xR4DRAMA

Extended Reality For DisasteR management And Media planning

H2020-952133

D3.9

Spoken and written language analysis techniques v2

Dissemination level:	Public
Contractual date of delivery:	Month 24, 31/10/2022
Actual date of delivery:	Month 25, 18/11/2022
Work package:	WP3 Analysis and fusion of multi-modal data
Task:	T3.3 Multilingual audio and written language analysis
Type:	Demonstrator
Approval Status:	Final version
Version:	1.0
Number of pages:	33
Filename:	D3.9_xR4Drama_LanguageanAnalysisTechniquesV2_v1.0.pdf

Abstract

This deliverable describes the advanced versions and outcomes of the audio and written language analysis components of xR4DRAMA developed in T3.3 of WP3. This component is responsible for (i) the transcription of spoken language into text and (ii) the analysis of textual and spoken (i.e., speech data transcriptions) material obtained from different sources, including communication data from citizens, textual information provided by location scouts, social media messages and online information.

that may be made of the information contained therein. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

co-funded by the European Union

History

Version	Date	Reason	Revised by
0.1	27-09-2022	Table of content	Jens Grivolla
0.2	18-10-2022	Semantic parsing section	Beatriz Fisas
0.3	25-10-2022	Formatting and integration	Jens Grivolla
0.4	31-10-2022	Restructuring and formatting, introductions	Jens Grivolla
0.5	02-11-2022	Geolocation section	Alexander Shvets
0.6	05-11-2022	Speech recognition section, integration and formatting, draft for internal review by NURO	Jens Grivolla
1.0	18-11-2022	Final revision after internal review	Jens Grivolla

Author list

Organization	Name	Contact Information
UPF	Jens Grivolla	jens.grivolla@upf.edu
UPF	Beatriz Fisas	beatriz.fisas@upf.edu
UPF	Alexander Shvets	alexander.shvets@upf.edu

Executive Summary

This deliverable describes the advanced versions and outcomes of the audio and written language analysis components of xR4DRAMA developed in T3.3 of WP3. This component is responsible for (i) the transcription of spoken language into text and (ii) the analysis of textual and spoken (i.e., speech data transcriptions) material obtained from different sources, including communication data from citizens, textual information provided by location scouts, social media messages and online information. We describe technical details and evaluation results of the improved versions of the included sub-components and overall system.

Abbreviations and Acronyms

ASR	Automatic Speech Recognition
CE	Concept Extraction
DSynt	Deep Syntactic
NER	Named Entity Recognition
NLG	Natural Language Generation
POI	Point Of Interest
SSynt	Surface Syntactic
UD	Universal Dependencies

List of Figures

List of Tables

Table of Contents

1 INTRODUCTION

Multilingual audio and written language analysis are integrated in the xR4DRAMA platform to analyse the material acquired in T2.1 or provided by the users from several sources for the derivation of abstract linguistic representations, which can be used by the language generation component that will export information relevant to the needs and requirements of the users. This deliverable describes the advanced version of the techniques and methodologies for the linguistic analysis in the project.

The position of the stress level detection component in the xR4DRAMA architecture is depicted in Figure 1.

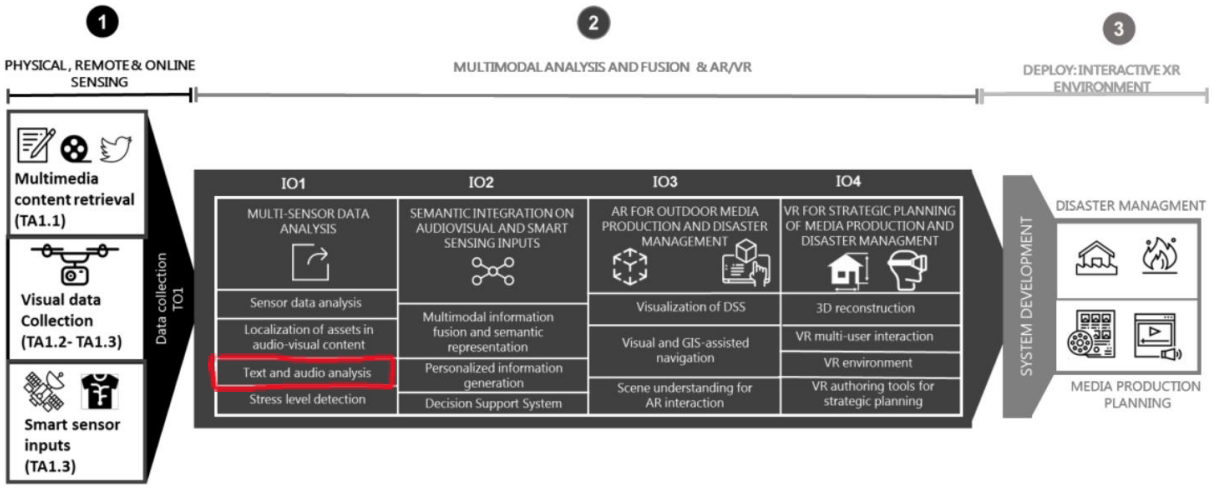

Figure 1: The text and audio analysis component in the xR4DRAMA architecture.

In Section 2 we describe the advances in Automatic Speech Recognition, with a special focus on its integration within the platform and the overall workflow.

Section 3 presents the advances in the analysis of textual and spoken (transcribed) material, in particular the extraction of domain-specific information for both use cases.

Section 4 concludes the deliverable.

2 AUTOMATIC SPEECH RECOGNITION

2.1 Task Definition and Related Work Summary

As described in more detail in D3.3, Automatic Speech Recognition (ASR) addresses the task of transcribing natural spoken language into text.

The ASR functionality is used within the xR4DRAMA project to make audio content in the form of speech recordings accessible to automatic analysis (as described in section 3), so that information received in that form can be presented efficiently without the need to have a person listen to each audio message. In the current prototype it is used primarily to obtain information from messages sent by citizens or first responders through their corresponding apps, though some work was also done on phone calls to emergency services as a possible application scenario.

2.2 Advances over the initial version

While there are constant advances and improvements in speech technology, the Wav2Letter++ framework we selected (and described in D3.3 along with project-specific evaluation results) has held up close to the state of the art over the last year, with rather good transcription quality. This has therefore remained our framework of choice, with the focus of our work mainly on the integration with the xR4DRAMA platform and making the functionality fit in the overall workflow (which has been successfully achieved).

