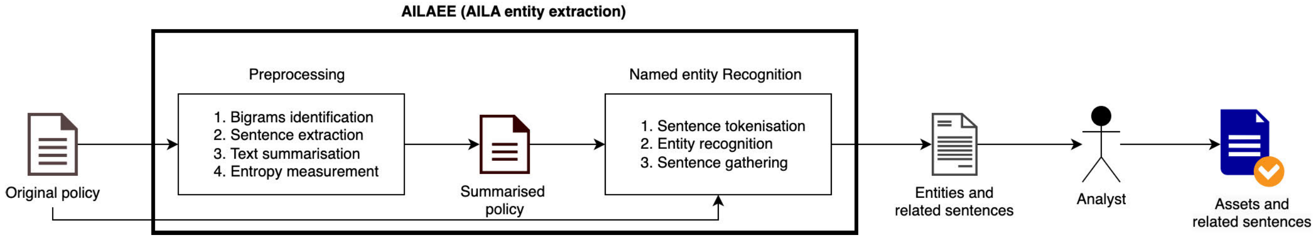


FIGURE 1. AILA Step 1 (AILAEE).

function. Therefore, it discards irrelevant sentences such as titles and subtitles.

3) TEXT SUMMARISATION

There are two approaches to summarise text: abstraction and extraction [32]. Abstraction involves deep learning techniques to shorten the original document by emulating a human. Extraction shortens a text by combining a subset of weighted words, which represent the most relevant parts of the text. Various kinds of algorithms can be used to calibrate the weights of the sentences and rank them according to their relevance in the document. A few tools are available for free. Having tried out some of them (such as Tldrthis [29], Autosummarizer [30], Tools4noobs [31]) to evaluate if entity recognition improves, it turns out that none of the examined tools is useful. This might be due to the fact that those tools are built to summarise narrative texts and not privacy policies. Therefore, we also implement a summarising tool in the AILA Preprocessor and find out that it consistently augments entity recognition through all our experiments of at least 13%. It rests on the summarise function, which receives a list of sentences linked to certain bigrams as input and performs the following steps:

1) it tokenises all sentences and calculates the frequencies of each word;
2) it calculates the score of each sentence by adding up the frequencies of the words in the sentence;
3) it extracts the sentence with the greatest score.

4) ENTROPY MEASUREMENT

The AILA Preprocessor ultimately evaluates whether the loss of information between the original set of sentences and the chosen sentence is negligible. It does so by calculating Shannon's entropy on the original set of sentences, then on the chosen sentence and by evaluating the difference. Should that difference be non-negligible, the preprocessor opts for the next candidate in the list of summarised sentences.

B. NER ENGINE

Named Entity Recognition (NER) is a subtask of information extraction that seeks to locate and classify named entities mentioned in unstructured text into predefined categories (i.e., names, locations, quantities, organisations, etc.). Entities may contain what we call an asset in a risk assessment process or, in other cases, may express a subcategory or property related to an asset. For this purpose, Dandelion [33] conveniently provides us with a set of NLP services that can be accessed via REST APIs. We use these to write the NER Engine of AILAEE, another Python tool that relies on the NLTK and dandelion-eu libraries. The output comes as a JSON file containing pairs formed by an entity and pertaining to sentences. These offer a great simplification of the original policy to the analyst's eyes, especially for the selection of the relevant assets from the distilled entities. The engine carries out three steps, detailed here.

1) SENTENCE TOKENISATION

The engine tokenises both the original text and the summarised version yielded by the AILA Preprocessor into individual sentences, by using the nltk.sent_tokenize function.

2) ENTITY RECOGNITION

The text summarisation described above improves the accuracy of the engine as it provides a shortened and most relevant input: the Dandelion Entity Extraction service performs very well also on short texts. For each tokenised sentence from the summarised text, the tool performs an HTTP request to the Dandelion APIs for the Entity Extraction service [34] through the DataTXT.nex method by specifying the inclusion of alternate_lables, namely synonyms. The service returns the entities and the synonyms of each.

3) SENTENCE GATHERING

The NER Engine gathers, for each entity, all the sentences that contain the entity or its synonyms in the original text. Precisely, for each sentence featuring at least a verb, the engine compares the entity and its synonyms with each word of the i-th sentence, previously tokenised by the nltk.word_tokenize function. The results are stored in a JSON file also in this case.

V. AILA STEP 2: AILA LIKELIHOOD DETERMINATION

The previous step is preparatory to our development of the AILA Classifier (AILAC), as depicted in Figure 2. As noted above (Section III), its core includes a dataset to train a ML model to label sentences with a likelihood value, then the model's practical use over real-world policies. Also, AILAC is available open source [4].

