Analysis Process Guidelines

version 1.2.0 · 17 December 2024

This document and its contents were created by Cesar Gonzalez-Perez, Martín Pereira-Fariña and Beatriz Calderón-Cerrato.

10 February 2025 18:13 · revision 1715

Table of Contents

Introduction

IAT/ML is a method for the combined analysis of discourse from ontological, argumentation and agency points of view. IAT/ML is based on Inference Anchoring Theory (IAT) and the ConML conceptual modelling language.

This document provides a process-oriented set of guidelines for discourse analysis using IAT/ML. Please see the www.iatml.org web site for more details, and the *IAT/ML White Paper* document for an introduction to IAT/ML. Also, see the following documents for specific guidance on each type of analysis:

- *IAT/ML Ontology Patterns Guidelines*
- *IAT/ML Argumentation Patterns Guidelines*
- *IAT/ML Agency Patterns Guidelines*

Overview

This section presents an overview of the process to be followed when using IAT/ML.

Phases and Processes

IAT/ML is organised around the phases and processes shown in Figure 1.
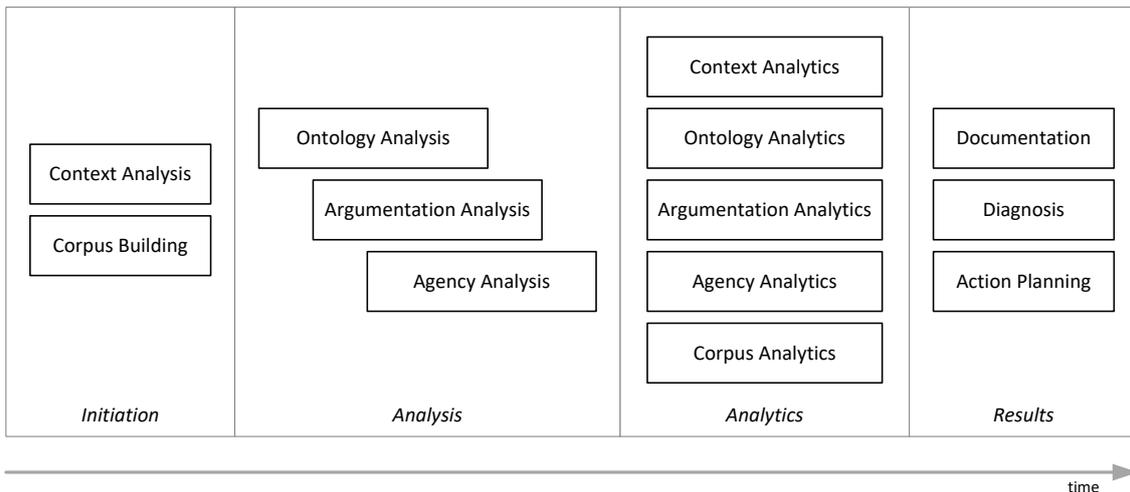
Figure 1. Phases (bottom, in italics) and processes (as boxes) of IAT/ML over time.

Each phase and process are described in the next sections.

Initiation

During this phase, you define the situation to be addressed by your project, and make the necessary preparations for analysis.

There are two processes that you must carry out during this phase:

- **Context Analysis**, to describe the issues that motivate your project and characterise the associated situation in terms of themes, positions and agents.
- **Corpus Building**, to gather the documents to be analysed.

These processes may occur in any order, and often overlap. Some projects start by defining a situation via context analysis, which establishes the needs for document gathering and corpus building. Some others start from an existing corpus and use the available documents to characterise the situation to address. Many projects do a bit of both, incrementally, by carrying out a preliminary context analysis, gathering some documents, studying them to refine the context analysis, gathering more documents, and so on and so forth. In any case, every project needs some context analysis (to define what the project is about) and some corpus building (to gather the documents to be analysed).

Analysis

During this phase, you manually analyse the texts gathered during initiation, and produce a set of discourse models. This phase is the core of IAT/ML.

There are three processes that you may carry out during this phase:

- **Ontology Analysis**, to create one or more ontology models (or, simply, ontologies) that describe what things in the world speakers refer to in the documents.
- **Argumentation Analysis**, to create an argumentation model for each document, which shows how speakers justify what they say.
- **Agency Analysis**, to create one or more agency models that describe the beliefs, desires and intentions of the speakers in each document.

You can carry out one or more of these processes. You can choose which kinds of analysis you want to do depending on the needs of your project. All these processes are carried out by hand and take significant time and effort for a human analyst to complete.

Also, analysis processes usually overlap in time, but they work better if you start by ontology analysis, do argumentation analysis next, and finish with agency analysis. This is so because each type of analysis becomes easier, faster and richer if you have carried out others before. For example, argumentation analysis can provide much more information if you connect it to an ontology model that you developed before. Similarly, agency analysis is easier to do if you start from existing ontology and argumentation models.

Analytics

During this phase, you apply automated analytical algorithms to the discourse models from the previous phase, and obtain quantitative and visual results.

There are five processes that you may carry out during this phase:

- **Context Analytics**, to process the context defined during initiation, such as context balance.
- **Ontology Analytics**, to process ontologies and obtain additional information about them, such as entity centrality.
- **Argumentation Analytics**, to process argumentation models and obtain additional information about them, such as denotation relationships, argumentation structure, or proposition centrality.
- **Agency Analytics**, to process agency models and obtain additional information about them.
- **Corpus Analytics**, to process the complete corpus or a subset of it and obtain additional information about it, such as intertextuality relationships, significance measurements, collocations, co-occurrences, or readability metrics.

You can carry out none, one or more of these processes, depending on the needs of your project. Some projects do not need any analytics at all, whereas others rely strongly on the quantitative and visual results that they provide. If you are using LogosLink, all these processes are fully automatic; you only need to provide some parameters and the computer executes the analytics for you.

Results

During this phase, you elaborate products that you can apply later on to address the target situation of your project.

There are two processes that you may carry out during this phase:

- **Documentation**, to produce a descriptive document that provides detailed information on your findings on themes, positions and agents.
- **Diagnosis**, to produce an interpretive document that provides detailed information and visualisations on the situation being addressed within its social context.

- **Action Planning**, to produce a recommendations document that outlines what actions may be taken to address the situation being addressed, and what consequences are likely for each action.

Other types of results processes are possible. Since different projects may have very different requirements in relation to what outcomes are expected, you can "plug in" your own results process at this point to obtain a customised result.

Tools

To apply IAT/ML, you can take a manual approach and use pen and paper or a whiteboard. However, software tools make some tasks much easier, faster and less prone to errors.

For text-based tasks, you can use a word processor or a plain-text editing tool. For analytical tasks, you can use different specialised tools.

One of these tools is LogosLink. LogosLink is a software tool specifically designed to support IAT/ML ontology and argumentation analysis, and has some corpus management features as well. You can download LogosLink for free from:

http://www.iatml.org/LogosLink/

In addition, Bundt can help with ontological analysis. Bundt is a collection of software tools for conceptual modelling using ConML, and although its scope is different to that of IAT/ML, it can complement LogosLink nicely for some ontology modelling tasks. You can learn more about Bundt, as well as download the software, from:

http://www.conml.org/Bundt/

Initiation

During this phase, you define the situation to be addressed by your project, and make the necessary preparations for analysis.

Usually, you must carry out context analysis and corpus building in parallel, as each of these processes is supported by the other.

Context Analysis

The purpose of this process is to describe the issues to address and characterise the associated situation in terms of themes, positions and agents. This usually involves the following tasks:

1. Characterise the situation to address
2. Describe the relevant issues
3. Determine the relevant themes
4. Determine the relevant positions for each theme
5. Determine the relevant agents for each theme

Although the context of your project is defined during initiation, it is perfectly normal to come back to it and make adjustments throughout the project. In particular, it is common that the analysis phase reveals new information that will make you want to revise and refine the positions and agents that you have initially defined.