Very recently, however, OpenAI[1] have released Whisper[2], a multilingual speech recognition system trained on 680,000 hours of speech recordings, which appears to produce impressive results for a large variety of languages. While we haven't been able to systematically evaluate it in time for this deliverable, our preliminary informal tests confirm the very high quality of results for our languages of interest. We are therefore in the process of integrating Whisper as a possible alternative to the currently integrated Wav2Letter++ system, and may use Whisper in the upcoming testing rounds.

2.3 Workflow and platform integration

The core analysis services, both for speech recognition and text analysis are packaged and deployed as Docker[3] containers, running on Docker Swarm[4]. They are accessible as REST-like[5] web services, with Swagger[6]-based documentation and an interactive web-based test interface available at **https://xr4drama.upf.edu/xr4drama-services/**.

However, integration in the xR4DRAMA platform requires some additional steps beyond exposing these core services, as detailed below.

[1] **https://openai.com/**

[2] **https://openai.com/blog/whisper/**

[3] **https://www.docker.com/**

[4] **https://docs.docker.com/engine/swarm/**

[5] **https://en.wikipedia.org/wiki/Representational_state_transfer**

[6] **https://swagger.io/**

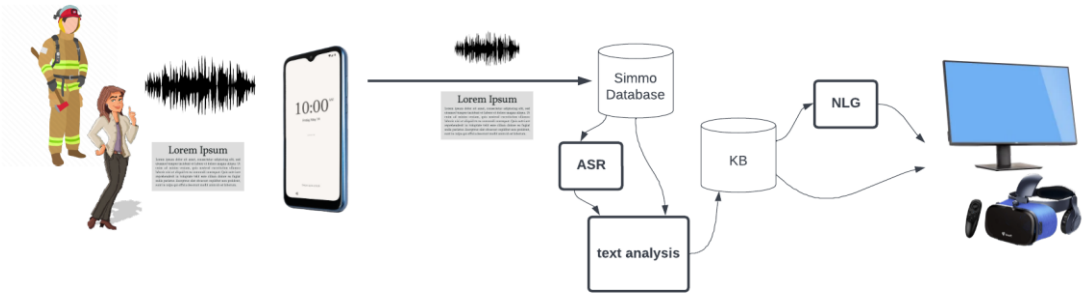

Figure 2: Information flow in the xR4DRAMA platform

As shown in Figure 2, messages from citizens or first responders are received by the crawler component that stores them in SIMMO[7] format in its database, as described in deliverables D2.2 and D2.5. The crawler puts the corresponding document in the text analysis queue using its internal identifier. The text analysis component processes its queue by retrieving each document and, depending on the document type, passing it through ASR as needed before analyzing the text and storing the extracted information in the KB. The information is then made available to the system's users through geolocated points of interest (POIs) shown on a map, or through reports generated to provide the information of interest, using the Natural Language Generation (NLG) component as required (see D3.12).

The full workflow was first successfully tested during the PUC1 pilot in Vicenza, with some refinements and improvements implemented for the later test runs.

[7] https://gitlab.com/xr4drama/xr4d_simmo

3 ANALYSIS OF TEXTUAL AND SPOKEN MATERIAL

Language Analysis in xR4DRAMA is a complex task that requires the combination of a large variety of components performing a series of steps, going from low-level linguistic analysis, such as tokenization, through higher levels of linguistic complexity, such as dependency parsing, to the extraction of entity-relation-entity triples. The language analysis module, thus, implements a pipeline of multiple components, described individually in the following subsection, each building upon the output generated by previous analysis steps.

In this section, we present the improvements over the initial version, with regard to different parts of the analysis pipeline. Most of the core pipeline described in detail in D3.3 has remained stable, with only minor fixes or improvements, in particular the linguistic analysis (segmentation, parsing, etc.) and the domain-independent semantic analysis (extraction and disambiguation of concepts, named entities, temporal expressions, etc.). This section therefore focuses on the geolocation component, which saw some improvements and adaptations to the geographical regions of interest, and the grammar-based information extraction also called "semantic parsing" which constitutes the main domain- and project-specific part of the analysis.

3.1 Geolocation

Extraction of location mentions in the messages from social media is one of the core text analysis tasks in xR4DRAMA. Knowing the spatial coordinates helps to locate the elements at risk in the reported emergency for PUC1 and the available facilities for PUC2.

The basic version of geolocation search is based on the assumption that the location mentions are presented in messages clearly in their canonical form and can be identified using generic named entity recogniser, entity linking (linking to the BabelNet entries), and linguistic dependency-based patterns. The basic algorithm to form a search query with a candidate location name consists of the following steps: i) if a place-indicating mention, such as "park", "avenue", "highway", etc. is linked via a NAME dependency to a proper name, then their concatenation is marked as a location; ii) if a BabelNet link has been obtained for a single- or multi-word mention, and it includes a reference to a DBpedia entry of the classes dbo:Place or dbo:SpatialThing, then the mention is marked as a location; iii) likewise if the mention under consideration has been tagged by the Named Entity Recognition (NER) tool as a location.

Considering real examples, we found out that proper names might be hard to detect with a standard part-of-speech tagger as they can be fully or partially written in lowercase which is often the case when the message is written in a hurry. For example, this happens in the sentence "*Traffico Vicenza ore 15:48 rotatoria viale Aldo moro direzione via quadri, strada chiusa per allagamento.*", for the location mentions "*viale Aldo moro*" and "*via quadri*". We also found out that sometimes location mentions are excluded from the main text and instead are enumerated in individual pseudo-sentences - each mention separated from others by a full stop. Since the concept extraction component in xR4DRAMA is robust to the decapitalization of words and it is also not computationally expensive to check short sentences in the geolocation index, we added two new conditions: (iv) free-form concepts connected to place-indicating mentions are marked as location candidates, (v) the entire sentence consisting of a single nominal group is a location candidate.

We also improved the search index by performing the detection of nominal subgroups in names and storing them as separate entries to account for shortened versions of names that might be used in informal language. To be able to retrieve coordinates of "*via quadri"* in the example above, which in fact corresponds to "*Via Giovanni Battista Quadri"*, we store in the inverted search index pseudo-names "*Giovanni Battista Quadri*", "*Battista Quadri*", and "*Quadri"* referring to the same object. The place-indicating mention "*via"* is used in order to rank candidates found in the search index.