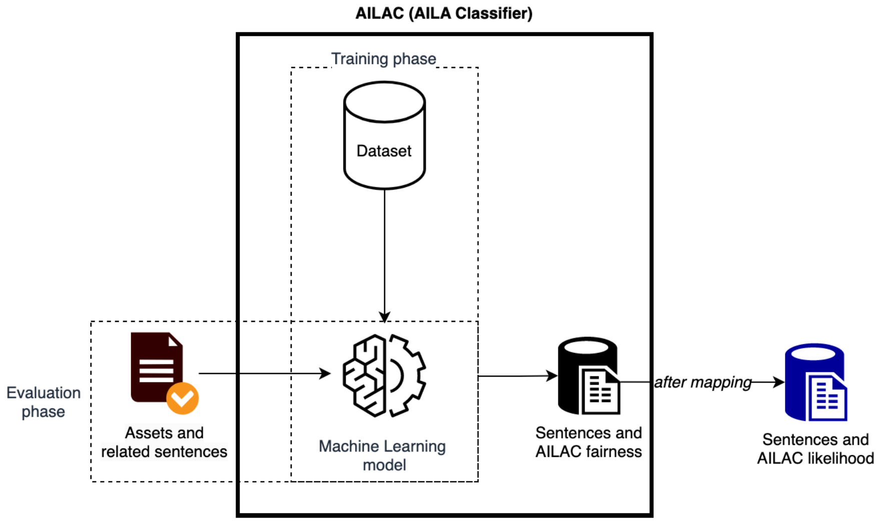


FIGURE 2. AILA Step 2 (AILAC).

While all techniques and tools discussed thus far are general, the specific dataset and ML model discussed below are tailored to privacy policies, hence their names. Equivalent versions could be similarly built for other properties.

A. THE AILA PRIVACY DATASET

As an ML model requires solid data for training, we created a dataset starting from a corpus of sentences already labelled by the ToS;DR community in terms of fairness. Such sentences are gathered from the following popular services: Amazon, Apple, Blizzard, CNN, DuckDuckGo, Facebook, Google, Khan Academy, Paypal, Pinterest, Pornhub, Quora, Reddit, Spotify, Startpage, WikiHow, Wikipedia, YouTube. The resulting corpus contains 500 sentences, which are then enriched by using text augmentation and synonyms following a standard practice [35]. This yields the AILA Privacy Dataset, a corpus of over 100.000 labelled sentences, which is released open source to foster future research.

B. THE AILA PRIVACY MODEL

Convolutional Neural Networks are normally represented by a sequential architecture, which is created by passing a list of layers [36]. We build the AILA Privacy model by using Keras, a Deep Learning API written in Python, as demonstrated in Figure 3. The model is trained on more than 75.000 labelled sentences, which are transformed into a 2-D feature matrix for the training step. Precisely, the Relu function is chosen as activation function of the first layer, the Sigmoid function for the second layer and the Adam function for the optimisation. The model gets trained for 15 epochs using the binary cross-entropy function as loss function, 0.0001 as learning rate, and 50 as batch size. The dataset is split into two parts: the first, corresponding to the 75% of the total corpus, is used to train the model; the rest is employed for the testing step.

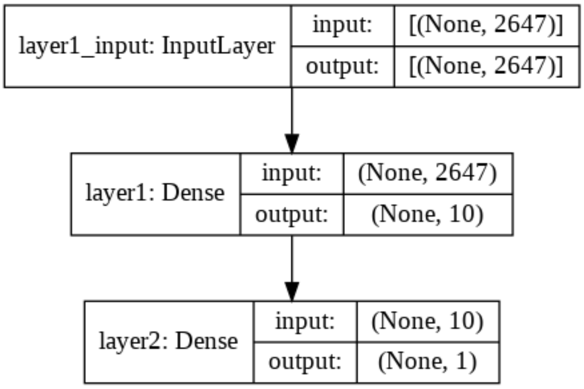
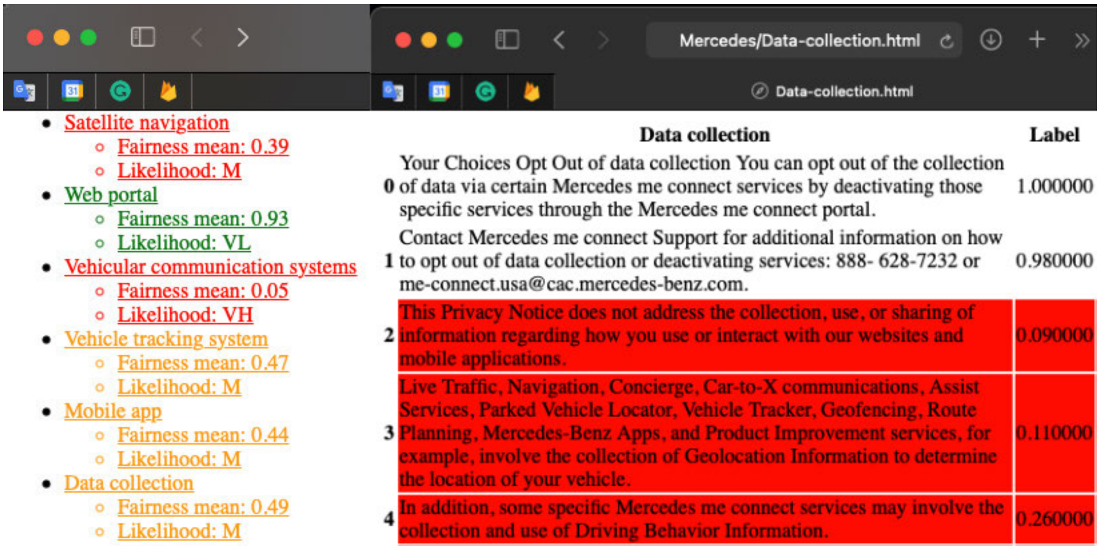