1. Characterise the situation to address

This task aims to describe the situation that you want to address within your project. A "situation" may be a social problem, a phenomenon, or any other matter that you think deserves attention through discourse analysis.

The outcome of the task is a brief written text describing the situation and the objectives of your project in relation to it. For example,

> In this project I want to address the coexistence problems in multi-ethnic neighbourhoods in the city of Malmö, and especially those problems associated to racism, educational challenges and gentrification.
>
> The aim of the project is to produce a thorough diagnosis of the situation plus some recommendations for municipal public policies.

2. Describe the relevant issues

This task aims to outline one or more issues that are relevant to the described situation. Issues correspond to questions or knowledge gaps that you would like to fulfil.

Any project is motivated by a drive to know more about something. The gaps between what you already know and what you want to know are usually expressed through research questions and project goals. You can use these motivators to express one or more issues that seem relevant to your project, each one formulated as a question. Not every issue that you express at this point will end up being fully address during the project, but it's good that you write it down nevertheless for the sake of context.

The outcome of the task is a list of issues to explore. For example,

> Issue 1: Why are immigrant kinds doing worse at school?

Issue 2: What kind of discriminatory practices exist in the job market, and how pervasive are them?

3. Determine the relevant themes

This task aims to identify and describe one or more themes that are relevant to address the outlines issues. A theme is a subject or idea to be investigated. Themes may be decomposed in sub-themes if necessary.

You can usually find relevant themes in the description of the situation and issues that you composed earlier on. Also, if you have a corpus or some documents already, you can browse them to gain an understanding of the major issues that need to be addressed.

The outcome of the task is a list of the themes that you will address during the project. For example,

- Racism
    - o Racism at school
    - o Racism in the job market
- Gentrification

4. Determine the relevant positions

This task aims to identify and describe the most common positions that people hold about each of the themes just described. A position is a belief that is supported by some and attacked by others, and which is relatively well known in the relevant social group.

You should be as fair as possible and collect positions that cover the whole spectrum of beliefs, not only those that you like better or are more widely accepted. Including a position in your project does not mean that you agree with it.

If you have a corpus or some documents already, you can read through them to find out positions commonly held, defended or attacked.

The outcome of the task is a list of positions for each theme. For example,

- For the "racism at school" theme:
    - o Students from ethnic minorities struggle to complete their studies because of language barriers
    - o Students from ethnic minorities struggle to complete their studies because of unfair discrimination by teachers and other students
    - o Students from ethnic minorities do not struggle more than others to complete their studies
- For the "gentrification" theme:
    - o People from ethnic minorities are being displaced to cheap and substandard housing due to rising rent prices
    - o Rent prices are being artificially inflated to expel ethnic minorities from some neighbourhoods

5. Determine the relevant agents

This task aims to identify and describe the relevant agents for each of the themes just described. An agent is an individual or group that has a voice, or should have a voice, about each of the themes that you described. Agents may be decomposed in sub-agents if necessary.

As in the case of positions, you should be as fair as possible and include agents that cover the whole spectrum of social diversity, not only those that you like better or align with. Including an agent in your project does not mean that you identify with it.

If you have a corpus or some documents already, you can read through them to find out who the relevant stakeholders and social actors are.

The outcome of the task is a list of agents for each theme. For example,

- For the "racism at school" theme:
  - Students:
    - Students from ethnic minorities
    - Students from ethnic majorities
  - Teachers:
    - Teachers from ethnic minorities
    - Teachers from ethnic majorities
- For the "gentrification" theme:
  - People from ethnic minorities
  - People from ethnic majorities
  - Real estate owners
  - The local Town Council

Corpus Building

The purpose of this process is to obtain the documents to be analysed in plain-text format. This usually involves the following tasks:

1. Determine the kinds, sources and numbers of documents to gather
2. Gather the documents
3. Convert the documents into plain-text files
4. Prepare the texts for analysis

Your corpus is expected to be quite well defined by the end of this process. However, it is normal that you need to add extra documents to the corpus, or remove others, as a consequence of downstream work during the analysis phase. Feel free to adjust the contents of your corpus as you advance through analysis.

1. Determine the kinds, sources and numbers of documents

This task aims to decide on what kinds of documents you will be using, where you will find them, and how many of them you need to analyse.

When deciding on document kinds and sources, make sure that you will be able to address the issues described and cover the full range of agents and positions that you have identified. Also, bear in mind that discourse analysis requires a lot of time. You can use the following numbers as rough guidance about how many documents or words can be realistically analysed in your project:

- For ontological analysis, you will need about 6 hours per simple document or set
- For argumentation analysis, you will likely need 1 hour per every 300 words or so
- For agency analysis, you will need about 8 hours per simple document or set

A document set, in this context, is a collection of documents that you analyse as a whole. It is common to carry out ontology and agency analysis only once for a set of related documents instead of once for each individual document.

The outcome of this task is a brief collection of criteria that will help you find the necessary documents. For example,

> I will use news articles and opinion pieces from online sources, as well as transcribed audio interviews with residents from ethnic minorities and majorities, including both school-aged children as well as their families. I aim to gather documents up to around 50.000 words.

2. Gather documents

This task aims to obtain the documents that you will analyse during the project and store them in a safe place.

Documents may be of any kind; in this sense, IAT/ML is suitable to analyse almost anything as long as it contains a textual discourse. Of course, different types of analysis will be more or less applicable depending on the nature of the document. For example, it is difficult to do argumentation analysis on a very descriptive text, but ontology or agency analysis may work well. In this regard, you should select the types of analysis that you will perform depending on the type of documents that you have, as well as your research goals.

Documents can include web pages, PDF files, audio or video files, etc., as long as they contain a discourse expressed in a human language.

The outcome of this task is a collection of document files containing the discourses to be analysed. This collection can be a formal corpus if you use a software tool with corpus management features such as LogosLink.

3. Convert documents into text

This task aims to produce, for each document, a plain-text file containing the same textual information (or mostly) as the associated document.

If the source document is a PDF file, a Word document, or some other document in a textual format, this task involves removing all formatting and layout, deleting images and other non-textual elements, and removing any fragments of text that are not interesting for the analysis, such as page numbers, footnote callouts, figure captions, etc.

If, on the other hand, the source document is an audio or video file, you must transcribe it into text. There are many free and paid tools specialising in audio transcription, as well as specialist companies that can do it.

In any case, you should store the plain-text files together with the source documents, using a corpus management software tool if possible.

The outcome of this task is a plain-text file containing the discourse to be analysed for each source document.

4. Prepare texts for analysis

This task aims to clean up the texts previously obtained and add markings to aid during analysis.

Start by cleaning up the text in each file, removing unnecessary sections such as comments, captions, footnotes, running headers or other text fragments that are not relevant for analysis. In general, fragments containing text with no discursive value are usually unnecessary and may be removed. It is a good idea to write down some criteria on what kinds of text elements are to be removed from texts and apply them throughout the corpus.

Then, structure the text in paragraphs so that text belonging to different speakers is clearly separated. Headings or section titles, if they exist, must be clearly visible, each in its own paragraph.

Then, if text involves two or more speakers, introduce speaker marks for each paragraph, following a well-defined format. You can do this by inserting the speaker's name or nickname at

the beginning of each paragraph, followed by a semicolon or a similar sign. Again, this is something you should homogenise if you are working with a corpus.