Thus, we have a two-stage procedure: first, location candidate identification, and second, the search of normalized surface forms of detected candidates in a key-value index. We make use of two geographical databases, OpenStreetMap[8] and GeoNames[9], which we convert into a direct and inverted search index and extend with possible shortened versions of names. To reduce the number of candidates, we pre-filter the search index according to the perimeter of the geographical area related to the pilot (i.e., Vicenza for PUC1 and Corfu for PUC2).

3.2 Establishing the relevance for users

The messages crawled automatically from social media within xR4DRAMA may not contain information related to the interest of the users. Specifically, this happens for the Disaster Management use case where some texts discuss topics in which flood emergency events play a significant role and affect the situation in some related domains (e.g., consequences for economics), but they do not provide any useful insights for the first responders. To avoid extracting misleading information, we trained a classifier that filters out unrelated messages.

We labelled 2500 messages about floods in Italian and randomly split them into training and test datasets. Statistics are provided in Table 1.

	Training set	Test set
Relevant messages	487	53
Non-relevant messages	1712	190

Table 1: The size of the dataset for detecting the relevance of the messages

The fact that only about 25% of the crawled messages are relevant for the users already shows the complexity of the task – the keywords in short social media texts are mainly similar even though the domains can be rather different. Limiting the keywords to the very special markers may lead to the loss of very important messages of generic discourse. Therefore, we opt for gathering more but rather noisy data and further leveraging the state-of-the-art language model for data validation.

The size of the created dataset is large enough to fine-tune a pre-trained transformer-based model that already learned peculiarities of the language and requires a significantly smaller

[8] https://www.openstreetmap.org/

[9] https://www.geonames.org/

number of examples to be adjusted to the downstream task. We chose a "bert-base-italian-uncased" model[10] for Italian trained on a large 2-billion-token corpus.

The chosen model was fine-tuned only with three epochs to avoid overfitting. The results are provided in Table 2.

	Precision	Recall	F1-score	# Examples
Negative class	0.99	0.96	0.98	190
Positive class	0.88	0.98	0.93	53
Macro average	0.94	0.97	0.95	243
Weighted average	0.97	0.97	0.97	243

Table 2: Results of classification on the test set

The model was successfully trained and we reached the F1-score (which combines precision and recall into a single value) of 0.93 for the target class. It is important that the recall value is very close to 1, i.e., only one relevant message was misclassified (having several public reports on the same event shall mitigate this misclassification). At the same time, the precision of 0.88 is also high, signifying that only a small portion of messages from other domains will be analysed with semantic parsing and, therefore, the processing time of the language analysis components in the xR4DRAMA pipeline will be mainly dedicated for extracting of the information needed for the first responders.

3.3 Semantic Parsing

Semantic dependency parsing (SDP) (Oepen et al., 2014) is defined as the task of recovering sentence-internal predicate–argument relationships for all content words,

As explained in the previous deliverable 3.3, in our pipeline semantic analysis generates structured representations with a set of graph-transduction grammars that perform semantic parsing on top of the Surface Syntactic (SSynt) representations. The pipeline outputs the semantic structures at two different levels of representation: deep-syntactic (or shallow-semantic) structures and semantic structures.

Finally, an additional module maps the semantic structures in the form of predicate-argument structures onto entity-relation-entity triples.

In this second period of the project, we have focused on extending the coverage and improving the rules that extract triples for two completely different use cases: Disaster Management (PUC-1) in Italian and Outdoors Media Production (PUC-2) in English This corresponds to the last two grammars in Table 1: 8-xr4drama-triples.rl and 8b-xr4drama-triples.rl

[10] https://huggingface.co/dbmdz/bert-base-italian-uncased

Table 3 sums up the current state of the graph-transduction grammars and rules for the mapping between SSynt structures and entity-relation-entity triples through Deep Syntactic (DSynt) and semantic structures.

Grammar	#rules*	Description
0-Ud_normalization.rl	81	Normalise UD structures
1-UD_Track2_preproc.rl	103	Identify nodes to be removed Identify verbal finiteness and tense
2-UD_Track2.rl	159	Remove idiosyncratic nodes Establish correspondences with surface nodes Replace determiners, modality, aspect, and voice markers by attribute-value features Identify duplicated core dependency labels below one predicate
3-UD_postproc.rl	94	Replace duplicated argument relations Identify remaining duplicated core dependency labels
4-UD2MTT.rl	122	Assign DSynt dependencies Identify conjunct nodes
5-DSynt-Sem.rl	95	Recover shared arguments Establish coord. conjuncts as predicates
6-Sem-postproc.rl	78	Assign PredArg arguments
8-xr4drama-triples.rl	89	Map PredArg structures onto entity-relation-entity triples for PUC-1
8b-xr4drama-triples.rl	142	Map PredArg structures onto entity-relation-entity triples for PUC-2

Table 3: Graph-transduction rules mapping

3.3.1 **Extraction of triples**

Triples are a combination of three kinds of information. In our graph transducer grammars, the three elements are represented in a graphical format as two nodes and a relation between them. The first node is one of the categories of the taxonomy developed for xR4DRAMA for each use case and constitutes a key information in the environment analysed. The relation is a feature of the category needed to classify the category and the second node holds the value of this property.

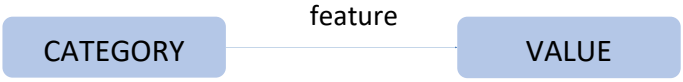

Figure 3: The structure of a triple

For xR4DRAMA we target some specific features that allow to categorize the reports from citizens according to the situation reported and to identify the elements at risk and their location for PUC1. For PUC2, we are interested in those features related to the availability and

accessibility of the specific facilities users are interested in, e.g., power outlets, access to Wi-Fi, toilets, mainly in restaurants, and cafés, etc.

This task is done by an additional component in the graph transducer that identifies the patterns in the predicate-argument structures that correspond to the targeted information and translates the entities through a simple dictionary lookup.

This process links the grammars' output representations to the Knowledge Base representations[11].

The output of the text analysis will provide structured information about the concepts defined in the taxonomy, with which Flood Reports can be generated in PUC-1 and Summaries for Media Planning in PUC-2.

Extraction of triples for the Disaster Management use case

In the flood scenario considered in this use case decision makers and first responders require as much information as possible to take the best and quickest decisions.