FIGURE 3. Structure of the AILA privacy model.

The resulting model manages to classify sentences with an accuracy of 96%.

As noted above (Section III), the AILA Privacy model classifies sentences with a fairness value in the range [0, 1]. For example, the sentence ''We will sell all your data'', will be classified with 0, whilst the sentence ''We will save your data at any cost'', will be classified with 1. The trained model is used to evaluate the fairness of the sentences extracted at the end of the NLP step. For each entity, we calculate the fairness of the related sentences and assign the mean fairness of these to the entity. The values are then used for the likelihood definition, as discussed in Table 2. AILAC stores this output in a folder containing the following files:

- an index HTML file with all identified entities and their respective fairness and likelihood;
- an HTML file for each identified entity with all the sentences related to that entity and the fairness per sentence.



- Satellite navigation
  - Fairness mean: 0.39
  - Likelihood: M
- Web portal
  - Fairness mean: 0.93
  - Likelihood: VL
- Vehicular communication systems
  - Fairness mean: 0.05
  - Likelihood: VH
- Vehicle tracking system
  - Fairness mean: 0.47
  - Likelihood: M
- Mobile app
  - Fairness mean: 0.44
  - Likelihood: M
- Data collection
  - Fairness mean: 0.49
  - Likelihood: M

	Data collection	Label
0	Your Choices Opt Out of data collection You can opt out of the collection of data via certain Mercedes me connect services by deactivating those specific services through the Mercedes me connect portal.	1.000000
1	Contact Mercedes me connect Support for additional information on how to opt out of data collection or deactivating services: 888- 628-7232 or me-connect.usa@cac.mercedes-benz.com.	0.980000
2	This Privacy Notice does not address the collection, use, or sharing of information regarding how you use or interact with our websites and mobile applications.	0.090000
3	Live Traffic, Navigation, Concierge, Car-to-X communications, Assist Services, Parked Vehicle Locator, Vehicle Tracker, Geofencing, Route Planning, Mercedes-Benz Apps, and Product Improvement services, for example, involve the collection of Geolocation Information to determine the location of your vehicle.	0.110000
4	In addition, some specific Mercedes me connect services may involve the collection and use of Driving Behavior Information.	0.260000

FIGURE 4. Sample outputs of the AILA classifier.

Figure 4 provides two small extracts of the output, where values are coloured through an obvious colour scale.

VI. AILA STEP 3: COMBINED LIKELIHOOD DETERMINATION

This step of AILA investigates the integration of the automated assignment of likelihood values that is discussed above with a state-of-the art tool for risk assessment such as PILAR [2]. PILAR is a commercial semi-quantitative tool that comes with extensive classes of assets and threats. The analyst manually inserts the desired assets and assigns one (or more) available class(es) to each asset. The analyst then evaluates all assets with a value (possibly on several dimensions), reflecting how the asset is relevant for the case, and PILAR provides guidance through different labelled options. As for threats, because the tool comes with threats per each class of assets, each asset of the analyst's automatically gets its threats. Remarkably, each threat for each asset comes with a predefined likelihood value, which therefore is ignorant of the specific case study although the analyst is given the chance to accept or modify it. With this information, the tool is able to automatically determine all the values needed to calculate impact and then risk. It is clear that the automated choice of assets and the determination of likelihood values from given policies explained above facilitate the analyst also through the interaction with PILAR. First of all, the analyst quickly determines the assets to insert in the tool but; most importantly, the analyst mechanically obtains likelihood values that are not only informed by the given policies but also free from subjective influence. The analyst may finally choose to update the original Pilar Likelihood values in the tool, for example, with the AILA ones or with an average of the two. However, PILAR is not open source, hence we have no access to the source code and, especially, to its algorithms. Therefore, the extent to which the likelihood values (and in particular the updated ones) influence the ultimate risk levels that the tool computes is far from obvious. We seek out to understand that beyond what empirical tests can demonstrate, hence also appeal to regression analysis below. As a result, we find out that the likelihood values affect risk levels linearly for risk levels above 2, hence come to the conclusion that likelihood variations are consistently reflected on risk levels according to a linear relation.

TABLE 4. PILAR levels map.

