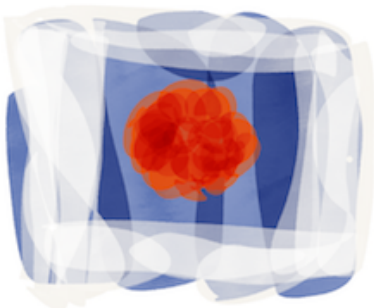
Deliverables 4.2 and 4.3

Lessons learned from documenting, with metadata, local information

Lessons learned from using mapping and geo referencing for the representation of local climate information

Author(s) and affiliation(s)	Date	Version
Didier Peeters[1] Anne De Rudder[2] Juan Baztan[3] Charlotte Da Cunha[3] [1] Université Libre de Bruxelles (ULB) [2] Institut royal d'Aéronomie Spatiale de Belgique (IASB) [3] Université de Versailles Saint-Quentin-en-Yvelines (UVSQ)	10/6/2021	1.0



The CoCliServ project benefits from funding obtained through the ERA4CS Joint Call on Researching and Advancing Climate Services Development.

CoCliServ is funded by the following national funding agencies: Agence Nationale de la Recherche (**ANR**), France; Service public fédéral de programmation politique scientifique (**BELSPO**), Belgium; Deutsches Zentrum für Luft- und Raumfahrt EV (**DLR**), Germany; Nederlandse organisatie voor wetenschappelijk onderzoek (**NWO**), the Netherlands; Norges forskningsrad (**RCN**), Norway.

Table of contents

Table of contents 2

Executive summary 4

Purpose of the document 4

Relationship to the Description of Work (DoW)....4

Foreword 5

1. Lessons learnt from building up a conceptual metadata scheme for narratives 5

1.1. Scope....6

1.2. Granularity 6

1.3. Confidentiality....9

1.4. Functionality....9

1.4.1. Traceability 10

1.4.2. Visualisation and dissemination 10

1.4.3. Comparing, sorting, aggregating 11

1.4.4. Data analysis 11

1.4.4.1. Narratives of change.... 12

1.4.4.2. Scenario design and development.... 12

1.4.4.3. Local climate information in context 12

1.4.4.4. Knowledge quality assessment 13

1.5. Compromises, challenges, limitations.... 14

1.5.1. Feeding the metadata scheme 14

1.5.2. Open evolutive model versus frozen form 15

1.5.3. Subjectivity of narratives 15

1.6. Conclusion 16

2. Lessons learned from using a GIS tool in the CoCliServ context 16

2.1. The challenge of storing social science information as data 17

2.2. The quality or ease of use dilemma 18

2.3. The development of the database 19

2.4. QGis 20

2.5. QGis/PostGis interconnection 21

2.6. The users 23

2.7. The data 24

2.8. The locations in the *informions* 26

2.9. Using the tool 27

2.10. Conclusion 30

References 31

Annex – CoCliServ metadata for narratives 32

Executive summary

CoCliServ developed and explored novel ways to transform state-of-the-art climate science into action-oriented place-based climate services. In this context, the project proposed to explore how metadata and Geographic Information Systems (GIS) could contribute to data collection and production in social sciences. This document summarises the metadata standardisation and mapping activities (Tasks 4.2 and 4.3) led during the project in an attempt to characterise, catalogue and represent graphically relevant information conveyed by the collected narratives, and derives considerations from this work.

Purpose of the document

The purpose of the document is to draw some thoughts from the documentation and mapping exercises conducted throughout the project, that could be of some use in future similar projects.

Relationship to the Description of Work (DoW)

D4.2 and D4.3 are expected deliverables of CoCliServ WP 4 (Local representations of a changing climate), respectively resulting from Task 4.2 (metadata standardisation) and Task 4.3 (dynamic mapping).

Foreword

Initially included in the Description of Work (DoW) to characterise climate data, Task 4.2 was, right at the project start, re-directed to describe narratives. Climate metadata standards already exist and are widely documented. It was estimated more interesting to explore the capacity of metadata to characterise the basic objects created and handled by the project, i.e. the collected narratives, which are its building blocks. This original approach established a natural bond to Task 4.3, the mapping activity, conceived as a way to provide visual glimpses of the information conveyed by the narratives. Since the graphic tool needed some underlying (meta)database to feed its various layers anyway, it offered a technical solution to the metadata encoding and storage. With this approach in mind, the coupled activities were outlined in the document achieving Milestone M4.1, *Available modes of representation*.

1. Lessons learnt from building up a conceptual metadata scheme for narratives

Starting from a basic skeleton, the CoCliServ narrative metadata scheme was further developed and took shape progressively, as discussions went on with the site and work package leaders. The final result (which would probably have evolved further, would the project have continued) is presented in the Annex – CoCliServ metadata for narratives – at the end of the current document.

Referring to the scheme in annex, the following sections account for different aspects of it that we had to examine in the course of the work. Along the way, they highlight some challenges we had to face and choices we had to make.

1.1. Scope

One of the first questions that arose was the type of objects that had to be described and categorised to serve the project objectives while exploring innovative ways forward. As the project elementary building blocks, narratives or sets of narratives were the most obvious candidates, but the possibility to create metadata for other types or sets of objects studied by CoCliServ was also envisaged. Such objects included scenarios (derived from the narratives) and chronotopes (understood as physical or conceptual space-time milestones). It was even proposed to develop metadata for managerial objects such as deliverables or project milestones. However, creating metadata for narratives revealed ample and challenging enough to monopolise the human resources in presence, and efforts exclusively focused on this task.

1.2. Granularity

The initial metadata scheme proposed in the *Available modes of representation* document considered three levels of metadata, corresponding to three levels of aggregation:

level 3 – metadata characterising any narrative;

level 2 – metadata characterising the set of narratives from one of the five project sites;

level 1 – metadata characterising the entire set of narratives collected by the project.

It soon appeared that in the context of a single project, only the level 3 metadata offered any interest, and Task 4.2 focused on the development of a metadata scheme where the described objects were the individual narratives.

In the context of Task 4.2, narratives were conceived as the original tales collected by the site leading teams: one narrative would be a story told, written, painted, illustrated, discovered by one individual or collectivity, who could be anonymous but would generally have a definite identity. Metadata were developed with this relatively non-constraining definition of a narrative in mind.

Toward the end of the project, one of the partners underlined the fact that what she called narratives were synthetic stories derived from a set of similar individual testimonials, retaining their common prominent features. In other words, narratives were not the individual pieces of raw material collected, but already resulted from some processing of this material, highlighting recurrent typical concerns through some disembodied or fictive story. It made more sense, for her work, to characterise such narratives than to describe in detail every singular story.

The point she raised was actually a matter of granularity, since the synthetic narratives she had in mind aggregated, somehow, sets of similar individual narratives. A large number of metadata fields required to characterise singular narratives, especially among the pieces of *contextual information* listed in the annex (such as the identity of the narrative's author or the place and circumstances where and when the narrative was collected), would not be relevant for synthetic narratives any longer.

There was no time left to change course, so that Task 4.2 stuck to the approach pursued up to that point. But it would be worth, for future projects similar to CoCliServ integrating a metadata application, to devote some time, at the beginning, to questioning which levels of granularity are desirable or useful for the work undertaken. In contexts where traceability (and hence posterity) is key, documenting individual pieces of raw material sounds like an appropriate decision. Where the focus is to highlight prominent features arising from some stage of the analysis, characterising datasets or abstractions derived from particular observations may be preferred.