Text analysis aims to extract information from social media in written form. It outputs triples (Category-feature-Value) which identify situations, agents, affected objects, number of affected objects and the severity degree of the danger. The categories in the taxonomy that have been identified are mainly people, animals, infrastructure, and other objects in danger.

Rules and Dictionaries

Several types of rules have been coded according to the type of information that is searched, and the Concepticon dictionary has been populated with the concepts of relevance in the domain of flood emergencies, so different lexical units can point to one same concept.

The Concepticon is a language agnostic dictionary, developed together with the analysis grammars, in which concept labels are mapped to specific lexical units in different languages and aims to be a cross-linguistic resource that allows the interface between the semantic output of a language analysis pipeline and a Knowledge base.

For example, if a semantic representation includes the verb ESONDARE or TRACIMARE or the adjective ESONDATO, or the noun ESONDAZIONE or PIENA or STRARIPAMENTO, a lookup to the dictionary will output in all cases the label "OVERFLOW" in the conceptual representation (triples), if the rules apply to the semantic graph.

[11] https://miro.com/app/board/uXjVOwc-oig=/

```
"overflow" {
 //xR4DRAMA
 EN = {
  lex = "overflow_VB_01"
 }
 IT = {
  lex = "esondare_VB_01"
  lex = "esondato_JJ_01"
  lex = "esondazione_NN_01"
  lex = "straripamento_NN_01"
  lex = "tracimare_VB_01"
  lex = "piena_NN_01"
 }
}
```

Figure 4: Entry for “overflow” in the Concepticon

During the first half of the project, 37 rules were coded for identifying Situations, Agents, Affected Objects, number of affected objects and locations.

In this second half of the project, the number of rules has increased from 37 to 89 and new categories are identified such as the Severity degree of the situation, the Stage in which the situation is, the Final State in case a change has been produced and the potential Risks in the described scenario.

Table 4 details the types and number of rules for the mapping between SSynt structures and entity-relation-entity triples in the last module for PUC_1.

Grammar 8-xr4drama-triples.rl	#rules
Rules for identifying SITUATIONS	22
Rules for identifying AGENTS	12
Rules for identifying AFFECTED OBJECTS	17
Rules for identifying the NUMBER OF AFFECTED OBJECTS	10
Rules for identifying LOCATIONS	14
Rules for identifying the SEVERITY DEGREE of a situation	2
Rules for identifying RISKS	9
Rules for identifying the STAGE in which a situation is	2
Rules for identifying FINAL STATE in a changing situation.	1

Table 4: Breakdown of the triples extraction rules for PUC_1

Extracted categories

Situations

The main information that must be clearly identified in a report is the SITUATION to which it refers to, and that must be solved as soon as possible. In xR4DRAMA PUC1 domain situations are emergencies such as Floods, Obstructions, Overflows, Collapses….

The graph transducer identifies in the predicate-argument structures the patterns defined in the rules and looks up in the Concepticon dictionary to output the appropriate concept label in the triple. The situations may be the Argument1 of one of the predicates in a set list, or Argument2 if predicates belong to a different list, they may be verbal or nominal, and they can in some cases have a nonCore relation with another element in the pred-arg structure.

These triples are graphically represented by a first node, named Situation, a relation, named Value, and the Concept Label, retrieved from the dictionary in each case.

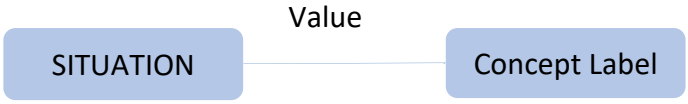

Figure 5: Situation triple structure

Affected Objects and Number of Affected Objects

As its own name suggests, the Affected Objects are those elements that are affected by the emergency situation. They might be humans, animals, infrastructure, or objects according to the taxonomy developed for xR4DRAMA by AAWA and DW[12].

As in the case of Situations, the Affected object may be the Argument1 or the Argument2 of different predicates that are defined in a list of sem values in each rule. They can also be nonCore dependents with or without a collapsed preposition. Each rule has its own particular list of sem values and all of them require the look up in the Concepticon to retrieve the correct concept label.

For example, if a rule tries to extract the Situation in a semantic structure where the node X is related to a node Y with an Argument2 relation, the rule will limit the values of the sem attribute of the node in a list to keep its application to the domain of interest, in this case, the Emergency domain.

[12] https://miro.com/app/board/uXjVOwc-oig=/

```
c:?XI {
  c:sem = ?sem
  c:lex = ?pn
  c:Argument2-> c:?Arg2 {
    c:sem = ?sem2
    c:lex = ?pn2
    c:id1 = ?id2
  }
}

(?sem == "allagare" | ?sem == "aggrappare" | ?sem == "bloccare" |
 ?sem == "colpire" | ?sem == "intrappolare"  |
 ?sem == "isolare" | ?sem == "ostruire" | ?sem == "trascinare"  |
 ?sem == "chiudere"| ?sem == "evacuare" |
 ?sem == "crollare" | ?sem == "intasare" | ?sem == "incastrare" |
 ?sem == "tamponare" | ?sem == "centrare" |
 ?sem == "esondare"
)
```

Figure 6: Part of the rule triple_situation_PredArg2, with a list of sem values

The grammar will not only identify the affected object, but also the number of the affected objects, as it is shown in the output of Ex PUC1_2, where the output graph includes a triple indicating that two persons were affected by an obstruction.

Agents

The Agent is the causer of the emergency. It may be water or mud if it is a flood, or a tree or some branches if it is an obstruction or any other element that could give rise to a hazardous situation.

Location

Although the messages received from citizens will be geolocated with position coordinates, some more information can be retrieved from the text.

Different patterns in the predicate-argument structure will match the rules that extract these triples, although most of them will include the presence of collapsed locative prepositions (a, dentro_a, in, su, dentro, sotto).

The following examples obtained from messages reported by citizens show the extracted triples with the categories as explained above.

In this example (Figure 7), an Obstruction situation is reported, its location (bridge), and the affected objects (person and car). Additionally, the number of persons affected (2).