Level	Value
0	1000
1	2150
2	4650
3	10000
4	21500
5	46500
6	100000
7	215000
8	465000
9	1000000
10	2150000

A. UNDERSTANDING IMPACT IN PILAR

PILAR's underlying methodology indicates impact depends on asset value and on its degradation [37]. The tool uses a table to map values with levels, shown in Table 4, stored in a file named "levels.xml". It can be seen that three levels correspond to one order of magnitude. Thus, we can deduce that a 10% of degradation would mean decreasing the value of the asset by three steps.

Therefore, it is easy to discover that PILAR calculates impact by the following equation:

$$I = (V \times d) \tag{1}$$

where I is the impact, V is the asset value and d is the degradation. Both I and V are expressed according to their corresponding value in the maps. Once the impact value is calculated, we can retrieve its level by an exponential fit. With tools such as WolframAlpha, it is possible to input a map as "exponential fit {{0, 1000}, {1, 2150}, . . . , {10, 2150000}}" and find an exponential function that approximates the trend of the sought function with a reliability index of 99%. In this case, we find the function:

$$y = 1002.75 \times e^{0.767241\ x} \quad (2)$$

It means that this function approximates the one that PILAR uses, with an error equal to 1%. Thus, for example, given $V = 6 (= 100000)$ and $d = 20\%$, we can calculate the impact by applying both Eq. 1 and the foreseen exponential fit equation:

$$I = (V \times d) = 100000 \times 20\% = 20000 \simeq_{(2)} 3.9 \simeq 4 \quad (3)$$

Moreover, if we are interested in calculating only the discrete value of the impact, it is possible to follow this reasoning: $d = 1\%$ means a decrease of the asset value of 6, $d = 10\%$ a decrease of 3, $d = 20\%$ a decrease of 2, $d = 50\%$ a decrease of 1 and $d = 100\%$ a decrease of 0.

B. UNDERSTANDING RISK LEVELS IN PILAR

The calculation of risk levels is more complicated as we do not know how PILAR exactly assigns likelihood values. The official glossary of the tool [38] only states that PILAR uses a heat map to calculate the risk level and that "*In a qualitative risk analysis, the relative likelihood is the relevant information*" but no additional information.

Likelihood may take five labels in the tool, VL, L, M, H, VH, but we are unaware about the numerical value that PILAR assigns to each of these. Moreover, we tried to derive the likelihood value from the simple equation $R = I \times L$, but empirical attempts demonstrated that the ratio R/I outputs different numbers even though the risk level is determined by the same likelihood label.

Therefore in order to understand how risk is calculated by the tool, we tried to reverse engineer its equation. Differently from the impact calculation, where we had the level map as a cornerstone, here the process is entirely based on experimental trials and empirical data. Therefore, we created a simple project in PILAR with only a few assets and only assigned the essential values. Then, we concentrated on assets that are relevant to privacy.

Since risk depends on two main factors, we start to observe its change in value by assigning different labels to the likelihood, by holding the impact first, and to the impact, by holding the likelihood later. As a result, we conjecture the following equation:

$$R = 0.6 \times I + L \quad (4)$$

where R is the risk level, I is the impact and L is the likelihood value according to the following map: $VL \approx -0.9$; $L \approx 0$; $M \approx 0.9$; $H \approx 1.8$; $VH \approx 2.7$. This map can be confirmed by observing, through systematic experiments, that the risk level changes by some 0.9 when the likelihood varies by one step. Figure 5 shows the risk map based upon the conjectured formula; negative values are set to 0 because risk levels range in [0, 9] in PILAR.

We may now compare the values calculated by the conjectured formula with those returned by PILAR by performing a linear regression analysis, with a sample of 59 elements that cover almost all possible outputs.

Once the two sets to compare are available, we proceed to perform a linear fit with the help of a tool, such as Google Sheets. By appealing to the *CORREL* function, a value for Pearson's correlation coefficient arises: $r = 0.9909792073$. Because a value of $+1$ means total positive linear correlation, 0 is no linear correlation and -1 is total negative linear correlation, we conclude that the two sets strictly correlate. Now, we calculate a linear fit with the help of the *LINEST* function, which returns the values of the intercept and the slope, and derive the following equation:

$$y = 0.97x + 0.15 \quad (5)$$

The comparison between the two sets is depicted in Figure 6. The ordinate indicates the risk level. It can be seen that the two sets follow a very similar trend. In particulars, the points are strictly close to each other starting from a risk level of 2. The red line also shows that the formula used by PILAR assumes a different behaviour near the lower values. On the other hand, the conjectured formula can assume negative values, which clashes with the domain of the risk levels. However for risk levels above this threshold, the conjectured formula expressed in Eq. 4 empirically seems to fit the actual values as well.