That said, as will be examined more closely in the Provided the personal aspect of most narratives to be collected, it was decided at the start that their contents would not be divulgated, not even to the project co-partners, apart from quotes representative of climate-related local concerns.

At the beginning, the intention to represent the narratives using metadata was met with reluctance by some partners, who feared that this approach would violate the confidentiality of information entrusted to them. They were however reassured by the possibility to tag selected metadata fields with a "confidential" stamp that would prevent them from being seen by internal or external users of the metadatabase: personal details about the authors, stored in the metadata, would be kept unveiled, as well as any information likely to reveal their identity. As will be seen in the second part of this document, the mapping tool developed at ULB offered easy technical solutions to achieve confidentiality, such as fuzzy geo-referencing.

Functionality **Erreur ! Source du renvoi introuvable.**section below (Comparing, sorting, aggregating), one may observe that metadata of singular objects may serve to filter or group these objects according to the criteria that are relevant for the defined research; that is, they may help to construct emblematic abstractions like synthetic narratives.

1.3. Confidentiality

Provided the personal aspect of most narratives to be collected, it was decided at the start that their contents would not be divulgated, not even to the project co-partners, apart from quotes representative of climate-related local concerns.

At the beginning, the intention to represent the narratives using metadata was met with reluctance by some partners, who feared that this approach would violate the confidentiality of information entrusted to them. They were however reassured by the possibility to tag selected metadata fields with a “confidential” stamp that would prevent them from being seen by internal or external users of the metadatabase: personal details about the authors, stored in the metadata, would be kept unveiled, as well as any information likely to reveal their identity. As will be seen in the second part of this document, the mapping tool developed at ULB offered easy technical solutions to achieve confidentiality, such as fuzzy geo-referencing.

1.4. Functionality

Beside the interest that may – or may not – lye in enunciating and naming for their own sake, metadata may serve a number of purposes.

In climate science, for example, where data are the numeric results of physical measurements, it has become traditional to distinguish between *discovery*, format and intrinsic metadata. Discovery metadata encompass elements informing on the existence, location and availability of the data, making the search for – and, but not mandatorily, the access to – the data possible, usually through a catalogue. Format metadata, especially relevant when data are stored in electronic files, inform on the data formal display, providing guidance on how to read the data once downloaded. Intrinsic metadata pertain to the data nature, allowing their understanding and informed use. One of the key pieces of information stored as intrinsic metadata is the data uncertainty, i.e. an assessment of their *closeness to truth*.

Although inspired from natural science practice and from libraries, CoCliServ metadata do not have a similar vocation, since the data collected by its partners are not, as such, intended for public dissemination. Yet, to the extent that they provide a common wording framework and support the data analysis, they can be far more than a gimmick for attractive representations.

1.4.1. Traceability

Metadata are a convenient instrument to keep, in a concise homogeneous form, the memory of the work accomplished, of the methods applied and of the original elements (here, the narratives) that led to the project conclusions. When used in a systematic way throughout a project, they can provide accurate information to later researchers interested in the subject, while avoiding diving into the bits and bobs of the actual stories, which are confidential anyway.

1.4.2. Visualisation and dissemination

By standardising the description of narratives with words and by providing the underlying information to the geographic mapping tool developed by Task 4.3, narrative metadata contributed to WP4 (Representations) and had a potential to support, through visual representations, the dissemination of the project outcome ensured by WP6 (Coordination and dissemination).

1.4.3. Comparing, sorting, aggregating

Metadata are an instrument to standardise the way an object is described. In retaining some features and excluding other ones, they somehow disarticulate the objects under study. When these objects are stories enunciated by humans, anthropologists may rightfully dislike the method, since it reduces vivid accounts to formal lists of factual elements and discards any unforeseen consideration – that would possibly throw some original light on the question tackled. The critic is justified. Yet, when it comes to forge a general image of the various views or wishes of people belonging to a certain community, or to establish statistics, metadata offer an appropriate medium to compare narratives to each other based on common criteria, sort them out in categories and aggregate them into archetypical narratives, as already mentioned.

1.4.4. Data analysis

When the work on metadata started, it was debated whether they should be used for other purposes than recording basic facts about the narratives. However, their potential to support the analysis performed in the various work packages became clear as the project proceeded. The scheme was progressively extended to include metadata fields informing the researchers on features relating to their part of the study. At the work package leaders' demand, the *analytical information* (see Annex), which at first only included an abstract, keywords, names of places mentioned by the narrator and quotes, was successively completed by the following metadata fields to support WP1 (Narratives of change), WP2 (Scenario design and development), WP3 (Local climate information in context) and WP5 (Knowledge quality assessment). In the end, this enlarged functionality of the metadata was not exploited, but it was recognized by the partners as a potentially useful analysis tool.

1.4.4.1. Narratives of change

Metadata value: *Yes / No*.

Mention of the future	desired future
	feared future
	without hope or fear overtone
Mention of climate	past climate
	present climate
	possible future climate impacts

1.4.4.2. Scenario design and development

Metadata value: tick / no tick.

Potential for a hinge point	unclear	
	no	
	yes	climate-related
		not climate-related
		relationship to climate undecidable
		can be controlled
		cannot be controlled
		control possibility unclear

1.4.4.3. Local climate information in context

Metadata value: tick / no tick.

Mention of meteorological physical parameter	air temperature
	sea surface temperature

	soil temperature	
	air humidity	
	cloudiness	
	precipitation (rainfall, snow, hail)	
	wind (speed and/or direction)	
Mention of climate-related phenomenon, event or landscape feature	storm(s)	
	drought(s)	
	natural fire(s)	
	flood(s)	
	freshwater, river system	
	sea level	
	coast line	
	salinity	
	other	please state
Mention of life aspect in relation to climate	air quality	
	water supply	
	agriculture	
	food supply	
	energy supply	
	ecology	
	vegetation, forestry	
	animal life	
	health	
	economy	
	education	
	culture	
	human population density	
	human migration	
	peace	
	politics	
	other	please state

1.4.4.4. Knowledge quality assessment

Metadata value: free text.

Knowledge types on which analysis and/or prediction of the site future is based	type 1	knowledge type
		reasons for trust

		identified uncertainties
	type 2	knowledge type
		reasons for trust
		identified uncertainties
	type 3	knowledge type
		reasons for trust
		identified uncertainties
Knowledge quality aspects	mentioned knowledge quality criteria	
	mentioned knowledge issues	

1.5. Compromises, challenges, limitations

Constructing and using a metadata model and a metadata tool such as the ones developed by Tasks 4.2 and 4.3 is not necessarily straightforward. The time-consuming aspects of preliminary tasks may mask the interest of the hypothetical future benefits. Some of the issues and challenges we met are addressed below.