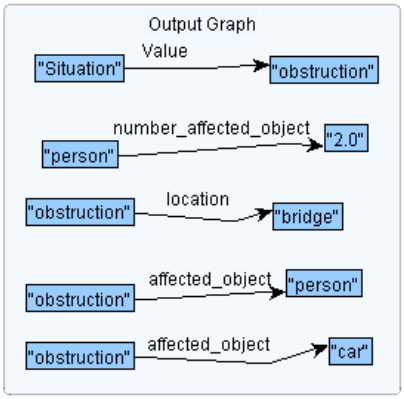

Figure 7: Output graph for *"Aiuto, c'è un'auto incastrata sotto ponte degli angeli con due persone dentro"*

In the following example (Figure 8), the citizen has reported 3 different but related emergency situations, that appear in the output with the labels “obstruction”, “flood” and “drag”, and gives some information about the affected objects (a “school bus”, a “parking lot” and a “car”) as well as which is the agent in the “drag” situation (“water”) as it is shown in the output graph.

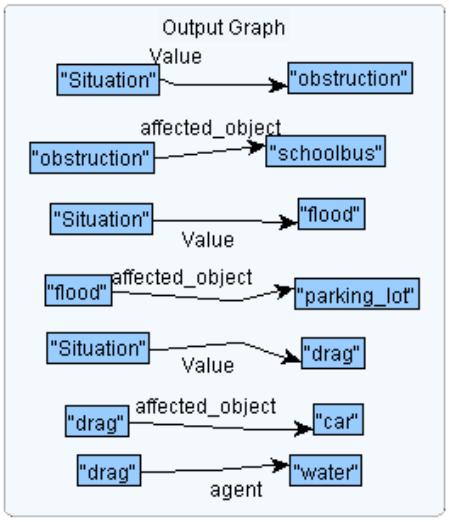

Figure 8: Output graph for *“Il parcheggio fogazzaro è allagato, ci sono macchine trascinate via dall'acqua e uno scuolabus bloccato”*

Risk

In order to differentiate a current emergency situation from a potential one, a series of rules output triples where the first node is named “Risk” (instead of “Situation”) and it is related with a Value relation to the second node, with the Concept Label that is retrieved from the Concepticon in every case.

Stage

Messages received during the emergency may give information about the stage of the situation, it may be starting, continuing, or ending. The stage is then extracted as a triple like in next example (Figure 9).

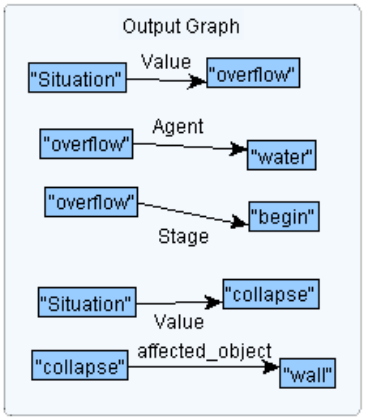

Figure 9: Output graph for *"E' crollato il muro di contenimento a monte del Ponte, e l'acqua ha iniziato a esondare"*

Final State

The stage differs from the final state in a change process. This information is also output if the Situation is Change, as it is shown in the output graph of following example (Figure 10).

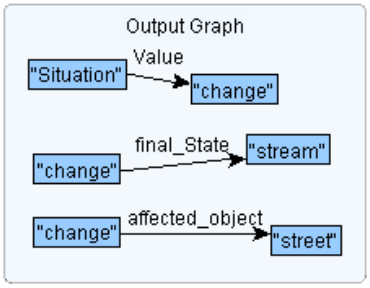

Figure 10: Output graph for *"La strada si è trasformata in un torrente carico anche di fango"*

Severity Degree

Finally, the severity degree of the emergency is also identified and extracted as a triple. The next examples extract "terrible" as the severity degree of the "traffic jam" (Figure 11), and "total" as the severity degree of the flooding situation (Figure 12).

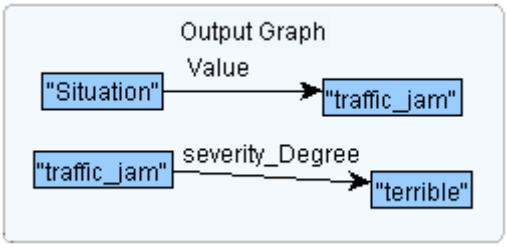

Figure 11: Output graph for *"Si è creato un ingorgo spaventoso"*

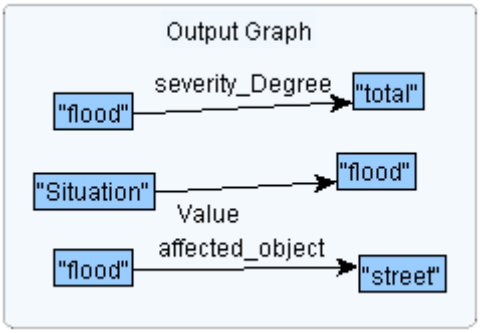

Figure 12: Output graph for *"Le strade attorno alla casa sono completamente allagate"*.

Post Processing

Affected Objects have been grouped into the categories defined in the xR4DRAMA taxonomy in a post process stage. In this way, both the fine-grained and the coarse-grained information are kept in the JSON file.

The Concepticon has been populated with these concept labels which have been grouped in the xR4DRAMA categories:

HUMAN = {baby, boy, child, citizen, cyclist, family, first responder, girl, inhabitant, man, old man, old woman, pedestrian, people, person, retired man, woman}

ANIMAL = {animal, cat, chick, dog, cow, horse, sheep, pig, goat, pony, donkey, duck, goose, hen, cock}

INFRASTRUCTURE = {apartment building , barn, beach house, coffee shop , cottage , factory, farm , house , mansion , manufactured home , office , office building , shelter , skyscraper , stable , terrace , tower , windmill , airport terminal , amphitheatre, bus station, care home, church , city hall, embassy , fire station , gas station , heliport , hospital , library, lighthouse , market, mausoleum , museum , opera house, palace , police station, post office, prison, school , stadium, stadium baseball, stadium football, subway station, theater, train station, train station platform , airport, bridge, embankment, field road , footpath, harbor, hotel , inn , motel , motorway , park , pavement , pier , plaza , power line, railroad track , restaurant , rope bridge , sidewalk , street , tunnel , underpass , viaduct , wind farm}

OBJECTS = {bicycle, boat, bus, car, drone, helicopter, lorry, minivan, motorbike, motorcycle, quad, school bus, scooter, taxi, train, tram, truck, van, vehicle, wagon, beach chair, beach umbrella, branch, ceilings, debris, manhole, rock, roof, tree, trunk, umbrella, windowpane}

Extraction of triples for the Outdoors Media Production use case

For PUC2, we are interested in extracting information from the social media (foursquare, reddit,…) about the accessibility of the specific facilities users are interested in, e.g., power outlets, toilets, drinking water, parking areas in or near to restaurants, and cafés, etc.[13] Frequently citizens post their experiences in reference to these aspects of the daily travelling life and report not only on the availability but also about the quality, location and quantity of the mentioned facilities.