Then, insert timestamp marks at the beginning of each paragraph if they are available. This is often the case for text that has been transcribed from an audio or video source, but is usually absent from textual sources.

Finally, identify and make note of the language of the text.

The outcome of this task is a properly marked up plain-text file, ready for analysis, for each source document. The following is an example taken from [1]:

> [00:01:21] Brennan: The New Yorker reported that when her name was first floated to you by- by Senator Daniel Patrick Moynihan, you had some misgivings. Why?
>
> [00:01:28] Clinton: I don't remember that. I, first of all, there were- everybody that I know is taking credit now for- for 27 years for nominating her, but I didn't have misgivings.

In the example, timestamp marks are included in brackets at the beginning of each paragraph, followed by the speaker's name and a semicolon before the actual discourse content. You don't need to follow this particular syntax but, in any case, a clear indication of speaker names is important or even necessary for argumentation and agency analysis. Timestamps, in turn, are important for some kind of analytics, but not as much as speaker names.

Analysis

During this phase, you manually analyse the texts gathered during initiation, and produce a set of discourse models. This phase constitutes the core of IAT/ML.

You can carry out ontology analysis, argumentation analysis or/and agency analysis, depending on what discourse models you are interested in. Bear in mind that some analysis processes are easier and producer richer results if you tackle them in a particular sequence: first, ontology, then, argumentation and, finally, agency analysis.

Ontology Analysis

The purpose of this process is to analyse a body of text to create an ontology model, often called simply an ontology. An ontology is a model that represents what things the text is about, what properties they have, and how they are connected to each other.

The text that you use as a basis to develop an ontology can pertain to a single document, to a set of documents that share some commonalities (for example, a common theme), or even a whole corpus.

Developing an ontology usually involves the following tasks:

1. Decide the level of analysis
2. Find key terms and phrases
3. Define entities, features and facets
4. Revise the model

Please see the *IAT/ML Ontology Patterns Guidelines* document for additional details.

1. Decide the level of analysis

This task aims to decide whether you will create a separate ontology for each document, a common ontology for all the documents, or one ontology for each document subset.

Each option has pros and cons. In general, creating an ontology per document will result in many ontologies, each of them quite small. To the contrary, if you create a single ontology for the whole corpus, you will produce a much larger ontology, but only one. There are intermediate options, such as creating an ontology for each subset of documents according to common themes.

It is not a good idea to create a separate ontology per position or agent, because ontological analysis aims to, precisely, reveal the range of things referred to by different agents and from different positions, as well as the connections between them. If you separate agents or positions into different ontologies, you lose this.

If you have a small or medium corpus where all the documents are about a common theme, you should probably create a single ontology. If your corpus uses topics or is clearly divided into different themes, then you probably should create an ontology per topic or theme. Unless your project wants to focus strongly on individual and singular documents, avoid creating a separate ontology per document.

The outcome of this task is a decision on what the level of analysis will be.

2. Find key terms and phrases

This task aims to use the selected text and build candidate lists of terms and phrases that can be used to define ontology elements later on.

You don't need to include too many terms in the lists. In fact, an ontology must be minimalistic. Include only terms and phrases that are central to what the text is about. Consider different kinds of things that may be being referred to, such as physical things (such as buildings or objects), places, events, people (including groups and individuals), situations and circumstances, abstract ideas (such as opinions or values), expressive elements (such as songs, performances or artworks), measurements, or time spans (such as moments or periods).

The outcome of this task is a set of lists of candidate categories, atoms, features and facets.

Highlight text

Gather the text to be analysed, either from a single or multiple documents, and start by highlighting the nouns, adjectives, verbs and adverbs, using different colours or highlighting styles. For nouns, you should highlight two kinds: atom and category nouns. Atom nouns refer to individual things, whereas category nouns refer to types of things.

Atom nouns include proper names, such as "Alice Abbott", "the Eiffel Tower" or "Microsoft", and also common names with determiners that indicate individual things, such as "a song" or "the first person on the Moon". Category nouns, on the contrary, often appear as common nouns in plural or accompanied by quantifiers such as "each" or "every". Also, category nouns may be simple words like "people" or "songs", or longer multi-word expressions such as "customer representatives" or "long-term projects". Adjectives and verbs may also be multi-word phrases, such as "is related to" (a verb) or "who expressed the strongest opinion" (an adjective).

When highlighting terms and phrases, consider that the focus of this task is to find candidate ontology elements, that is, things in the world that are relevant to the text being analysed. You should not highlight every single noun, adjective or verb, but only those that you find relevant to your research goals. Consider the following example:

> A database was used to manage and keep track of a set of archaeological sites in the UK. Features and artefacts found in each site were given a code and a description. Documented artefacts were stored in boxes and labelled accordingly. Sites that had been excavated were marked as such. Boxes from the High Street site were sent to the local museum.

Here, atom nouns have been highlighted in pink, category nouns in blue, adjectives in yellow, and verbs in green.

Make lists

Then, create separate lists for atom nouns, category nouns, adjectives, and verbs, sorting them alphabetically and removing duplicates. Often, you will need to decide whether two similar terms are in fact the same or not. For example, "archaeological sites" and "sites" in the example above probably refer to the same thing, so the term is included in the list only once.

The list of atom nouns for the example text would look like this:

- A database
- The High Street site
- The local museum
- UK

Proper names are left as they appear, or reworded slightly if necessary.

The list of category nouns would look like this:

- Archaeological Site
- Artefact
- Box
- Code
- Description
- Feature

Category nouns are kept in singular and with initial capitals.

The list of adjectives would look like this:

- Documented (Artefact)
- Excavated (Archaeological Site)

Here, the noun to which each adjective applies in indicated in parenthesis.

Finally, the verb list would look like this:

- (Boxes for the High Street site) Are Sent To (the local museum)
- (Artefact) Is Found In (Archaeological Site)
- (Feature) Is Found In (Archaeological Site)
- (Many archaeological sites) Are In (the UK)
- (Archaeological Site) Is Marked as Excavated
- (Artefact) Is Stored In (Box)
- (A database) Keeps Track Of (some archaeological sites)
- (A database) Manages (some archaeological sites)

In this list, verbs are expressed in present tense, and the subjects and objects to which they refer, which must be either atom or category nouns, are included in parentheses.

Check and simplify lists

Now, look for and remove redundant items across lists. For example, the "Excavated (Archaeological Site)" adjective and the "(Archaeological Site) Is Marked as Excavated" verb actually refer to the same thing, so that we can remove the verb. Similarly, the "(A database) Keeps Track Of (some archaeological sites)" and "(A database) Manages (some archaeological sites)" verbs seem to refer to the same thing being expressed in slightly different ways, so we can remove the latter and keep the former. Finally, the "Documented (Artefact)" adjective seems to be unnecessary, as non-documented artefacts are, by definition, not in the database, so that all artefacts in the boxes would be documented. We can remove it.

Also, you must make sure that lists are well integrated. Check that all the nouns in parentheses in the adjective and verb lists exist in the noun lists, and add those that do not.

Add types for instances

In an ontology, it is recommended that every atom has a corresponding category, so start by adding items to the category noun list for items in the atom noun list that have no matching category. In our example above, the following category items are added:

- Database
- Museum
- Country

Similarly, it is necessary that every adjective or verb that refer to an atom also has a category-level equivalent. In our example, the "(Boxes for the High Street site) Are Sent To (the local museum)" verb connects two atoms. Similarly, "(A database) Manages (some archaeological

sites)" or "(Many archaeological sites) Are In (the UK)" connect atoms together. In this manner, the equivalent category-level verbs must be added:

- (Box) Is Deposited In (Museum)
- (Archaeological Site) Is In (Country)
- (Database) Manages (Archaeological Site)

We have changed the "send to" verb to "deposit", because "send to" indicates an event whereas "deposit" indicates a change of state. Verbs that indicate more or less stable states are much easier to represent in ontologies than verbs indicating events or rapid changes, and are often more valuable, so rewordings like this are usually recommended. For example, "has met" is better than "meets" when analysing acquaintance relationships between people.