1.5.1. Feeding the metadata scheme

Encoding metadata for the whole lot of narratives revealed too tedious a task for most partners, whether via the Excel spreadsheets provided or directly into the QGis-based tool developed by Task 4.3, even with the recourse to templates that could considerably simplify the task. Later advantages remained too unclear – and were indeed not demonstrated – to convince the project partners to really dive into this effort. The timing was also not ideal to do so, since the narrative collection and the metadata scheme proceeded and matured in parallel. The site leaders had already started to use their traditional methods of analysis when the possibilities that the metadata offered in this respect appeared more clearly. However, the prospective of visual representations of results was an incentive.

1.5.2. Open evolutive model versus frozen form

In many respects, we opted in this project for a resolutely open metadata model:

- no field was mandatory;
- new options could be added to drop-down menus;
- new metadata fields could be proposed;
- numerous fields accepted free text.

The obvious advantages of such a model are its flexibility, the freedom left to the metadata provider and the high degree of nuances that can be expressed. But this approach has also strong drawbacks. The fact that the scheme expands as time goes by hinders the comparison of narratives for which metadata have been recorded at different stages of its life. Empty fields may result in discarding a narrative from statistical studies. Free text cannot be automatically read and makes comparison between various narratives arduous. To guarantee a sensible information content, an easy analysis and a minimal comparability, a balance between flexibility and rigor must be achieved. Flexibility may be preferred as long as the scheme is under development (which was actually the case during CoCliServ) but at some point, the model must be frozen in order to be useable.

1.5.3. Subjectivity of narratives

Narratives are of a subjective nature. This makes them essentially different from the data treated by the physical sciences for which metadata schemes have been established for a long time. But even more crucially, the subjectivity of the researcher who will document a narrative will influence the metadata content, which will usually not be the case for metadata pertaining to physical measurements. Whether a number lies under a given threshold, for example, will not depend on the metadata provider, whereas an element of a narrative may be interpreted by one researcher as an allusion to a potential climate-related hinge point, and not by one of her/his colleagues. This provides metadata for narratives a relative nature of which users must be aware.

1.6. Conclusion

Due in part to the novelty of the approach, in part to the very different styles of practice, concepts and vocabulary of the researchers in their respective fields, developing a metadata scheme for narratives involved substantial communication and lasted for the duration of the project. The scheme that was progressively completed to meet the partners' expectations was only tested on some examples, not numerous enough to conclude if the concept was adequate, even less to document all the collected narratives and serve as an analysis or dissemination tool. Yet, it is hoped that the outcome of this pioneering exercise can be the starting point of future other projects.

2. Lessons learned from using a GIS tool in the CoCliServ context

Deliverable D4.3 presents the processes and results of Task 4.3 (*Dynamic mapping*) initially described as : "develop a spatial database, in relation with regular databases, to support WP1-2-3 activities with mapping: spatial representation of global trends for use in WP1-2-3; interactive dynamic visualisation support tool for WP1-2-3, sharing WP1-2-3 output between case studies (if needed) and WPs, graphical interface for visualizing spatial objects and for database interactions, provide and centralise material for mapping purposes."

The resulting work is a tool made to be used by regular users in a research context, allowing, thanks to the metadata standardisation achieved in Task 4.2, to share (anonymous) data between study sites, between work packages or with other analogous projects.

2.1. The challenge of storing social science information as data

GIS brings together various more or less sophisticated techniques of which cartography is probably the best known. Cartography produces rich and relatively easy-to-understand graphics to enlighten a discourse, but also opens the way to certain analyses by allowing the spatial dimension of studied problems to be taken into account. Geographers start by mapping a phenomenon and then use this spatialization to look for correlations with other aspects or other phenomena also mapped. As soon as the mapping is done, the locations are digitized and the associated data can benefit from computer processing of the digitized data. Although the process is common in human geography, it is only exceptionally used in the other human sciences. By attempting a hybridization between humanities and natural sciences, the CoCliServ project wanted to explore different unusual paths, including GIS.

The ambition of CoCliServ in relation of mapping was more than just produce beautiful illustrative maps for brightening up reports or to support meeting discussions with local communities. These were made as well of course, maps are a common tool in scientific discussions about environmental issues, and are presented in other deliverables. The plan was to explore the feasibility of producing a tool to dynamically share field observations from many types of sources (in fact, not only narratives) and ensure a maximum preservation of the collected information. In social science not every information can be saved and standardized, there is always a limit, even literature fails at capturing the ultimate personal feelings.

The ambition here was to develop a database model to push further the limits of the data capture from informal social sciences outputs. Therefore our first logical step was to elaborate a metadata scheme with the field investigators (see task 4.2), the second was to produce a database model and the third one was to make a user-friendly interface that could be used by different types of users, dealing with different types of information. Of course this three theoretical steps were explored simultaneously with lots of back and forth during the project. "Whether such a dream may really be a nightmare is another topic" (Peuquet & Marble, 1990).

2.2. The quality or ease of use dilemma

As explained above, we needed to marry an IT tool with users who are not familiar with such technology in their usual activities. It has been tried before of course by other researchers (see Caquard, 2011, or Vivant et. al, 2014), and it is a never ending task, since on one side the needs are always too specific and on the other side a versatile solution would most probably require a high level of computer literacy as it would remain something very abstract. The challenge is therefore to find the right balance between solutions that are very simple to grasp but quickly overwhelmed by the difficulty of taking into account the specificities of human reality, and very sophisticated solutions with a very good potential for abstraction but which are too complicated to learn and use.

We have come to a (temporary ?) conclusion that mixing those opposite constraints might be feasible if:

- users are ready for a little learning.
- there are a few users responsible for some of the more difficult tasks
- someone is there to manage the system on a technical level

And our proposition is described below.

2.3. The development of the database

So the central element of our system is a database. The concept of "database" ranges from any set of data, like an address book or an MS Excel file for example, to a computer highly mathematical structure like the well known Oracle, MS SQL or MySQL, which are (Relational) Data Base Management Systems (RDBMS). Because they are very efficient tools these systems are now everywhere on all computers or smartphones, in all softwares and most web pages rely on them. They ensure a great safety for the data, they are very resistant to data corruption, they control very efficiently the accesses and allow to define specific privileges to the different users.

As said earlier this is the kind of tool we are using and we have chosen PostgreSQL which is also the choice of many important businesses and administrations around the world (like Google Earth, the French and Swiss statistical national institutes). PostgreSQL itself can manage regular data (numbers, text, dates, ...) and it has an extension named PostGis that gives it the ability to manage geographical or geometric data (points, lines, polygons). PostgreSQL complies with open standards, is open source itself (meaning it does not depend on private technologies), uses the SQL language. A database developed inside PostgreSQL can therefore be managed by anyone having access to it and having the skills for it. The schema itself (the database structure and all the functions added) can be exported as a pure text file and installed on any other PostgreSQL host, or transformed in another RDBMS format. The data can be exported as well together with the schema or in a single compressed file (a "dump") or in individual tabular files (like "csv").

So we created a database containing tables and functions[1]. The general structure of the tables are based on the work made in the metadata part of CoCliServ; they are of course a little adapted to comply with the SQL logic and with the interface tool we made with QGis (see below). The database structure remains simple. The functions were added to ease the work of the user, either by adding some information like the name of the user, the date of the recording, the study site based on the location, re-organise the stage numbers inside a recorded journey (see below), etc, or to do some post-processing, like inserting new values to drop-down lists. Without these code elements, data recording would be more complicated and more error-prone.