[13] These have been selected according with the taxonomy in **https://miro.com/app/board/uXjVOwc-oig=/**

The aim is to contribute with the extracted triples to add real life information to the remote production planning to build a better understanding of the filming location.

Text analysis will provide this information in the format of Category-feature-value triples as described in the Disaster Management use case. Table 5 shows a breakdown of the types of rules that have been coded, and in the following subsections an explanation and examples of the extracted categories will make clear the LAS output.

Grammar 8b-xr4drama-triples.rl	#rules
Rules extracting information about PLUGS	46
Rules extracting information about WIFI	47
Rules extracting information about 3G	2
Rules extracting information about DRINKING WATER	7
Rules extracting information about PARKINGs	37
Rules extracting information about TOILETSs	3

Table 5: Breakdown of the triples extraction rules for PUC_2

Extracted categories

Power Supplies

In the extraction of information about power supplies, the goal has been to identify places where there is access to plugs so the working group in the media production can, for example, charge their phones or laptops. If there is information about the quantity, type, and location of those plugs, then this information is also extracted. The process has been thought to be applied on user or customer comments in social media as foursquare or reddit, but it can also work on tweets or Facebook posts.

Rules extracting information about Power Supply	46	
ACCESS		22
QUANTITY		8
TYPE		3
LOCATION		7
USB		6

Table 6: Breakdown of the rules for the Power Supply category

Comments in the social media can be a sum of short sentences, like in the first example in Figure 13, or a single longer sentence (Figure 14). In both cases, the LAS will go through the chained texts or the single sentence looking for any pattern that matches the rules which have been created for extracting new information of interest for the Media Production scenario.

Access to Power Supplies

Different patterns have been considered to detect the access to power supplies:

- Expletive structures (“there are plugs”)
- prepositional phrases introduced by the preposition with and its variants (“w/ power plugs”)
- sentences with the verb to have ("...it has plugs"),
- the relation Argument1 of some adjectives such as “available” with one of the elements of the plug quasi-synonym list,
- if a lexical unit of the quasi-synonym plug list is the Argument2 of a find predicate.
- Negation and coordination must also be considered as coexistent with the patterns described and the output will be different in the first case.

In the first example (Figure 13), the rules named triple_plug_yes_NP_inside and triple_plug_location_NP_inside match with the phrase if they find a node which contains a location (such as “inside” or “outside”) that has an Argument1 relation with a node with any of the sem values that refer to a power supply (plug, outlet, socket, wall plug, wall outlet, wall socket, plug adaptor, power plug, power outlet or power socket) and there is no relation with a negation element. They will then output 2 triples: Power supply –access→ yes and Power supply –location → inside.

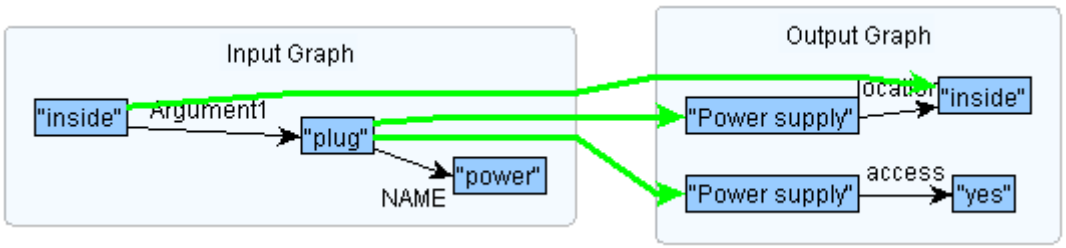

Figure 13: Mapping of pred-arg structure to triples category-relation-value for “Good music in the background. Sometimes jazzy. Nice place for studying. Power plugs inside.”[14]

In the next sentence, both a negation and a predicate with “find” are involved so the output of our grammars will show that there is NO ACCESS to a power supply.

[14] https://es.foursquare.com/v/piquet/4c7e8325fb74236ac551f7b9?tipsQuery=plugs

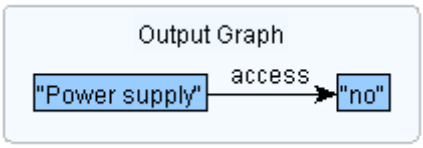

Figure 14: Output graph for "This lounge must have become worst, this lounge is quiet, but food selection is mediocre, and you can't find power outlets anywhere." [15]

Quantity of Power Supplies

Some comments in the social media include information about the quantity of power supplies that can be found in a certain location (café, restaurant, hotel…), so some rules look for this quantitative or qualitative information.

The quantity may be expressed as a cardinal number or with a quantitative expression or determiner (many, lots of, loads of, not enough, few, not many…)

For example, in the following sentence (Figure 15), the quantity is expressed with a qualitative expression "loads of…", which appears in the output with the label "many".

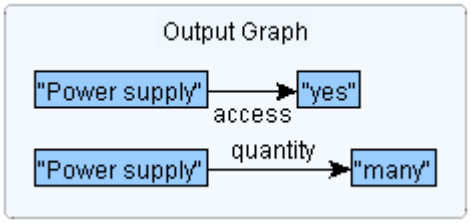

Figure 15: Output graph for *"At last, a no smoking coffee shop with loads of plugs to connect your laptop."* [16]

Types of Power Supplies

Information about some types of plugs (UK, US, EU, MK) can also be extracted if the author of the post in the social media has reported about this.

Location of the Power Supplies

Sometimes, the comments in foursquare or reddit are more detailed and they may refer to where are the power supplies located.

[15] https://es.foursquare.com/v/melina-merkouri-lounge/4da2746a63b5a35d356dfe19?tipsQuery=power%20outlets

[16] https://es.foursquare.com/v/starbucks/4d8a3f786daeb60c7d9d60e0?tipsQuery=plugs

The rules mainly search for patterns where one of the lexical units in the plug list is related to a node with a collapsed preposition indicating location (such as “in”, “at”, “by”, “on”,” towards”) or where the predicate “have” has an Argument2 which is in the plug list and an Argument1 which is the location. Also, the adverbs “inside” and “outside” are considered as possible locations, as you can see in the first example (Figure 13).