Introduce supertypes

Then, consider whether groups of two or more categories are in fact subtypes of a common supertype, and introduce this. In our example, "Artefact" and "Feature" are clearly subtypes or a larger supertype, as both artefacts and features are described in the text as being treated in a similar way ("were given a code and a description"). The text doesn't provide a term for this more abstract concept, so we must come up with one, such as "Archaeological Element".

Once a supertype is introduced, lists must be revised to make sure that it is replaced for its subtypes were appropriate. For example, the verb list in our example would be simplified as follows:

- (Box) Is Deposited In (Museum)
- (Archaeological Element) Is Found In (Archaeological Site)
- (Archaeological Site) Is In (Country)
- (Artefact) Is Stored In (Box)
- (Database) Manages (Archaeological Site)

Identify features

Some nouns in the categories list may not indicate proper categories but properties or associations of categories. This is usually easy to detect as these items represent things that don't exist by themselves but as properties or associations of others. In our example, "Code" and "Description" are properties of features and artefacts, so we can decorate them with their "owner" in parentheses. Also, and since we have just introduced the category "Archaeological Element" to cater for both features and artefacts, we would add them as follows:

- Archaeological Element
- Code (of Archaeological Element)
- Description (of Archaeological Element)

3. Define entities, features and facets

This task aims to produce a first version of the ontology for the text or texts being analysed.

The outcome of the task is an ontology (or ontological model) based on the term lists previously obtained, and expressed in terms of entities, features and facets.

Create categories

Categories in the ontology can now be created from items in the category noun list. In our example, we would create the following categories:

- Archaeological Element
- Archaeological Site

- Artefact (subtype of Archaeological Element)
- Box
- Country
- Database
- Feature (subtype of Archaeological Element)
- Museum

When creating categories, think if any of them represent any of the themes, positions or agents in the associated context. If this is the case, record the connection. For example, the Museum category in the above list may correspond to a Museum agent defined during context analysis.

Create features

Similarly, now we can add features (properties and associations) to the just created categories by looking at the category noun list. In our example, we would add "Code" and "Description" properties to the "Archaeological Element" category.

Also, the adjective list usually contains indications of additional features. In our example, we would add an "Is Excavated" property to the "Archaeological Site" category.

Finally, the verb list contains useful information to add associations between categories. In our example, these associations would be created:

- (Box) Is Deposited In (Museum)
- (Archaeological Element) Is Found In (Archaeological Site)
- (Archaeological Site) Is In (Country)
- (Artefact) Is Stored In (Box)
- (Database) Manages (Archaeological Site)

Note that associations always connect two categories, but association names refer to one direction only. For example, "(Artefact) Is Stored In (Box)" connects the "Artefact" and "Box" categories via the "Is Stored In" association, which is named from the perspective of "Artefact" and towards "Box". The inverse name would be, for example, "Contains". It is a good idea to add associations in both directions for the sake of completeness.

Create atoms

Once categories and features have been added to the ontology, we can create atoms by looking at the atom noun list. Every atom must be added to the ontology as an instance of its corresponding category. In our example, these atoms would be created:

- A database (instance of Database)
- The High Street site (instance of Archaeological Site)
- The local museum (instance of Museum)
- UK (instance of Country)

As in the case of categories, think if any of the created atoms represent any of the themes, positions or agents in the associated context. If this is the case, record the connection. For example, the local museum atom in the above list may correspond to a Museum agent defined during context analysis.

Create facets

We can now add facets (values and references) to the just created atoms. Like for the atoms themselves, every facet in an ontology must be an instance of a corresponding feature; values are instances of properties, whereas references are instances of associations.

In our example, there is only one property, "Is Excavated" of "Archaeological Site", and since we don't know whether the only atom of this category, "The High Street site", is excavated or not, we would not create any values. Also, there are no clear information in the text about references between atoms, so we wouldn't create any.

4. Revise the model

This task aims to revise the ontology as a whole, assess its integrity, coverage and balance, and carry out the necessary amendments to improve it.

The outcome of the task is an improved ontology. You can revise the ontology again if you feel the need, until it eventually stabilises.

You may also want to revise the context that you defined during project initiation, and update its positions and agents according to the information revealed during ontology analysis.

Check integrity

Integrity refers to the structure of the ontology being free from issues. For example, it is common that some associations are added to an ontology during the initial steps without thinking about the inverse association names. This can be checked and completed as necessary during revision.

Similarly, you should make sure that category and feature names are concise, self-explanatory and unambiguous. They should also follow the same naming convention and style.

Check coverage

Coverage refers to the ontology being complete, covering everything that needs to be addressed. Revise the ontology for clearly missing elements and add them. In our example we have a "Country" category, but there is no connection between this country and other elements in the ontology. We should add, for example, an association "(Archaeological Site) Is Located In (Country)" to make the ontology a single connected mesh.

Also, revise categories for clearly missing properties. For example, we may want to add "Name" properties to the "Country" and "Archaeological Site" categories in our example, since it makes sense to think that every country and archaeological site will have a name that we can use to refer to them.

Check balance

Balance refers to the ontology being equally detailed across the concepts it represents. For example, sometimes texts describe some concepts with much detail, and some others just barely. We can decide now to re-balance this and add (or even remove) details as necessary.

Figure 2 shows a sample diagram of a finished ontology.
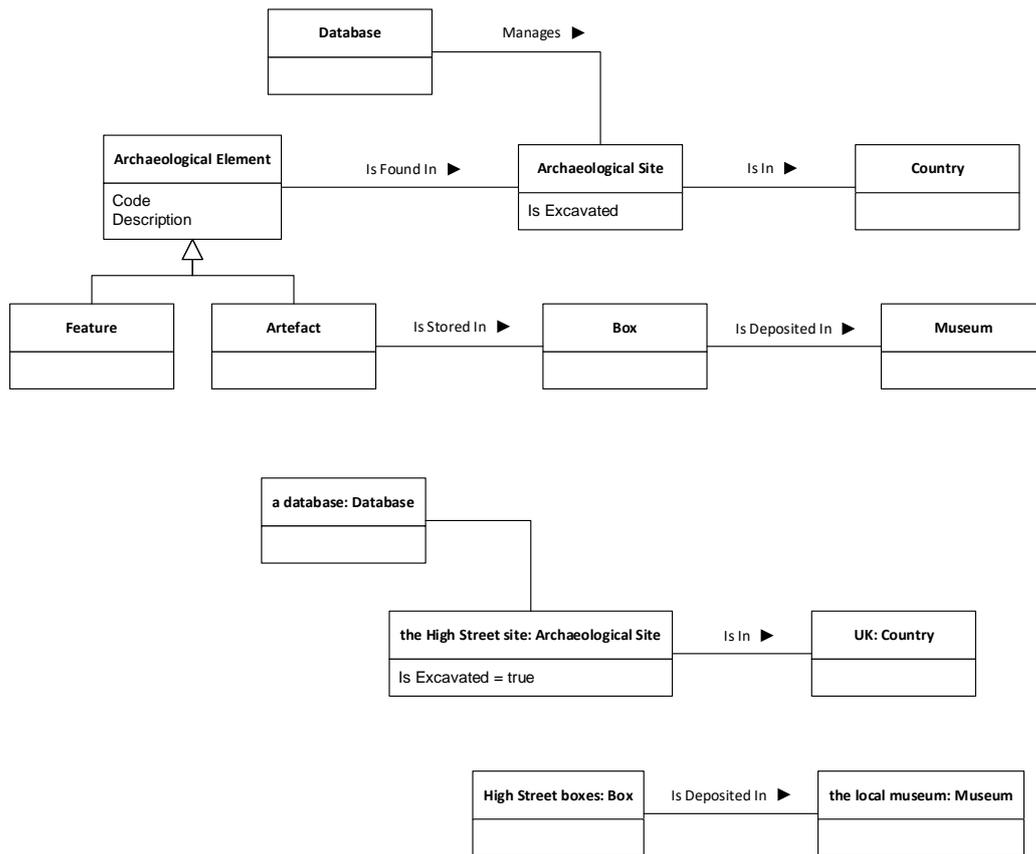
Figure 2. Sample ontology, depicted as a diagram. Boxes represent atoms or categories. Text inside boxes represents values or properties. Lines connecting boxes represent references or associations.