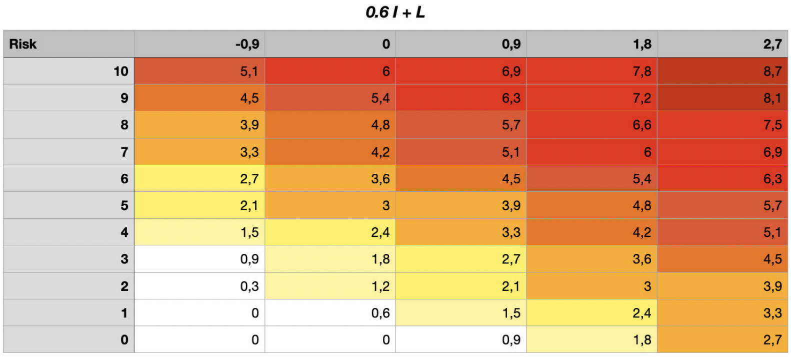
Before concluding this Section it is also important to highlight that the risk levels calculated by the tool vary by a few decimals — depending on the type of threat considered. This might be due to what the documentation defines as the "*relative likelihood*" of a threat and, logically, it might also be related to the class which the asset belongs to, as well.

TABLE 5. Asset extraction from policies.

Policy	Original words	Words after summarisation	Entities	Assets
Toyota	3526	768	52	19
Mercedes	1800	402	57	17
Tesla	6860	1164	72	21

TABLE 6. Correlation coefficients between the AILA Likelihood and the ENISA Likelihood.

Coefficient	Value
Pearson (r)	0.93
Spearman (r_s)	0.91
Statistical significance (p−value)	0.00026



0.6 I + L

Risk	-0,9	0	0,9	1,8	2,7
10	5,1	6	6,9	7,8	8,7
9	4,5	5,4	6,3	7,2	8,1
8	3,9	4,8	5,7	6,6	7,5
7	3,3	4,2	5,1	6	6,9
6	2,7	3,6	4,5	5,4	6,3
5	2,1	3	3,9	4,8	5,7
4	1,5	2,4	3,3	4,2	5,1
3	0,9	1,8	2,7	3,6	4,5
2	0,3	1,2	2,1	3	3,9
1	0	0,6	1,5	2,4	3,3
0	0	0	0,9	1,8	2,7

FIGURE 5. Conjectured risk map.

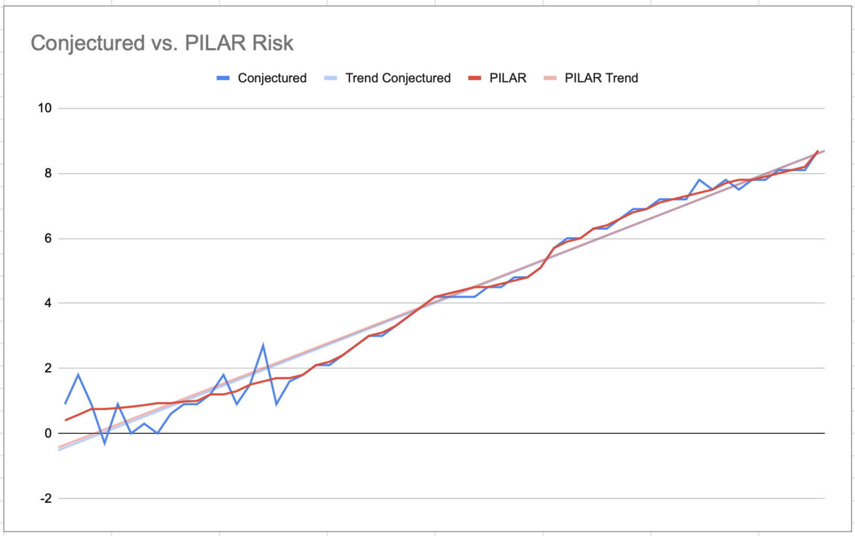


FIGURE 6. Linear Regression of PILAR Risk with Conjectured Risk.

VII. CASE STUDIES: TOYOTA, MERCEDES AND TESLA

Cars are increasingly complex and interconnected, treating a variety of personal data such as cabin preferences, music preferences, GPS coordinates and sensor data including camera streams. Car manufacturers therefore are data controllers that are called to comply, at least in Europe, with the GDPR. It is thus not surprising that car manufacturers' privacy policies are very developed. We demonstrate the outcomes of AILA on the privacy policies of Toyota and Mercedes, the first two car manufacturers in Interbrand's 2020 Best Global Brands (BGB) Report [8] (7th and 8th places, respectively, in the overall classification, which accounts for other areas too). We also add Tesla's privacy policy as a third case study due to the brand's pioneer role on electric cars. Table 5 which summarises the automated asset extraction step, demonstrates the simplifications for the analysts by showing how AILA reduces the original word numbers to a few dozen entities, out of which the analyst may conveniently choose the assets deemed relevant for the specific risk assessment exercise. After having obtained the likelihood of the relevant assets thanks to AILA for each of the three car brands, it can be seen that Tesla's average likelihood (and, arguably, corresponding risk level) is medium (or 3, or M), corresponding (by Table 2) to an average fairness level of 0,41; Mercedes's average likelihood is high (or 4 or H), corresponding to an average fairness level of 0,26; Toyota's average likelihood is very high (or 5 or VH), corresponding to an average fairness level of 0,14. On one hand, the fact that these values may appear to be high may be due to the stringent criteria that ToS;DR adopts for their labelling, which our ML model inherits through our training dataset. On the other hand, the significance of the very numbers through qualitative risk assessment is limited, while the relative differences are most relevant: Tesla is found to come with the fairest privacy policy. Finally, Table 8, Table 9 and Table 10 show one example assets per car brand and its respectively associated PILAR class and threats. The asset