2.4. QGis

As presented above, QGis is a free and open source software, quite easy to use and very rich in functionalities. It can be used to create regular maps but offers also lots of analytical functionalities, automation of processes, a programming environment in Python. It can use many different types of data and of course it can connect to database servers like PostgreSQL/PostGis.

So the data are either stored in files or in connected servers and are all organized in layers put one above the others, like in an image editor. The data in the main window are overlapping the ones from the layers below.

[1] The functions are written in Pl/Pgsql and the tables are in SQL.

Among its user-friendly tools there is the ability to create forms connected to a data layer, so that by clicking on a feature belonging to a layer a window will open and display the specific attributes of that feature in a way pre-defined in the layer properties. This is the functionality we use in our tools, knowing that the form can contain many tabs, containing themselves several blocks, that can be hidden or displayed according to the value of other attributes. So the form will evolve along with the filling of the attributes by the user, hiding blocks of attributes that would not be relevant anymore and showing the ones that could for instance be useful to add some more relevant information. For example imagine that the user is recording information about a book, then the block of data containing everything needed to describe a book would appear (title, author, publication date, content description, etc) while the blocks containing the attributes useful for describing a migration journey would be hidden. The form is versatile.

In QGis all the data are organized in a 'project' which is where all the information regarding the list of layers, the formatting (the symbols, the colors, the forms, etc) are stored. The layers (the 'real' data) are stored elsewhere, in files or in web servers. Since we are working with a dedicated server the data and the project itself are also stored on the server (except the base maps which are provided by public servers like OSM or Google), so that the user can change of computer without having to carry everything with him/her, all he/she needs is to remember his/her credentials.

2.5. QGis/PostGis interconnection

As is often the case in IT, the tasks are shared between front-end and back-end. Qgis offers the front-end, i.e. the system with which the user interacts and which contains a minimum of the logic, while PostGis ensures the real data management.

In this way, we could, if necessary and with minimal workload, replace or supplement QGis with another system, such as a smartphone or tablet application, and continue to work with the same database. We could also connect analytical tools, such as statistical software, to PostGis and perform analysis or mix data with other sources. Access is concurrent, which means that different users can work at the same time doing similar or different things. Still for now we use QGis because it fits the project requirements which were above all cartographic.

So the work would go like this:

- The user opens QGis and the CoCliserv project, which opens the connection to the server, which then identifies the user,
- The user chooses a task by selecting it in a list within a little window, like recording a new case (whatever it is) or updating one or adding information.
- If the task is recording a new case the user starts by clicking somewhere on the map, supposedly on a place in relation with the case, otherwise the user clicks on the case to modify, then a form predefined for the kind of task opens.
- At the end the user saves its work by clicking on 'save' then the database processes the data which are then organized according to the database structure. Possibly some symbol appears or changes on the map, on the user' screen and on any other user's.

A QGis project is personal to each user. The users are free to modify it, like replace symbols with others, change the colors, even the dialog in the forms, as long as he/she knows what he/she is doing. Having the projects stored on the server allows the system manager to create a new project whenever a new user joins, or alter existing projects whenever new functionalities are added or issues are fixed.

2.6. The users

We have mentioned above the 'users' but we need to clarify who they are and what they are allowed or invited to do. There are fundamentally three kinds of users:

- The regular users who are simply using QGis and the CoCliServ tools, and who only need to understand the basic use of QGis and who can only use the system 'as is'.
- 'Power users' with the rights to modify the system parameters,
- The system manager, who fixes the tools and who is above all in charge of the server, and therefore must have a good technical background with spatial databases and GIS.

Of course daily work can be done by regular users without any assistance, possibly there may be a need for power users and the system manager is only required for developing the tools or adding new functionalities.

The 'Power users' can change the system parameters because they are allowed to add/modify content inside specific tables. For instance there is a table containing the different languages the data can refer to, or a table with allowed keywords, another one with media that can be used for recording purpose, etc, and the need might appear to add one item in one of those tables, in order to make it available for use.

One of the features that was requested and that we included in the tool was the ability to extend easily a predefined list of categories: the interface of the tool offers fields to be filled, and some offer the possibility to make a selection between different values, but it can happen that the correct answer is not included in the list and therefore we have added the ability for the regular user to add new values that will immediately be added to the default list for the other users as well. The advantage is that these predefined lists are progressively elaborated by the users themselves but the drawback is that it could lead to duplicates or the to the inclusion of not relevant answers. This feature should therefore be evaluated carefully before enabling it or not.

2.7. The data

As said above, a list of attributes was elaborated in the metadata part of this work package, and the database reflects accurately this list but the general organization of the information is adapted to comply with the relational model and the QGis interface requirements.

First of all, the reader might be confused by the use of the words 'data' and 'metadata'. The metadata are nothing else but data about data, meaning usually that they define a set of data by specifying the source of the data, the time they were elaborated, the context of the collection, etc. In CoCliServ we elaborated an important set of metadata in order to describe - initially - each narrative that would be collected in the work package 1 because we didn't want to store and publish those narratives as is. In short, the metadata would describe who was interviewed, where it happened, when, by whom, in what context, what were the topics, what was the feeling about the climate topic, etc. So instead of having metadata describing a set of data, we have metadata describing each item in our collection, and we don't store the data itself which would be in the case of a narrative a conversation. And these metadata are our data. In our tool none of the attributes are mandatory, especially the identity can be a pseudo. The location recorded in the tool can also be altered afterward in order to scramble the information and strengthen anonymity.

It is important to note that the information collected is not necessarily the information published afterward.

The metadata model elaborated contains an important number of attributes which makes the work with a graphical interface difficult because of the need to display all the fields to fill in a single form, and tricks had to be found to bypass the cluttering of the display.

A second challenge came from taking into account sources other than narratives, like books, tv broadcast, news in newspapers, and so on. Some attributes can fit several types of sources but not all, like for instance a ISBN number. We had therefore to elaborate a form that can adapt itself according to an initial choice of source.

This is when we invented the generic word "*informion*" to designate the different sources of information. A narrative is an informion, a book is an informion, a paper is an informion, well at least they are in our database and our QGis interface.

Figure 1. List of glossaries.

So we have informions, and we have other types of data in our tables, i.e. what we call the parameters and the glossaries. As mentioned above in the users description the glossaries contain the possible answers to questions when there is a limited list of possible answers, they typically are what we see in dropdown lists. These are important because they somehow limit the freedom of the user, in order to avoid apparently different answers which are in fact synonyms, or out of topic answers. Let's remind here that no question in our tool is mandatory and the tool is made to record *meta*data only.