Argumentation Analysis

The purpose of this process is to analyse a body of text to create an argumentation model. An argumentation model is a representation of the statements made by the speakers in the text plus the relationships that they have with each other regarding how each one is justified.

The text that you use as a basis to develop an argumentation model usually pertains to a single document. It rarely makes sense to create an argumentation model for a set of documents.

Developing an argumentation model usually involves the following tasks:

1. Segment the text
2. Reconstruct the argumentation
3. Add argumentation relations
4. Add denotations
5. Revise the model

Please see the *IAT/ML Argumentation Patterns Guidelines* document for additional details.

1. Segment the text

This task aims to divide the text to be analysed into meaningful chunks called locutions, connected by transitions.

The outcome of this task is an initial argumentation model, containing a collection of locutions and transitions that represent the discourse to be analysed.

Create locutions

A locution is a block of text that is spoken by a single speaker, and which expresses a single statement. Often, locutions correspond to sentences, but sometimes a single sentence can include multiple locutions. For example, take the following sentence:

> The speed at which the virus spread surprised us and, although we do not yet have all the answers to the question, in recent weeks very important discoveries have been made that help us understand what is happening.

This sentence should be segmented as four separate locutions:

> LO1. The speed at which the virus spread surprised us
>
> LO2. we do not yet have all the answers to the question
>
> LO3. in recent weeks very important discoveries have been made
>
> LO4. that help us understand what is happening.

Note how each locution contains a single statement about the world. Note also that some connector words, such as "and" or "although", have been removed, as they will be encoded as transitions.

A good rule of thumb is to start by dividing the text according to sentence boundaries and then revise each sentence to further divide it as needed.

Create transitions between locutions

Once the text has been divided into locutions, these must be connected by using transitions. A transition indicates a discursive relationship between locutions. A transition always starts at one locution and ends in one or more locutions. The locutions linked by a transition may belong to the same speaker or different ones.

If the discourse contains timestamps, then the initial locution in a transition must be prior in time in relation to the final locutions of the transition. In other words, transitions must be compatible with the time ordering of locutions over time, even if the text does not contain timestamps. All locutions in the model must be connected into a single mesh, that is, no locution "islands" should remain.

There are different kinds of transitions, depending on what kind of discursive relationship exists between the related locutions:

- **Adding**: the speaker continues talking by adding a new locution right after the previous one.
- **Alternating**: the speaker continues talking by offering an alternative locution to the previous one.
- **Contrasting**: the speaker continues talking by contrasting a new locution in relation to the previous one.
- **Embedding**: the speaker embeds a locution into another, by e.g. apposition.
- **Reporting**: the speaker reports a locution from another speaker.
- **Turn Taking**: the speaker takes a turn right after the previous speaker.

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to use each of these.

2. Reconstruct argumentation

This task aims to generate the necessary propositions so that the contents of each locution are expressed in a clear, explicit and self-contained manner.

The outcome of this task is an augmented argumentation model, which contains a collection of propositions as well as the previously added locutions and transitions, plus the necessary connecting illocutionary forces.

Create propositions

A proposition is a state of affairs about the world, that is, a statement that expresses something. A proposition is always anchored on a particular locution, in the sense that the locution constitutes the way in which a speaker has expressed the associated proposition. The work necessary to obtain a proposition from its associated locution is called "reconstruction".

Propositions must be self-contained, so any deictic or phoric references in the locution must be resolved. Deictic references include pronouns referring to people or things in the world that are not explicitly mentioned, such as "She expressed some concerns". Phoric references include references to other propositions in the discourse, such as "That is true". However, some conventional deictic or phoric references must not be resolved, such as generic first-person plural pronouns that refer to everybody (such as "the pandemic surprised us") or "it" in reference to the weather (such as in "it is sunny today").

Also, propositions must have a form that is as simple as possible, and it is recommended that a subject plus verb plus complements structure is used for better understandability. Finally, a proposition must express what the speaker apparently intended to express through the associated locution. Usually, the reconstructed proposition looks very much like the original locution, although sometimes it can be quite different.

Consider the locutions identified in the previous section:

LO1. The speed at which the virus spread surprised us

LO2. we do not yet have all the answers to the question

LO3. in recent weeks very important discoveries have been made

LO4. that help us understand what is happening.

These could be reconstructed as propositions in this manner:

PR1. The speed at which the virus spread surprised us.

PR2. We do not yet have all the answers to the question of why the virus spread so fast.

PR3. Very important discoveries have been made in recent weeks.

PR4. The discoveries made in recent weeks help us understand what is happening.

Note how PR1 is very similar to the original LO1. Note also that, in PR2, "the question" has been expanded to avoid the anaphoric reference to "why the virus spread so fast". Finally, note that, in PR4, which corresponds to the subordinate clause in LO4, a subject has been added for explicitness. In this manner, the obtained propositions constitute self-contained sentences that make statements about the world.

Some locutions, such as questions or colloquial expressions, may be difficult to reconstruct into propositions by applying the guidance in this section. Also, locutions with no argumentative relevance should not be reconstructed into propositions at all. Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to tackle different situations.

Finally, when creating propositions, consider their adherence to each of the positions identified during context analysis. If a proposition shows a clear adherence to a position, whether to support or oppose it, record this. For example, proposition PR2 in the example above opposes the position "The mechanisms by which SARS-CoV-2 spreads are well known".

Characterise propositions

If relevant to the project, propositions can be characterised through several variables: statement type, factual aspect, ontological aspect, modality, tense and truth value. Some of these variables affect others, so it is recommended that they are tackled in this order. Characterising propositions can help with some aspects of agency analysis, such as determining the tone of the discourse or detecting what is good or bad for different agents.

Please see the *IAT/ML Technical Specification* document for detailed definitions of each of the following variables and options.

Statement type

The statement type of a proposition indicates whether the proposition expresses a fact or a value. Before this can be determined, you must decide what is the group of people where the proposition takes place. This usually corresponds to the speaker plus the intended audience. Once the group is clear, you must evaluate whether the proposition is free from values that are not fully shared by all the members of this group. If it is, then it is a fact (e.g. "This car is white", "I have a headache"); otherwise, it is a value (e.g. "The Beatles are the best band ever").

Factual aspect

The factual aspect of a proposition indicates what the proposition is about. There are two major kinds of propositions in this regard: static, which express a relatively stable state of affairs, and dynamic, which express a process or changing situation. Static propositions can be further classified depending on the type of fact that they express: existence, identity, predication (attribution or relation), classification or subsumption. Dynamic propositions can be classified depending on the kind of action they describe: activity, accomplishment or achievement. In general, the factual aspect of a proposition is given by its main verb.