WiFi

Media Production teams are interested in knowing where they can connect to a WiFi network in the area where the filming takes place. Our LAS system will extract this information about the access, quality, and location of a WiFi network in a series of triples.

Rules extracting information about WiFi	47	
ACCESS		24
QUALITY		11
LOCATION		7
REQUIREMENT		5

Table 7: Breakdown of the rules for the WiFi category

Access and Quality of the WiFi

Rules have been developed to match a series of patterns that identify the access (or not) to a wifi:

- Expletive structures
- When WiFi (or quasi-synonyms) is Argument2 of the predicate “find”
- When WiFi (or quasi-synonym) is Argument 1 of the predicate “work” and there is a relation with a lexical unit that describes the quality of the WiFi
- If WiFi (or quasi-synonym) is qualified with any of the elements of the list: (good, great, fast, excellent, brilliant, fast, reliable, nice, easy, bad, slow, intermittent, poor, unstable, terrible
- Negative expressions (with “no”, “not”, “any”, “without”, “non-existent”)

The availability of accessing a WiFi and its quality (fast) is extracted from the following foursquare post (Figure 16).

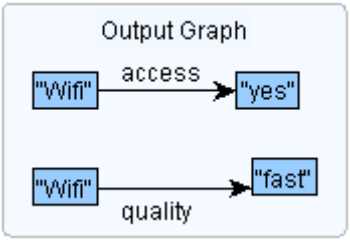

Figure 16: Output graph for "*Very friendly, free fast wifi.*[17]

Information about power supply and Wi-Fi can appear in the same social media post, as in Figure 17:

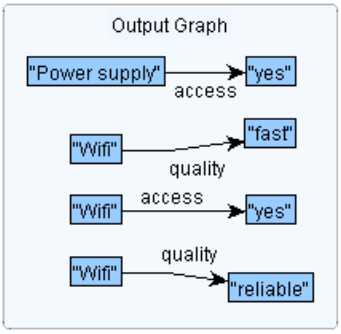

Figure 17: Extracted triples from "*Fast, reliable Internet, power plugs, and good desks for work. Great coffee and brunch. Prices slightly high*".[18]

Wi-Fi's location

As in the case of power supplies, some users report more details about the area in which one can connect to the Wi-Fi. So, some rules look for matches for this type of patterns:

- Argument1 of the predicate HAVE or OFFER, if Argument 2 is Wi-Fi or a quasi-synonym.
- A noncore relation links a location to a node whose sem value is Wi-Fi (or one of the values in the Wi-Fi list) and this node has a collapsed preposition "with"
- If Wi-Fi (or quasi synonym) is the Argument 1 of an adverb such as "downstairs" or "upstairs"

[17] https://es.foursquare.com/v/skalinada/4beedef0e8c3c9285e489992?tipsQuery=fast%20wifi

[18] https://foursquare.com/v/the-underdog/549c4bb5498e683bd43ea38c?tipsQuery=plugs

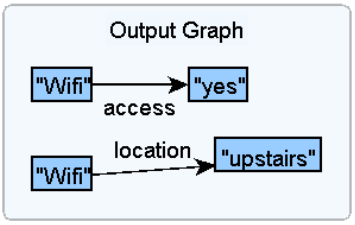

Figure 18: Extracted triples from "*You may be able to get free WiFi upstairs.*"

Requirement for WiFi

In some cases, Wi-Fi is accessible only under certain conditions, and the grammars will output this information in a triple with a REQUIREMENT relation.

For example, in Mikro Café in Corfu, the Wi-Fi is accessible only upon registration, as it is reported in the following sentence (Figure 19).

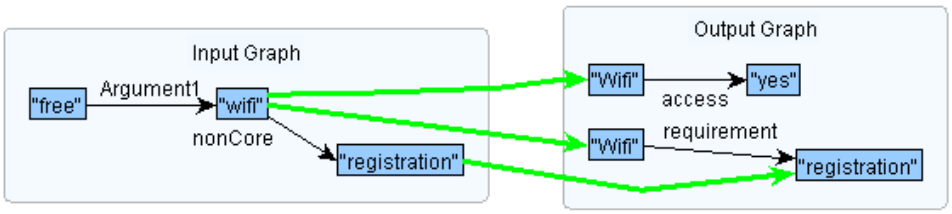

Figure 19: Mapping of pred-arg structure to triples category-relation-value for "*Perfect location to chillax while taking a coffee. Try the homemade lemonade. Free wifi upon registration.*" [19]

3G

Some comments in the social media report about the accessibility and quality of the 3G connection. So, a few rules have been coded to retrieve this information, which should be also extended to 4G and 5G connections.

Rules extracting information about 3G	2	
ACCESS		1
QUALITY		1

Table 8: Breakdown of the rules for the 3G category

[19] https://es.foursquare.com/v/%CE%BC%CE%B9%CE%BA%CF%81%CF%8C-%CE%BA%CE%B1%CF%86%CE%AD--mikro-cafe/4c2f35a266e40f474473c18b?tipsQuery=free%20wifi

Available Drinking Water

Where to find drinking water during the time the media production takes place is also an issue for the filming team. Therefore, some rules have been ready for this information extraction. Nevertheless, the social media users do not seem to post much about drinking water.

Rules extracting information about DRINKING WATER	7	
ACCESS		5
LOCATION		2

Table 9: Breakdown of the rules for the Drinking Water category

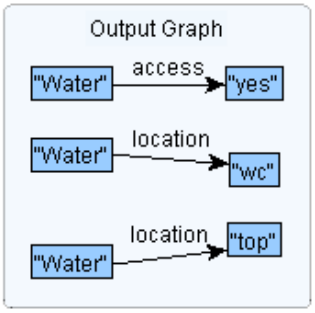

Figure 20: Extracted triples from *"12€ the adult ticket. Be careful it is slippery up there. Don't forget your bottle of water (free fountain at the top and wc). The renovation is planned to finish in 2020…"* [20]

Parking Areas

Rules extracting information about PARKINGs	37	
ACCESS		18
TYPE		8
LOCATION		9
RELATIVE POSITION		2

Table 10: Breakdown of the rules for the Parking Area category

Although the information about parking areas is also accessible in table format or even on maps [21] we have coded some rules to recover more information that users have reported in the social media.