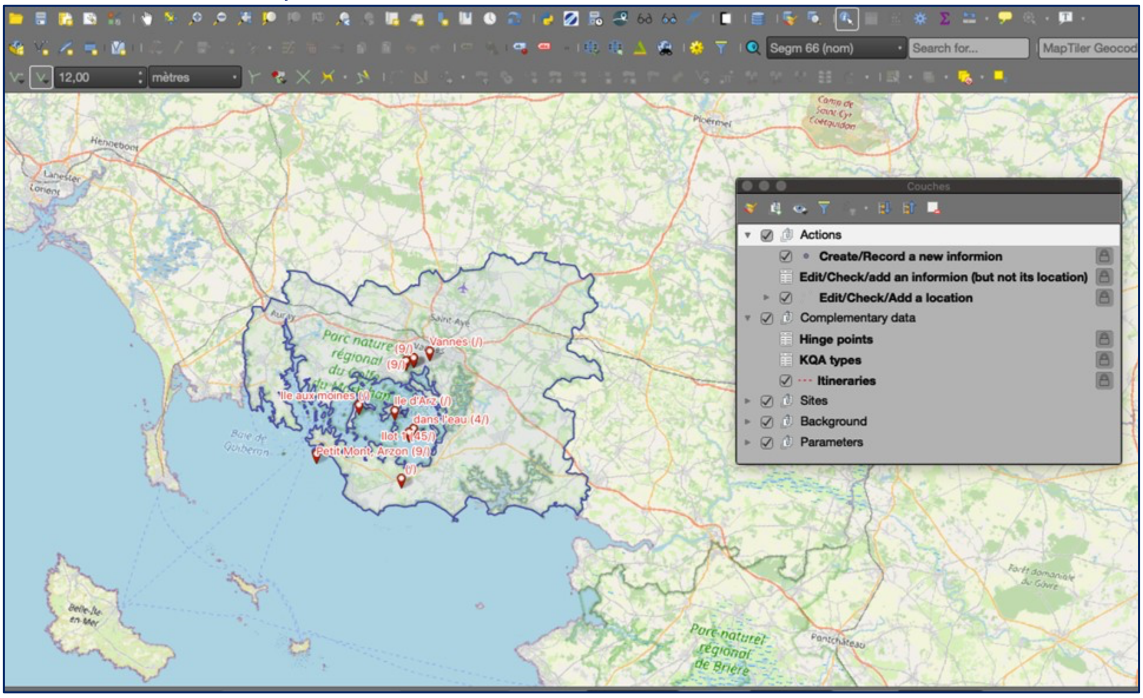
2.8. The locations in the *informions*

So we are here talking about the possibility to document field studies taking into account the spatial dimension. The informions therefore (usually) have a location and it took us time to decide what would be the 'place' of this geographical dimension, i.e. what location to record, and above all what would be its meaning for the informion. Is it the place were the surveyor would have met the respondent ? The place where lives the respondent ? A place the respondent is talking about ? And in this case, what if the respondent talks about a country or the planet ?

We have answered these unsolvable questions by creating a '&-to-n' relation between the informions and the locations, and we allow thus to associate one single informion to several places. We have also added the possibility to specify the spatial dimension of the place; is it a point or an area ? And if it's an area it is possible to specify it's size. Finally we have added the possibility to link the places together in a sequence, in order to represent a journey, like for instance the migration of a refugee.

Places also have their own characteristics, such as a description of the environment, the place (birth place, residence, work place, ...) in the interviewee's life, a name, sensitivity to climate change, etc

As said above the metadata model developed in WP 4.2 lists an important number of attributes, but these are now shared between informions and locations.



2.9. Using the tool

On opening the project (the QGIS project) the user sees the map showing his/her work area and the already recorded information and places. There is also a background map which can be chosen and modified by the user, the places have labels. There is also a panel listing the layers composing the map set and the actions that the user can do.

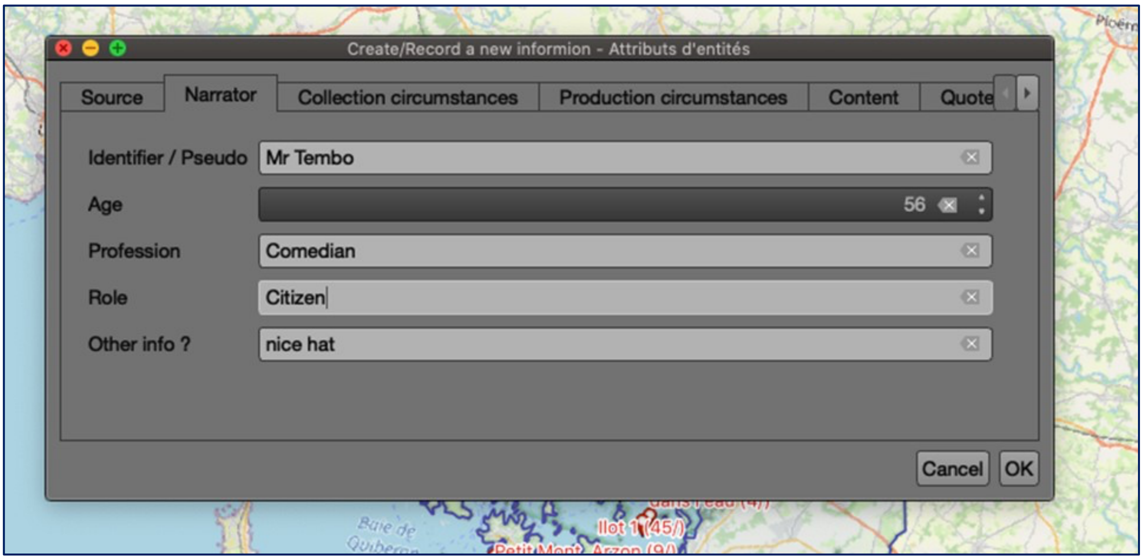
The main actions are :

- Create/Record a new informion
- Edit/Check/add an informion (but not its location)
- Edit/Check/Add a location

So the user can record, modify or delete the collected information in different ways, and also of course check the stored information by clicking on a symbol on

Figure 3. The informion form has many tabs grouping the different types of data to record.

the map. Doing so a form will open and display the data. The same form can be used to edit or create new items.



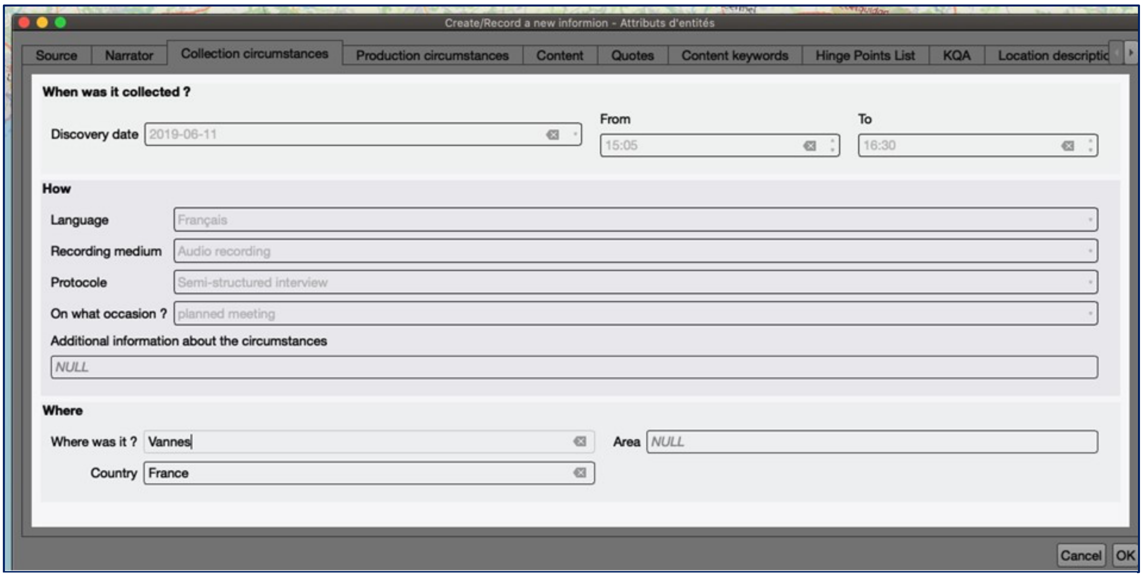


Figure 4. Describing the collection circumstances.