Once you have determined a proposition's factual aspect, you can determine its *propositional centre*, that is, the main thing that the proposition is about. For example, the centre of relation propositions (e.g. "Children go to school") is the relation itself, not the things being related. Knowing the propositional centre is useful to determine a proposition's ontological aspect.

Ontological aspect

The ontological aspect of a proposition conveys two closely related pieces of information: the ontological domain of the proposition, and its possibility aspect. The ontological domain refers to whether the propositional centre (as determined by the proposition's factual aspect) pertains to the logical, physical or social world. In this regard, the rules of logic are supposed to be unchangeable across possible worlds, the rules of the physical world are supposed to be changeable across possible worlds but unchangeable in the world where the proposition takes place, and the rules of society are supposed to be changeable across societies within this world.

The possibility aspect, in turn, refers to whether the proposition is expressing something that is impossible, possible, necessary or contingent.

For example, "Children go to school" is a socially necessary proposition. Firstly, its propositional centre, the *go to* association, is determined by social mandate, regardless of the fact that children and schools are physical things. Therefore, the proposition pertains to the social domain. Secondly, the proposition is stating something that is socially mandatory, and hence the proposition is classified as socially necessary.

Modality

The modality of a proposition describes the relationship between the proposition and the world. Roughly, this corresponds to whether the proposition is realis (i.e. expresses something that is, was or will be) or irrealis (it expresses something that may or must be). Realis propositions can

be classified as indicative (expressing a fact about the world) or declarative (changing the world through words). In turn, irrealis propositions can be further classified as epistemic (expressing an agent's beliefs) or deontic (expressing a state of the world that an agent considers desirable), and several additional subtypes within these.

Modality establishes the *direction of fit* of the proposition, that is, whether it is the words which fit the world, as in e.g. indicative or noetic propositions, or whether it is the world which fits the words, as in e.g. declarative or deontic propositions. The direction of fit is useful to determine a proposition's tense.

Tense

The tense of a proposition indicates the time frame that the proposition is about. To determine the tense, you must consider the proposition's grammatical tense of the main verb but also its direction of fit, as determined by the proposition's modality. For example, a proposition may have a verb that is not in the future tense but nevertheless refers to the future, as in for example "They should build a theatre in my neighbourhood". Possible tenses are past, present, future and atemporal. Atemporal propositions are those that make general statements about the world, such as laws or patterns (e.g. "Babies are born after nine months").

Truth value

The truth value of a proposition indicates whether the proposition is true, false or indeterminate. This must be determined by looking at the proposition plus the context whether it takes place. Indeterminate propositions are those for which neither true nor false make sense, such as deontic propositions (e.g. "They should build a theatre in my neighbourhood").

Note that ontological aspect and truth value are related. Necessary propositions are always true; impossible propositions are always false; contingent propositions may be true or false.

Add illocutionary forces

Once propositions have been obtained, illocutionary forces must be added to connect them to their associated locutions. A locution-based illocutionary force is a connection between a locution and a proposition, in terms of speaker intent. In other words, illocutionary forces describe the intention of the speaker when uttering each locution and producing each proposition.

Most locutions, and almost every proposition, must be involved in an illocutionary force. There are different kinds of illocutionary forces that can connect a locution to a proposition, each corresponding to a specific kind of intent by the speaker:

- **Asserting**: the speaker intends to communicate that they believe the associated statement.
- **Questioning**: the speaker intends to obtain new information.
- **Challenging**: the speaker intends to dispute a previous statement.
- **Popular Conceding**: the speaker intends to communicate that they believe a well-known and commonly accepted statement.

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to use each of these.

3. Add argumentation relations

This task aims to generate the necessary argumentation relations to connect the propositions in a meaningful argument network.

The outcome of this task is an augmented argumentation model that contains argumentation relations in addition to the previously created elements.

There are three kinds of argumentation relations:

- **Inferences**, which indicate that one or more premises are provided by a speaker to support a conclusion.
- **Conflicts**, which indicate that a proposition is provided by a speaker as a reason why a target argumentation element is invalid.
- **Rephrases**, which indicate that a proposition is provided by a speaker as a reformulation of another proposition.

As opposed to locutions, it is common that propositions "islands" exist in a model. In other words, not all propositions must necessarily be connected to each other, and it is normal that many of them will remain unconnected.

The outcome of this task is an augmented argumentation model, which adds argumentation relations to the existing propositions as well as the previously added locutions and transitions, plus the necessary connecting illocutionary forces.

Create inferences

An inference is an argumentation relation that indicates that one or more propositions (called the premises) are provided by a speaker to support another proposition (called the conclusion). An inference must have at least one premise, but may have more, and always one conclusion. Often, inferences occur as serial arguments, that is, chains of inferences where the conclusion of an inference works as premise for another inference.

Also, inferences may occur in convergent or linked arguments. A convergent argument is a structure in which one proposition works as conclusion of multiple independent inferences, so all of them converge on the conclusion. A linked argument is a structure composed by a single inference that has two or more premises. Divergent arguments may also occur, when one or more propositions work as premises for multiple independent inferences towards different conclusions.

Inferences can be classified into multiple subtypes, depending on the rationale that allows the speaker to move from premises to conclusion.

Inferences are often marked in the discourse by specific lexical particles, such as "therefore" or "because". Some other times, inferences may lack any explicit marking. Consider the following text:

> We cancelled the morning trip as it started to rain. We phoned mum to let her know.

This can be reconstructed as a serial argument around the following propositions:

> PR1. We cancelled the morning trip.
>
> PR2. It started to rain.
>
> PR3. We phoned mum to let her know that we had cancelled the trip.

Arguably, the speaker is making an inference using "It started to rain" as premise and "We cancelled the morning trip" as conclusion, which is marked by the "as" particle. In other words, the speaker is saying that they cancelled the trip *because* it started to rain. In addition, there is an inference using "We cancelled the morning trip" as premise and "We phoned mum to let her know that we had cancelled the trip" as conclusion. This inference isn't marked by anything but a full stop, but it is evident from the text that the speaker means to convey the fact that they called mum *because* they decided to cancel the trip.

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to create inferences in different situations.

Create conflicts

A conflict is an argumentation relation that indicates that a proposition (called the source) is provided by a speaker as a reason why another argumentation unit (called the target) is invalid. The target may be another proposition or an argumentation relation such as an inference, another conflict or a rephrase. This means that conflicts may be used to attack not only propositions but also relations between propositions.

A rebutting conflict occurs when a proposition directly attacks another proposition, such as in "Immigrants are beneficial to our economy" attacking "Immigrants are a threat to the country".

An undermining conflict occurs when a proposition attacks another indirectly by rebutting one of its premises. For example, consider the inference in "We cancelled the morning trip as it started to rain", and the response "But it didn't rain until the following day". This statement is rebutting "It started to rain", and since this is used as a premise, it also attacks, albeit indirectly, "We cancelled the morning trip". In other words, the second speaker is undermining the statement "We cancelled the morning trip" by attacking its premise.

Finally, an undercutting conflict occurs when a proposition attacks another indirectly by attacking an inference having it as conclusion. For example, consider the inference in "We cancelled the morning trip as it started to rain", and the response "But you could have gone anyway". This statement is attacking the reasoning behind the inference, that is, it is challenging the idea that one must cancel a trip if it rains. Thus, it is undercutting the statement "We cancelled the morning trip" by attacking an inference that leads to it.

Conflicts are sometimes marked by lexical particles such as "but".

Consider the following example:

> [00:01:21] Brennan: The New Yorker reported that when her name was first floated to you by- by Senator Daniel Patrick Moynihan, you had some misgivings. Why?
>
> [00:01:28] Clinton: I don't remember that. I, first of all, there were- everybody that I know is taking credit now for- for 27 years for nominating her, but I didn't have misgivings.