TABLE 7. AILA and ENISA Likelihood related to Mercedes assets.

Asset	AILA Fairness	AILA Likelihood (1-Fairness)	AILA Likelihood	ENISA Likelihood
Geolocation	0.23	0.77	High	High (9)
Maintenance	0.38	0.62	Medium	High (9)
Vehicle tracking system	0.4	0.6	Medium	Medium (8)
System	0.1	0.9	Very High	High(9)
Bill	0	1	Very High	High(10)
Payment Information	0.05	0.95	Very High	High(10)
Mobile Application	0.44	0.56	Medium	Medium (8)
Data collection	0.49	0.51	Medium	Medium (7)
Web portal	0.93	0.07	Very Low	Medium (6)
Emergency service	0.33	0.67	High	Medium(8)

TABLE 8. Toyota privacy policy.

PILAR Class	AILA Asset	PILAR Threat	PILAR Likelihood	AILA Likelihood	Combined Likelihood
Software	Application	Hardware or software failure Software vulnerabilities Defects in software maintenance /updating Malware diffusion Software manipulation	3 3 4 3 3	4	3.6
Communication	Location	Accidental alteration of the information Information leaks Unauthorised access Traffic analysis Eavesdropping Deliberate alteration of information Destruction of information	3 3 3 3 3 3 3	5	4

TABLE 9. Mercedes privacy policy.

PILAR Class	AILA Asset	PILAR Threat	PILAR Likelihood	AILA Likelihood	Combined Likelihood
Software	Mobile app	Hardware or software failure Software vulnerabilities Defects in software maintenance /updating Malware diffusion Software manipulation	3 3 4 3 3	3	3.1
Communication	Vehicle tracking system	Accidental alteration of the information Information leaks Unauthorised access Traffic analysis Eavesdropping Deliberate alteration of information Destruction of information	3 3 3 3 3 3 3	3	3

TABLE 10. Tesla privacy policy.

PILAR Class	AILA Asset	PILAR Threat	PILAR Likelihood	AILA Likelihood	Combined Likelihood
Software	Mobile App	Hardware or software failure Software vulnerabilities Defects in software maintenance /updating Malware diffusion Software manipulation	3 3 4 3 3	4	3.6
Communication	Information	Accidental alteration of the information Information leaks Unauthorised access Traffic analysis Eavesdropping Deliberate alteration of information Destruction of information	3 3 3 3 3 3 3	3	3

different yet similar among the brands, pertain to the same PILAR classes, Software. Therefore, it can be noticed that the associated threats and Pilar Likelihood values remain unvaried across the brands. This is the best the tool can do without the analyst's intervention. By contrast, AILA Likelihood values vary because they are tailored to the very contents of the reference privacy policy. The total 57 assets for the three car brands along with their likelihood values are not presented here but are available online [4].

VIII. VALIDATION

We validated AILA against a methodology promoted by ENISA [9]. The latter estimates the risk level for a personal data processing operation by guiding the analyst through a (rather lengthy) partially pre-filled form. The biggest implication is that, contrarily to AILA, the relevant data is entirely human-generated, hence the analyst must first read the privacy policy as accurately as possible to acquire the relevant information.

To compare the outcomes of the two approaches, we engaged in individual readings of Mercedes's privacy policy, then familiarised with ENISA's methodology and answered its questions during a focus group session. The resulting AILA Likelihood and the ENISA likelihood are summarised in Table 7 in relation to all extracted entities. The detailed ENISA outputs are not presented here but are available online [4]. While the outcomes may look remarkably similar, Pearson's coefficient r and Spearman's rank correlation coefficient r_s, along with the level of statistical significance $p-$value, quantify the correlation precisely. Table 6 shows that both r and r_s are very close to 1, while the $p-$value tends to null. These numbers highlight a substantially positive correlation between the AILA likelihood and the ENISA likelihood, confirming that AILA performs well with respect to ENISA's methodology. However, it is noteworthy that our methodology reaches its outcome in a fully automated way.