Places can have different roles in an informion (Figure 5), and additionally to their specificities they can be part of a sequence like for instance they can be stopovers in a journey, like for refugees but not necessarily, the sequence can be anything. This can be recorded by giving a sequential number to them. The system will sort them and draw a line connecting all the stopovers for a same informion (Figure 6 and Figure 7).

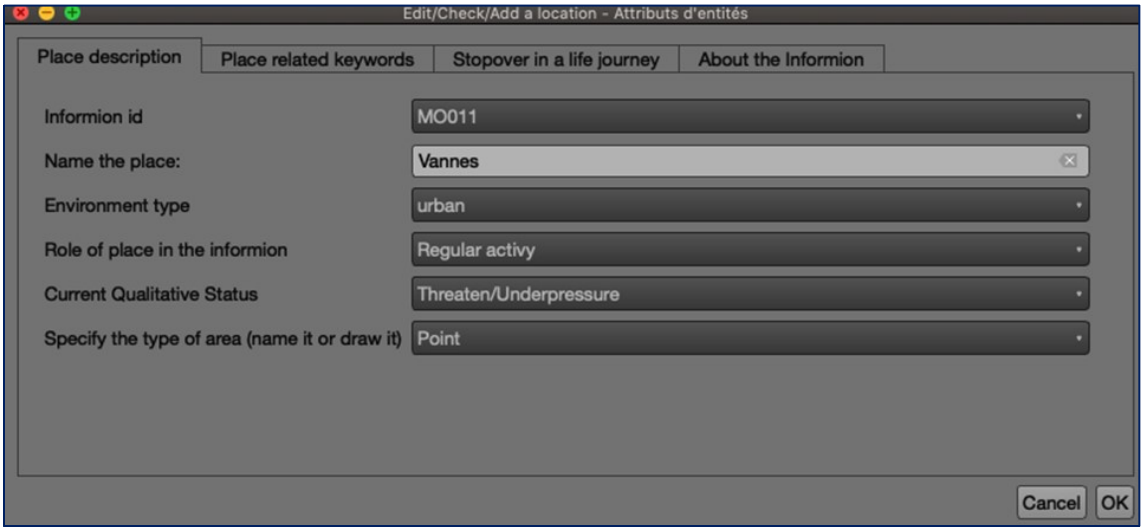


Figure 5. Recording a new location for an existing informion.

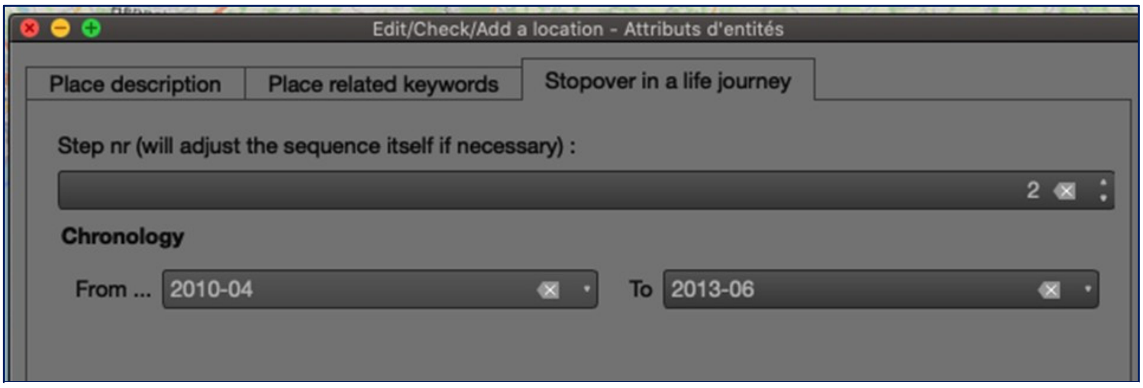
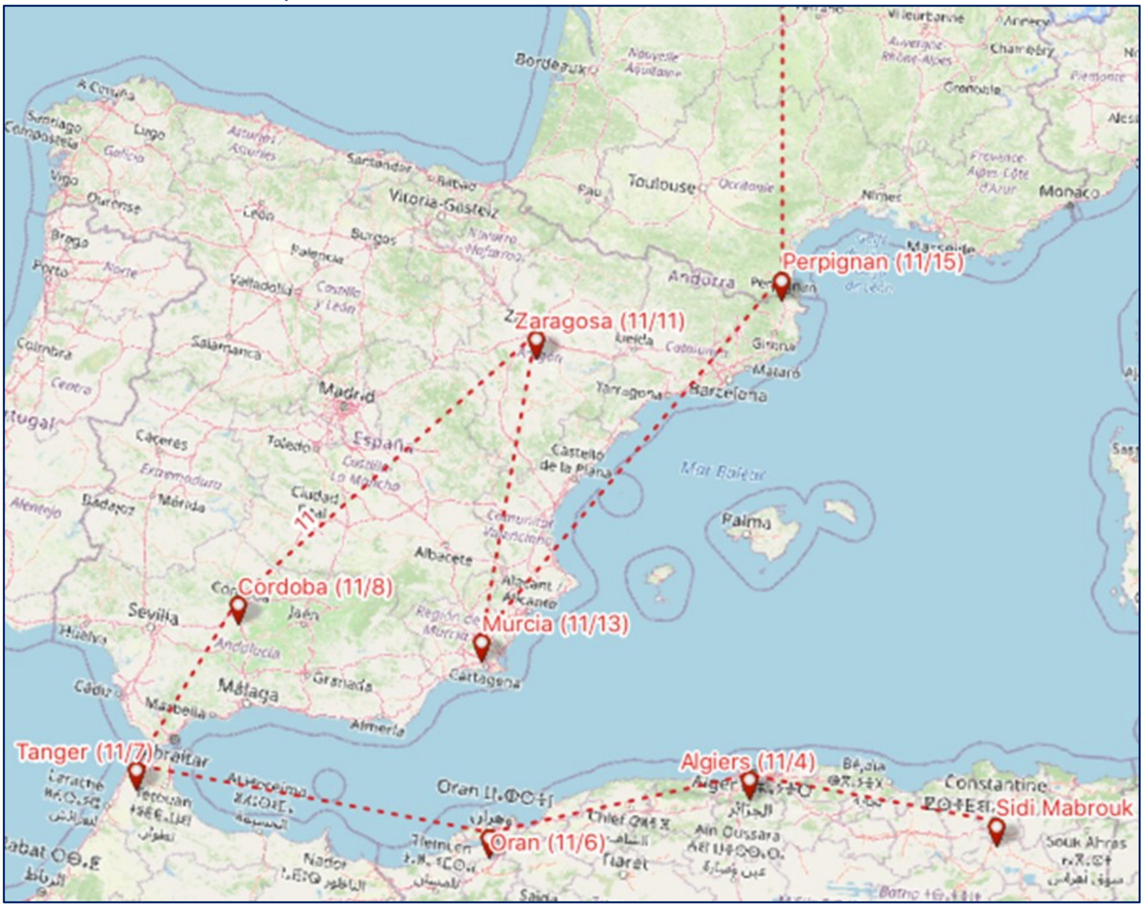


Figure 6. Specifying the journey's stopover for a place in an informion.