Some of the propositions that may be reconstructed from this text are:

> PR1. Clinton had misgivings about nominating her. (Reconstructed from "you had some misgivings" by Brennan)
>
> PR2. Clinton didn't have misgivings about nominating her. (Reconstructed from "but I didn't have misgivings" by Clinton)

Here, PR2 is clearly attacking PR1; there is a rebutting conflict using PR2 as source and PR1 as target.

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to create conflicts in different situations.

Create rephrases

A rephrase is an argumentation relation that indicates that a proposition (called the source) is provided by a speaker as a reformulation of another proposition (called the target). Usually, the source proposition is very similar to the target, but it often involves some variation. Rephrases are often used to provide nuance, add extra details, summarise or reconduct the dialogue.

There are many subtypes of rephrases:

- **Abstraction**: the speaker repeats the target proposition but raising the level of abstraction by e.g. generalising or classifying. An example would be "Construction of the

cathedral started in the 10th century" being rephrased as "Most cathedrals in the area were built around that time".

- **Concretion**: the speaker repeats the target proposition but lowering the level of abstraction by e.g. exemplifying or decomposing. An example would be "Dogs often catch canine influenza" being rephrased as "My dog had influenza when she was three".
- **Definition**: the speaker unpacks a term in the target proposition by describing its meaning. An example would be "My husband works 60-hour weeks" being rephrased as "The family provider works a lot".
- **Naming**: the speaker provides a term to name an idea in the target proposition. An example would be "This town concentrated commerce, arts and influx of people for many decades" being rephrased as "The town was a regional capital".
- **Agreement**: the speaker expresses agreement with the target proposition. An example would be "The cathedral needs urgent restoration" being rephrased as "Indeed it does".
- **Answer**: the speaker answers the question in the target proposition. An example would be "The painting was recovered, or the painting was not recovered" being rephrased as "The painting was never recovered".
- **Repetition**: the speaker literally repeats the target proposition, or a part of it, for emphasis. An example would be "Immigration benefits labour markets" being rephrased as "Immigration does benefit labour markets".
- **Paraphrasis**: the speaker repeats the target proposition by recasting it in different words that result in a mostly lexical or syntactic change, but not that much in a rhetorical or pragmatic sense. An example would be "Immigration benefits labour markets" being rephrased as "Immigration is a boost for labour markets".
- **Reinterpretation**: the speaker reinterprets the target proposition by changing its contents but without frontally contradicting it. This often includes mechanisms such as analogies, adding emotional nuance, straw man fallacies, etc. An example would be "Children should not be allowed to play in the street by themselves" being rephrased as "Children should not leave the home by themselves".

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to create rephrases in different situations.

Add illocutionary forces

Once propositions have been interconnected via argumentation relations, illocutionary forces must be added to connect these to their associated transitions. Transition-anchored illocutionary forces connect a transition to an argumentation relation, in terms of speaker intent. In other words, these illocutionary forces describe the intention of the speaker when making each transition and producing each argumentation relation.

Most transitions, and almost every argumentation relation, must be involved in an illocutionary force. There are different kinds of illocutionary forces that can connect a transition to an argumentation element, each corresponding to a specific kind of intent by the speaker:

- **Arguing**: the speaker intends to support an inference.
- **Agreeing**: the speaker intends to react affirmatively to a proposition through a rephrase.
- **Disagreeing**: the speaker intends to react negatively to a proposition through a conflict.
- **Restating**: the speaker intends to recast a proposition through a rephrase.

Please see the *IAT/ML Argumentation Patterns Guidelines* document for detailed guidance on how to use each of these.

4. Add denotations

This task aims to enrich propositions by using denotations so that each proposition is connected to the relevant ontology elements.

The outcome of this task is an augmented argumentation model that contains denotations for each proposition.

If an ontology is available for the text being analysed, denotations can be added at this stage. Denotations connect terms in propositions (and the associated locutions) to ontology elements. For example, proposition PR1 "The speed at which the virus spread surprised us" in the example above refers to "the virus", meaning the SARS-CoV-2 virus. If there is an atom if the ontology for this, then we should add a denotation connecting the fragments "the virus" in PR1 to the SARS-CoV-2 atom in the ontology.

Adding denotations can be time-consuming, so you may want to do it only if you are interested in a denotational analysis or a joint evaluation of argumentation and ontological issues in your text.

5. Revise the model

This task aims to iterate the model so far constructed, revise its semantics, and modify it as needed to fix mistakes, complete missing details, or capture additional nuances.

The outcome of this task is an improved model. You can revise the model again if you feel the need, until it eventually stabilises.

Figure 3 shows a sample diagram of a finished argumentation model.
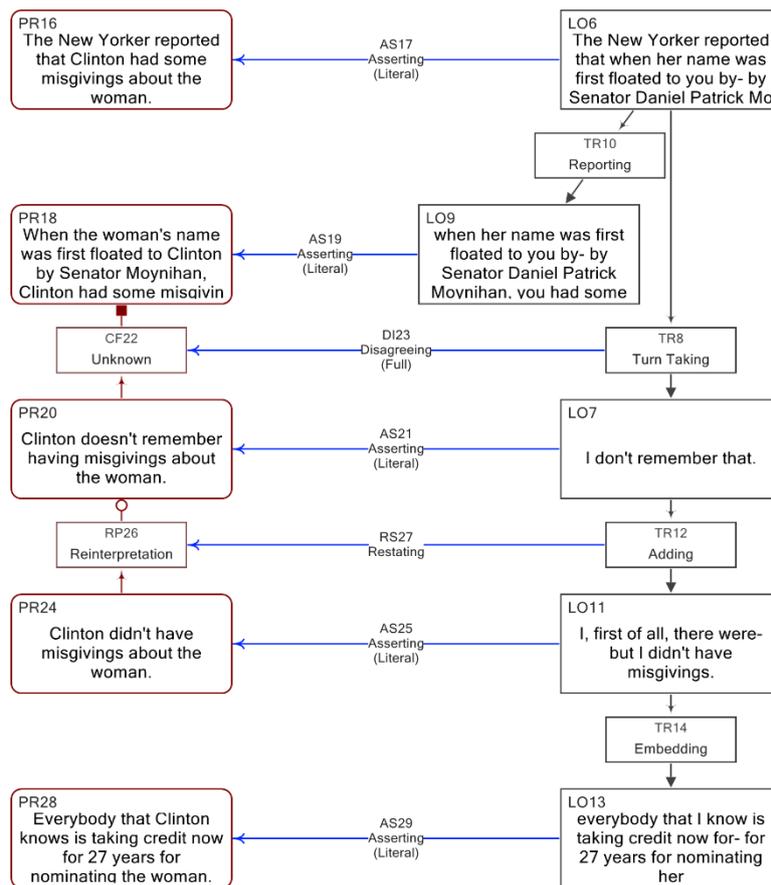
Figure 3. Sample argumentation model, depicted as a diagram. The right-hand side column of boxes represents the locutions and transitions of the speakers. The left-hand side column represents the propositions reconstructed by the analysts plus the argumentation relations among them. Horizontal arrows depict illocutionary forces.

You may also want to revise the context that you defined during project initiation, and update its positions and agents according to the information revealed during argumentation analysis.

Agency Analysis

The purpose of this process is to analyse a body of text to create an agency model. An agency model is a representation of the beliefs, desires and intentions of the speakers in the text and their linked context agents.

The text that you use as a basis to develop an argumentation model usually pertains to a single document. In some situations, it may make sense to create an agency model for a set of related documents, but this is rare.