IX. CONCLUSION

This article advanced AILA, an innovative methodology to reduce human subjectivity through risk assessment, and applied it to the assessment of given threats related to privacy. However, AILA is general for any risk assessment exercise relying on likelihood assignment upon the basis of information stated in prose in a policy document. AILA responds somewhat positively to the stated research questions, which pertain to how to automate the entity — asset, after the expert validation — extraction process from a policy, how to automate the likelihood assignment to given threats for those assets and how integrate the above with a tool of the state of the art. AILA's main software components, the AILA Entity Extractor and the AILA Classifier are released open source to promote the widespread development of the area. AILA's integration with PILAR is conceptually simple now that we understand how the latter calculates impact and risk levels, but cannot be completed because PILAR is not open source. The application of AILA to the automotive field was profitable on all three case studies, Toyota, Mercedes and Tesla, showing how to reduce a few thousand words to only a few dozen entities, hence facilitating asset extraction dramatically. AILA also conveniently automated the assignment of the likelihood values for all assets, offering significant reduction of subjectivity. A limitation of AILA is that it can be used only when a non biassed and labelled dataset is available in order to train the ML model; still, this is an inherent limitation of ML in general. Future work includes deeper semantic analysis of the policies to tune the likelihood values, which are currently assigned per asset, more finely per asset and per threat. Another useful direction would be to write an open-source risk assessment tool from scratch to truly integrate PILAR with AILA. Open sourcedness would spark off a community of developers as well as additional research, and the new AILA Methodology could bear the (at least de facto) standard tool for risk assessment.

ACKNOWLEDGMENT

The authors would like to thank Dandelion Support Team, for providing them a non-profit plan, and Jose A. Manas, for the precious suggestions about PILAR and in general risk assessment.

DECLARATIONS

COMPLIANCE WITH ETHICAL STANDARDS

This article does not contain any studies with human participants or animals performed by any of the authors.

COMPETING INTERESTS

The authors declare that they have no conflict of interest.

RESEARCH DATA POLICY AND DATA AVAILABILITY STATEMENTS

All the data used for this article are available online [4].

REFERENCES

[1] Organizacion Internacional de Normalizacion. (2008). *ISO/IEC 27005: Information Technology-Security Techniques Information Security Risk Management*. [Online]. Available: https://books.google.it/books?id=K1HbZwEACAAJ

[2] Ministero De Administractiones Pùblicas. *Ear/Pilar*. Accessed: 2012. [Online]. Available: https://pilar-tools.com/en/index.html

[3] CASES Team. *MonarC*. Accessed: 2022. [Online]. Available: https://www.monarc.lu

[4] G. Bella, C. Daniele, and M. Raciti. (2021). *Aila Source Code*. [Online]. Available: https://github.com/cristiandaniele/AILA-source-code

[5] M. Roehling, "'Extracting' policy from judicial opinions: The dangers of policy capturing in a field setting," *Personnel Psychol.*, vol. 46, no. 3, pp. 477–502, 1993. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.1993.tb00881.x

[6] R. Nagpal, C. Wadhwa, M. Gupta, S. Shaikh, S. Mehta, and V. Goyal, "Extracting fairness policies from legal documents," 2018, *arXiv:1809.04262*.

[7] *Term of Services; Didn't Read*. Accessed: 2022. [Online]. Available: https://tosdr.org

[8] Interbrand. (2020). *Interbrand's 2020 Best Global Brands (BGB) Report*. [Online]. Available: https://www.interbrand.com/best-global-brands/

[9] *On-Line Tool for the Security of Personal Data Processing*. Accessed: 2020. [Online]. Available: https://www.enisa.europa.eu/risk-level-tool/risk/

[10] G. Bella, C. Daniele, and M. Raciti, "The AILA methodology for automated and intelligent likelihood assignment," in *Proc. 6th Int. Conf. Cryptogr., Secur. Privacy (CSP)*, Jan. 2022, pp. 119–123.

[11] *RM/RA Tools*. Accessed: 2022. [Online]. Available: https://www.enisa.europa.eu/topics/threat-risk-management/risk-management/current-risk/risk-management-inventory/rm-ra-tools

[12] *Magerit*. Accessed: 2012. [Online]. Available: https://www.ar-tools.com/magerit/index.html

[13] *Trick*. Accessed: 2018. [Online]. Available: https://www.trickservice.com

[14] E. Costante, Y. Sun, M. Petković, and J. D. Hartog, "A machine learning solution to assess privacy policy completeness: (Short paper)," in *Proc. ACM Workshop Privacy Electron. Soc. (WPES)*. New York, NY, USA: Association for Computing Machinery, Oct. 2012, pp. 91–96, doi: 10.1145/2381966.2381979.

[15] W. Ammar, S. Wilson, N. Sadeh, and N. A. Smith, ''Automatic categorization of privacy policies: A pilot study,'' School Comput. Sci., Lang. Technol. Inst., Pittsburgh, PA, USA, Tech. Rep. CMU-LTI-12-019, 2012.

[16] F. Liu, S. Wilson, P. Story, S. Zimmeck, and N. Sadeh, ''Towards automatic classification of privacy policy text,'' School Comput. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-ISR-17-118R CMULTI-17-010, 2018.