Figure 7. A journey made by connecting several stopover places.



2.10. Conclusion

Over the course of the project, we completely recreated the tool several times before reaching this final state, and we feel that we now have something that reflects the data collection work of the anthropologists in the project quite well, and thus could be used in the future for similar work. Above all, it is a very flexible and versatile architecture, which allows for a certain diversity of uses.

When the data are recorded and when they are found in a large enough quantity, it could be envisaged to carry out statistical analyses taking into account the spatial dimension.

References

1. Sébastien Caquard, Cartography I: Mapping narrative cartography, Progress in Human Geography, URL: https://journals.sagepub.com/doi/10.1177/0309132511423796, 2011
2. Sébastien Caquard URL : http://mappemonde.mgm.fr/121_as1/#englishversion
3. D J Peuquet, D F Marble, Introductory Readings In Geographic Information Systems, CRC Press, 1990.
4. Elsa Vivant, Burcu Ozdirlik et Nadia Arab, « L'artiste, la carte et le territoire : détourner et retourner les représentations », Belgeo [En ligne], 3 | 2014,. URL : http://journals.openedition.org/belgeo/13286 ; DOI : https://doi.org/10.4000/belgeo.13286

Annex – CoCliServ metadata for narratives

The table below displays the CoCliServ narrative metadata scheme, as agreed at the end of the project. The last two columns in the table provide two examples, which are entirely fictive. They were made up to provide the project partners with a flavour of what metadata fields are and how they can be used. The fields are divided in two categories: contextual information documents the conditions in which the considered narrative was collected (top lines in the table). Analytical information refers to elements of knowledge conveyed by the narrative itself (bottom lines in the table).

All free fields may be left empty (i.e. with no value).

Fictive examples

	Definition / explanation	Field name	Sub-field name	*Additional info & technical aspects*	Example 1	Example 2
Contextual information	*CoCliServ site to which the object (narrative) relates.*	CoCliServ site		*To be picked in list of five: Bryggen (Bergen), Isle of Dordrecht, Wadden Sea, Gulf of Morbihan, Saint-Pierre/Kerourien (Brest).*	Isle of Dordrecht	Gulf of Morbihan
	Identification of the entity or object described by the metadata.	Object	CoCliServ code name	*Automatically generated upon metadata submission.*	DO001	MO002
			nature	*To be picked in a list.* *Only "narrative" is possible for the moment. Other categories (e.g. "scenario", "milestone") may be added at a later stage.*	narrative	narrative

Definition / explanation	Field name	Sub-field name		*Additional info & technical aspects*	Example 1	Example 2
CoCliServ dataset to which the described object directly belongs.	Parent dataset			*Automatically generated upon metadata submission.*	CoCliServ Dordrecht Narratives	CoCliServ Morbihan Narratives
Info on who collected the narrative.	Collector(s)	Collector 1	name	*It could be several people…*	Arjan Wardekker	Charlotte Da Cunha
			institution	*… and several institutions.*	Copernicus Institute of Sustainable Development	Université de Versailles Saint-Quentin-en-Yvelines
Info on the circumstances when the narrative was collected.	Collection circumstances	location name			City Library Dordrecht	Gulf of Morbihan
		extreme Western longitude		*In decimal degrees, modulo 360.* *May be provided as a pair of coordinates (longitude, latitude), i.e. as a dot on a map, or as the coordinates of the corners of a rectangular bounding box (min & max longitude, min & max latitude).* ***The coordinates of the dot or bounding box may be*** *directly provided or* ***derived automatically from drawing on an interactive map.***	4,69	-2,98
		extreme Eastern longitude			4,69	-2,70
		extreme Southern latitude			51,81	47,53
		extreme Northern latitude			51,81	47,65
		area name				Britany
		country		*ISO 3166/MA English short name or Alpha-2 code: https://en.wikipedia.org/wiki/ISO_3166-1*	Netherlands	France

Definition / explanation	Field name	Sub-field name	*Additional info & technical aspects*	Example 1	Example 2
		date	*Gregorian calendar. dd.mm.yyyy*	17.06.2018	18.06.2018
		start time	*hh:mm Local time. Or do we prefer UT?*	14:00	14:00
		end time		16:00	17:00
		occasion		Bibliographic research.	Boat trip.
Reference(s) of any standard procedure or ad hoc professional protocol followed in collecting the narrative, whether internal or external to the CoCliServ project.	Applicable document(s)	Reference 1	*Several documents may be applicable.*	CoCliServ WP1 – M1.1. Dordrecht – Protocol for initial mapping of narratives, January 2018.	CoCliServ WP1 – M1.1. DRAFT protocol for the Initial mapping and In-depth analysis of narratives for the Gulf of Morbihan study site, January 2018.
Information on who enunciated (uttered, wrote) the narrative.	Narrator / group of narrators	name or nickname		Cornelis de Witt	Chirp
		age	*[Years].*	49	37
		professional activity		attorney, soldier	accountant
		solicited (interviewed or invoked) as		Mayor of Dordrecht	bird watcher

Definition / explanation	**Field name**	**Sub-field name**	***Additional info & technical aspects***	**Example 1**	**Example 2**
		additional relevant information		Accompanied Lieutenant-Admiral Michiel de Ruyter in his raids against the British fleet on River Medway (June 1667) and at Sole Bay (June 1672).	
Info on the circumstances when the narrative was enunciated, i.e. uttered or written. These may be identical to the circumstances when the narrative was collected, e.g. for interviews. They may be different from the former in situations where the narrative existed before it was collected, e.g. for press articles.	Production circumstances	location name		Den Haag	Gulf of Morbihan
		extreme Western longitude	*See above.*	4,19	-2,98
		extreme Eastern longitude		4,42	-2,70
		extreme Southern latitude		52,01	47,53
		extreme Northern latitude		52,13	47,65
		area name			Britany
		country	*Where relevant, ISO 3166/MA English short name or Alpha-2 code: https://en.wikipedia.org/wiki/ISO_3166-1*	Republiek der Zeven Verenigde Nederlanden	France
		date	*Gregorian calendar. dd.mm.yyyy - or Month yyyy - or yyyy*	July 1672	18.06.2018
		start time	*See above.*		14:00
		end time			17:00

Definition / explanation	Field name	Sub-field name	Additional info & technical aspects	Example 1	Example 2
		occasion		Author's imprisonment by the Orangist party after the battle of Solebay.	Boat trip.
Features that characterise the situation of the narrator, or of the collectivity, or place, or period of time where he/she is immersed (e.g. familial conditions, way of life, socio-economic characteristics of her/his city, historical background).	Thematic context		*This field does not target the theme(s) tackled by the narrative, which are supposed to be captured by the summary and keywords.*	Holland's commercial Golden Age. Naval war between the Dutch and allied English and French fleets. Orange party conspiracy.	Farming (breeders, fishermen, oyster farmers, etc.).
Language in which the narrative is expressed.	Language		*Using the ISO norm 639-1 - see https://fr.wikipedia.org/wiki/Liste_des_codes_ISO_639-1*	nl	fr
Medium on which the narrative is	Medium	format		manuscript	sound recording