[20] **Ακρόπολη Αθηνών (Acropolis of Athens) - Hill en Αθήνα (foursquare.com)**

[21] For example:

https://visit.corfu.gr/category/parking-areas-in-corfu-town/

https://www.google.com/maps/search/parking+corfu+town/@39.6255562,19.9126433,15z/data=!3m1!4b1

Access to Parking Area

Similarly to the rules for detecting the possible access to power supplies and Wi-Fi, some rules have been prepared to extract this information for parking areas.

So if one of the sem values in the list of quasi-synonyms for parking is the Argument 1 in an expletive sentence, or it is the Argument 1 of an adjectival predicate (limited to a closed list, that can be extended), or it is the Argument2 of a predicate "have", or it is linked to a location with a collapsed preposition, then the rules will extract a triple with an ACCESS relation.

Type or Characteristics of the Parking Area

Some characteristics reported in the social media describing the parking may be useful information to be extracted. The rules are designed to output triples where the first element is "Parking" (or a quasi-synonym), the relation is TYPE and the value can be a miscellaneous type of information: if it is small, or expensive, or it is underground or in the sun or the shade, or it is a valet parking or it is tight, or tricky…

This information could be split into different categories, but it has been named TYPE to group them and keep the same nomenclature used in the extracted information about Wi-Fi or power supplies.

Location of the Parking Area

Location rules for parking areas are similar to those coded for extracting information about where one can find a power outlet or a Wi-Fi. Rules mainly try to match with structures where one of the sem values in the quasi-synonyms parking list has a nonCore relation with a location with a collapsed preposition among a series of values ("in", "on", "at", "by", "towards", "behind") or if the predicate "have" has two arguments, Argument1 being a location and Argument2 is one of the possible parking sem values.

Relative position of the Parking Area

Finally, in this category of information (Parking), the information conveyed by the collapsed preposition might be relevant so a new type of relation between the 2 nodes in the triple has been defined and named RELATIVE POSITION.

In the following example (Figure 21) you can see that not only knowing that the hotel (location) has a parking, is a useful information for the production team, but also that the parking is BEHIND (relative position) the hotel.

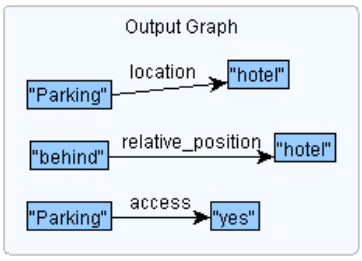

Figure 21: Extracted triples from *"There is a parking lot just behind the hotel"*

Toilets

Rules extracting information about TOILETS	3	
ACCESS		2
TYPE		1

Table 11: Breakdown of the rules for the Toilets category

When travelling to film in outdoors settings, finding toilets might be an issue for the filming team.

As with parking areas, this information is accessible in some webs[22], although it is usually not in a textual form but rather in a graphical form on maps.

A few rules have been coded to extract some more information from the users if it has been reported in the social media posts: if there are toilets in a location (ACCESS) or if there are any comments about them (named as TYPE).

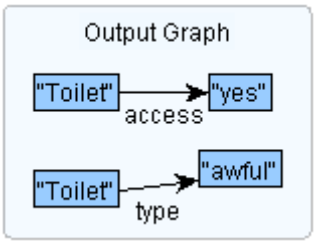

Figure 22: Extracted triples with information about toilets from *"Long sandy beach. 7 eur 2 sunbeds and umbrella. Showers and change cabin. Awful toilet. Sunday can be quite busy.*[23]

Slight changes in the sentence will produce different output, as it is shown in the following example, where instead of "awful toilet" the post reports "awfully busy toilet":

[22] For example: https://pee.place/en/p/public-toilet/fiKhwsk6qTCqnvgyc

[23] https://es.foursquare.com/v/%CF%80%CE%B1%CF%81%CE%B1%CE%BB%CE%AF%CE%B1-%CE%AF%CF%83%CF%83%CE%BF%CF%85-issos-beach/4c5a9725ec2520a104c95112#:~:text=Long%20sandy%20beach.%207%20eur%202%20sunbeds%20and%20umbrella.%20Showers%20and%20change%20cabin.%20Awful%20toilet.%20Sunday%20can%20be%20quite%20busy.

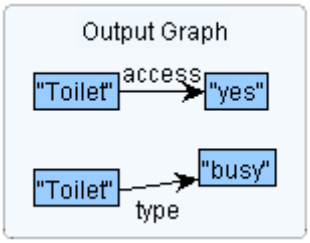

Figure 23: Extracted triples from "*Long sandy beach. 7 eur 2 sunbeds and umbrella. Showers and change cabin. Awfully busy toilet. Sunday can be quite busy*"

3.3.2 **Summary and Improvements in the Language Analysis System**

To sum up, the LAS improvements during the second half of the project have been:

- the extension of the coverage of the extraction of information in the form of triples, where a category is linked through a relation to a concept label, for the Disaster Management scenario (PUC1).
  - The number of rules has increased from 37 rules to 89 rules,
  - Some categories already covered in the first half of the project (Situations, Agents, Affected Objects, Number of Affected Objects and Locations) have been enhanced
  - Other categories have been added: Severity Degree, Risk, Stage, Final State
  - The Concepticon dictionary has been populated with the concept labels and corresponding lexical units needed for the extraction process
- the development of the 142 rules to extract information for the Outdoors Media Production scenario
  - The categories that have been searched are: Power Supplies, Wifi, Drinking Water, Parking Areas and Toilets
  - The features of interest have been: Availability, Type, Quantity, Quality, Location and Relative Position, depending of each category.

4 CONCLUSIONS

In this deliverable we presented the advanced version of the xR4DRAMA language technology modules for Automatic Speech Recognition (Section 2) and Language Analysis (Section 3), along with their integration within the xR4DRAMA platform. The modules now extract relevant information for both use cases, from both written and spoken sources. The extracted information is made available to users through textual reports created with Natural Language Generation (NLG) techniques (as described in D3.12), or as points-of-interest (POIs) that can be viewed on a map in the different xR4DRAMA applications, again with the support of the NLG module, or as searchable metadata associated with the POIs.