[17] P. Story, S. Zimmeck, A. Ravichander, D. Smullen, Z. Wang, J. R. Reidenberg, N. C. Russell, and N. M. Sadeh, ''Natural language processing for mobile app privacy compliance,'' in *Proc. AAAI Spring Symp. Privacy Enhancing AI Lang. Technol.*, 2019.

[18] S. Ghosh, D. Elenius, W. Li, P. Lincoln, N. Shankar, and W. Steiner, ''Arsenal: Automatic requirements specification extraction from natural language,'' in *NASA Formal Methods*, S. Rayadurgam and O. Tkachuk, Eds. Cham, Switzerland: Springer, 2016, pp. 41–46.

[19] K. M. Sathyendra, F. Schaub, S. Wilson, and N. M. Sadeh, ''Automatic extraction of opt-out choices from privacy policies,'' in *Proc. AAAI Fall Symp.*, 2016, pp. 1–5.

[20] R. N. Zaeem, R. L. German, and K. S. Barber, ''PrivacyCheck: Automatic summarization of privacy policies using data mining,'' *ACM Trans. Internet Technol.*, vol. 18, no. 4, pp. 1–18, Aug. 2018, doi: 10.1145/3127519.

[21] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, and J. Serna, ''I read but don't agree: Privacy policy benchmarking using machine learning and the EU GDPR,'' in *Proc. Companion Web Conf. Web Conf. (WWW)*. Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2018, pp. 163–166, doi: 10.1145/3184558.3186969.

[22] H. Ou, Y. Fang, Y. Guo, W. Guo, and C. Huang, ''Viopolicy-detector: An automated approach to detecting GDPR suspected compliance violations in websites,'' in *Proc. 25th Int. Symp. Res. Attacks, Intrusions Defenses (RAID)*. New York, NY, USA: Association for Computing Machinery, Oct. 2022, pp. 409–430, doi: 10.1145/3545948.3545952.

[23] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin, and K. Aberer, ''Polisis: Automated analysis and presentation of privacy policies using deep learning,'' in *Proc. 27th USENIX Conf. Secur. Symp. (SEC)*. New York, NY, USA: USENIX Association, 2018, pp. 531–548.

[24] R. Zaeem and K. Barber, ''Comparing privacy policies of government agencies and companies: A study using machine-learning-based privacy policy analysis tools,'' in *Proc. 13th Int. Conf. Agents Artif. Intell. (ICAART)*. Setúbal, Portugal: SciTePress, 2021, pp. 29–40.

[25] J. M. D. Alamo, D. S. Guaman, B. García, and A. Diez, ''A systematic mapping study on automated analysis of privacy policies,'' *Computing*, vol. 104, no. 9, pp. 2053–2076, Sep. 2022, doi: 10.1007/s00607-022-01076-3.

[26] *WordNet*. Accessed: 2005. [Online]. Available: https://wordnet.princeton.edu

[27] N. M. Alsharman and I. V. Pivkina, ''Generating summaries through unigram and bigram: Text summarization,'' *Int. J. Inf. Technol. Web Eng.*, vol. 15, no. 1, pp. 64–74, Jan. 2020.

[28] E. Villatoro-Tello, L. Villase nor-Pineda, and M. M.-Y. Gómez, ''Using word sequences for text summarization,'' in *Proc. Int. Conf. Text, Speech Dialogue*. Berlin, Germany: Springer, 2006, pp. 293–300.

[29] S. Gard. *TLDR This*. Accessed: 2021. [Online]. Available: https://tldrthis.com

[30] (2013). *Automatic Text Summarizer*. [Online]. Available: https://autosummarizer.com

[31] (2007). *Tools 4 Noobs, Online Summarize Tool*. [Online]. Available: https://www.tools4noobs.com/summarize/

[32] N. Andhale and L. Bewoor, ''An overview of text summarization techniques,'' in *Proc. Int. Conf. Comput. Commun. Control Autom. (ICCUBEA)*, 2016, pp. 1–7.

[33] *Dandelion*. Accessed: 2017. [Online]. Available: https://dandelion.eu/

[34] *Dandelion—Entity Extraction API*. Accessed: 2017. [Online]. Available: https://dandelion.eu/docs/api/datatxt/nex/v1/

[35] J. Wei and K. Zou, ''EDA: Easy data augmentation techniques for boosting performance on text classification tasks,'' 2019, *arXiv:1901.11196*.

[36] S. Albawi, T. A. Mohammed, and S. Al-Zawi, ''Understanding of a convolutional neural network,'' in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.

[37] Ministero De Administractiones Pùblicas. *Magerit Book I—The Method*. Accessed: 2017. [Online]. Available: https://www.pilar-tools.com/

[38] Ministero De Administractiones Pùblicas. *Pilar Glossary of Terms*. Accessed: 2017. [Online]. Available: https://www.pilar-tools.com/en/glossary/index.html