	Definition / explanation	Field name	Sub-field name	*Additional info & technical aspects*	Example 1	Example 2
Analytical information	*recorded (sound recording, written notes, print, ...).*		bibliographic reference		de Witt, C., Overwegingen over het belang van zeestrijdkrachten voor de welvaart van onze goede Staten van Holland, monograph, Dordrecht, 28 July 1672.	
	Summary of the narrative or principal themes tackled by the narrator.	Synthetic abstract or principal themes			A powerful fleet warrants commercial prosperity. Peace with France benefits the Dutch people. Johan de Witt has been a clever Raadpensionaris.	Modifications in the usual life mode and journey of migratory birds may be related to local or remote possible climate changes.
	Keyword : word or expression conveying a concept or denoting a fact identified by the collector as a	Keyword(s)	Keyword 1	*As many keywords as desired may be provided. A keyword may be picked in the list of previously proposed keywords or added to the list. Keywords to be provided in English unless they refer to some concept specifically meaningful in the native language.*	Dutch history	bird
			Keyword 2		*Gouden Eeuw*	aquatic birds
			Keyword 3		Anglo-Dutch war	migratory birds
			Keyword 4		French-Dutch war	ecology

Definition / explanation	Field name	Sub-field name		Additional info & technical aspects	Example 1	Example 2
particularly significant component of the narrative. It may but must not be a recurrent term used by the narrator.		Keyword 5			naval battle	climate change impact
		Keyword 6			trade	
Places to which the narrative relates (a building,a neighbourhood, a city, region, country, a mountain, a continent,…).	Place(s)	Place 1	place name		Sole Bay	Gulf of Morbihan
			extreme Western longitude	*See above.*	1,67	-2,98
			extreme Eastern longitude		1,67	-2,70
			extreme Southern latitude		52,31	47,53
			extreme Northern latitude		52,31	47,65
			area name		Suffolk	Britany
			country	*See above.*	GB	France
		Place 2	place name		Dordrecht	Marais de Séné

Definition / explanation	**Field name**	**Sub-field name**		*Additional info & technical aspects*	**Example 1**	**Example 2**
			extreme Western longitude	*See above.*	4,69	-2,78
			extreme Eastern longitude		4,69	-2,69
			extreme Southern latitude		51,81	47,59
			extreme Northern latitude		51,81	47,66
			area name			Gulf of Morbihan
			country	*See above.*	Netherlands	France
		Place 3	place name		Republiek der Zeven Verenigde Nederlanden	
			extreme Western longitude	*See above.*	3,34	
			extreme Eastern longitude		7,22	
			extreme Southern latitude		51,01	
			extreme Northern latitude		53,59	

Definition / explanation	Field name	Sub-field name		*Additional info & technical aspects*	Example 1	Example 2
			area name			
			country	*See above.*	Republiek der Zeven Verenigde Nederlanden	
Explicit or implicit (i.e. derived by collector from the narrator's words) mention of - or allusion to - the future and/or climate.	Mention of the future	desired future		*Answer: "Yes" or "No".* *If yes, explicit or implicit?* *NB Past, present and future refer to nowadays as the present (not necessarily the narrator's present).*	yes, explicit	yes, explicit
		feared future			yes, explicit	yes, explicit
		without hope or fear overtone			no	no
	Mention of climate	past climate			yes, implicit	yes, explicit
		present climate			no	yes, explicit
		possible future climate impacts			no	yes, explicit
Does the narrative suggest anything that might be used as (or result in) a hinge point in the future development of the case study area? *If so, is the key agent related to climate? Can it be controlled?*	Potential for a hinge point	unclear		*One of the 3 options "Unclear", "No" and "Yes" must be ticked.*		
		no				
		yes	climate-related	*If "Yes", one of the 3 options about climate must be picked.*		
			not climate-related			
			relationship to climate undecidable			
			can be controlled	*If "Yes", one of the 3 options about the possibility of control must be picked.*		
			cannot be controlled			
			control possibility unclear			

Definition / explanation	Field name	Sub-field name	*Additional info & technical aspects*	Example 1	Example 2
Explicit mention, in the narrative, of a meteorological parameter (e.g. temperature, wind direction or speed, relative humidity), a meteorological phenomenon (e.g. rainfall) or a meteorological event (e.g. flood). The mention can arise as a general statement (e.g. on the frequency or intensity of extreme events), a particular story (e.g. the winter of 1947), a hypothesis, a question, a wish of information, etc.	Mention of meteorological physical parameter	air temperature	*To be ticked if relevant.*		
		sea surface temperature			
		soil temperature			
		air humidity			
		cloudiness			
		precipitation (rainfall, snow, hail)		☐	
		wind (speed and/or direction)		☐	
	Mention of climate-related phenomenon, event or landscape feature	storm(s)		☐	
		drought(s)			
		natural fire(s)			
		flood(s)			
		freshwater, river system			
		sea level			
		coast line			
		salinity			
		other please state			
Same as above but for aspects of plant life, animal life, human life which are explicitly connected, in the narrative, to	Mention of life aspect in relation to climate	air quality	*To be ticked if relevant.*		
		water supply			
		agriculture			
		food supply			
		energy supply			

Definition / explanation	Field name	Sub-field name		*Additional info & technical aspects*	Example 1	Example 2
climate or climate change.		ecology				
		vegetation, forestry				
		animal life				
		health				
		economy				
		education				
		culture				
		human population density				
		human migration				
		peace				
		politics				
		other	please state			
When existing, mentioned types of knowledge used in analysing and/or predicting the future of the site, with the reasons why the knowledge is trusted and whether it is recognised to be uncertain.	Knowledge types on which analysis and/or prediction of the site future is based	type 1	knowledge type	*Free text*		
			reasons for trust	*Free text*		
			identified uncertainties	*Free text*		
		type 2	knowledge type	*Free text*		
			reasons for trust	*Free text*		
			identified uncertainties	*Free text*		
		type 3	knowledge type	*Free text*		

Definition / explanation	**Field name**	**Sub-field name**		***Additional info & technical aspects***	**Example 1**	**Example 2**
			reasons for trust	*Free text*		
			identified uncertainties	*Free text*		
Does the narrator refer to any knowledge quality issues or criteria?	Knowledge quality aspects	mentioned knowledge quality criteria		*Free text*		
		mentioned knowledge issues		*Free text*		
Excerpts from the narrative, judged interesting by the collector according to his/her own criteria or analysis perspective (illustrative quality, pertinence, meaningfulness, poetical quality, sense of humour, …).	Quotes	Quote 1		*As many quotes as desired.* *Should we keep the quotes in their original language?*	We won at Sole Bay.	Birds, marvellous birds.
		Quote 2			My brother is not a treator.	A redshank ringed in Séné has been located in Iceland two years later.
		Quote 3			Long life to the States of Holland.	We fear that level rise will make marshes disappear and sweep swamp life away.
	ADDITIONAL FIELDS ?					