Developing an agency model usually involves the following tasks:

1. Define the questions to ask
2. List scopes
3. Develop responses to the questions
4. Summarise results
5. Revise the model

Please see the *IAT/ML Agency Patterns Guidelines* document for additional details.

1. Define questions

This task aims to produce a set of questions that will be asked in relation to the text to be analysed. When working with a corpus, a single set of questions is often used to analyse all or a large subset of the documents.

A question in agency analysis is intended to be "asked" to the text, such as "What are the major agents in the text?". Different questions address different issues, and working with between 10 and 30 questions is not uncommon.

IAT/ML comes with a standard question set (see the *IAT/ML Agency Patterns Guidelines* document), which you can readily use. However, you can add, modify or remove questions depending on your research goals.

The outcome of this task is a list of questions, usually organised in groups, the responses of which are expected to provide insights into the text being analysed.

Design question groups

Questions are better arranged in groups so that different aspects of the agency analysis are addressed in turn. Also, bear in mind that some question groups are better tackled once other groups have been responded, so that groups should follow an order of increasing complexity and degree of subjectivity.

The standard question set in IAT/ML uses the following structure:

- **Direct questions**. This group includes questions that require no or little interpretation, and whose responses are usually objectively available from the text itself.
  - **Form**. This subgroup includes questions related to the surface form of the text, such as emphasis or highlights.
  - **Themes**. This subgroup includes questions related to theses and themes dealt with by the text.
  - **Rhetoric**. This subgroup includes questions related to the tone, lexical choices and metaphors used in the text.
- **Interpretive questions**. This group includes questions that require significant interpretation, and whose responses are defined by the text but also by its environment, speakers, situation and other contextual aspects.
  - **Agents**. This subgroup includes questions related to the agents that appear in the text.
  - **Situations**. This subgroup includes questions related to the situations in which agents are portrayed to be in the text.
  - **Rhetoric**. This subgroup includes questions related to strategies, hypotheses and intentions of the speakers.

This structure can be altered depending on the research needs.

Write questions

Once a question group structure has been decided, questions for each group and subgroup must be developed.

Some questions pertain to the text as a whole, such as "What agents or agent types are referred to by the text?". Questions like this are called unscoped, are "asked" to the text as a whole, and a single response is expected.

Some other questions are related to each speaker in the text, in two possible ways:

- The question is about what each speaker says in the text, such as "What are the major cognitive metaphors in the text according to each speaker?". These are called subject

speaker-scoped questions, because the speakers act as subjects, i.e. the question focuses on the speakers' views as expressed in the text. Questions like these are often indicated by the wording "according to each speaker" or similar.

- The question is about how each speaker is portrayed in the text, such as "What terms are used to address each speaker?". These are called object speaker-scoped questions, because the speakers act as objects, i.e. the question focuses on the speakers' characterisation. Questions like these are often indicated by the wording "for each speaker" or similar.

Other questions may be scoped to each agent that appears in the text, such as "What is beneficial for each agent?". These are called object agent-scoped questions. Questions like these are often indicated by the wording "for each agent" or similar.

Other kinds of entities, such as organisations or places, may also work as scopes, e.g. in "What does each place provide to each agent?". These are called object entity-scoped questions. Questions like these are often indicated by the wording "for each X", where "X" is the kind of entity it refers to.

In general, scoped questions must be responded for each scope combination. For example, "What does each place provide to each agent?" must be responded once for each place and agent combination.

Questions must be clear, succinct and self-contained. They must be formulated as interrogative sentences that require a short response. The expected response may be a short text, such as in "What is the intention of the speaker?", a list of short texts, as in "What is good for each agent in the text?", one or more options from a predefined list, such as in "What is the genre of the text?", or one or more agents from the linked context, such as in "What agents are portrayed in a positive light by each speaker?".

For each question, you should provide some general guidance on how to respond it, as well as specific guidance on how responses to the question contribute towards describing the beliefs, desires and intentions of the speaker. Not all questions are expected to contribute to all three aspects.

2. List scopes

If the question set that you are using includes scoped questions, you must list the relevant scopes before you attempt to respond the questions, as follows:

- **Subject or object speaker-scoped**. If your question set contains either kind of speaker-scoped questions, identify and list the speakers in the text.
- **Object agent-scoped**. If your question set contains object agent-scoped questions, identify and list the agents that are mentioned in the text.
- **Object entity-scoped**. If your question set contains questions scoped to other kinds of entities, populate the associated entity lists. For example, if there is a "Places" entity list, find and list the relevant places that are named in the text.

Note that you may need to alter or refine your lists later on, while responding questions.

The outcome of this task is a set of lists, which may include speakers, agents and other entities.

3. Develop responses

This task aims to respond to the questions that have been previously defined. The *IAT/ML Agency Patterns Guidelines* document lists the questions in the standard question set and provides specific guidance on how to respond each. You can use this standard question set as is, modify it, or create your own set altogether.

When working with a corpus, you may choose between responding all the questions for one document, then moving to the next document, and repeating this until all the documents have been processed; or you can instead respond the first question for each document in the corpus, then move to the next question, and repeat this until all questions have been responded. You can also choose to group the documents in the corpus, or in a subset of the corpus, as a single text, and respond the questions for the aggregated text.

When working out a response for each question, you must consider what kind of response is expected (text, itemised, option list or agent), as well as whether the question requires a separate response for each of the applicable scopes. For example, "Who are the friends of each agent?" requires a separate itemised response for each agent in the text. Bear in mind that speakers may work as subject and object scopes, or even both at the same time. For example:

- "What is the genre of the text?" is an unscoped question, so you must develop a single response.
- "What are the central entities according to each speaker?" is a subject speaker-scoped question, so you must develop a response for each speaker in the text by capturing what the speaker says about central entities.
- "What terms are used by each speaker to describe other speakers?" is a subject speaker-scoped question and also an object speaker-scoped question, since speakers are involved twice: as subjects who employ terms to describe others, and as objects who are described by others. In consequence, you must develop a response for each pair of speakers.
- "What things are forbidden for each agent according to each speaker?" is a subject-speaker-scoped question and also an object agent-scoped question. You must develop a response for each speaker and agent combination.

In any case, responses must always be anchored on the text. In other words, when developing a response to a question, make sure that you indicate what elements in the text support your response. If the text doesn't contain enough information to develop a clear response, you should probably skip the question rather than trying to guess a response.

When you progress through the questions by responding each one, you will probably develop a hypothesis about the speakers' beliefs, desires and intentions. As you advance through the question set, you will be able to test your hypothesis against questions that become more and more interpretive. If at any point you find that your hypothesis is wrong, you must go back and revise the responses obtained so far to attempt to generate a new hypothesis. This process is repeated until a solid hypothesis is obtained and responses to the questions acceptably support it.

Finally, when creating responses and their parts for each of the applicable scopes, consider their adherences to the positions identified during context analysis. If a response part shows a clear adherence to a position, whether to support or oppose it, record this. For example, imagine that the response to "What are the central entities according to each speaker?" for a particular speaker is "Immigration, Unemployment, Crime". This supports the position that "Immigration and crime are important social issues for the citizens".

The outcome of this task is one or more responses for each question.

4. Summarise results

This task aims to summarise the responses previously obtained in terms of beliefs, desires and intentions of each individual speaker in the text, and then grouping them according to their linked agents.

To develop these summaries, take one speaker at a time and focus on each of the three agency aspects (beliefs, desires and intentions) in turn, revise the responses for questions pertaining to the target aspect, and write an adequate summary. For example, take the first speaker in the text and start by focusing on their beliefs. Find the responses to those questions that contribute to identifying agent beliefs, and draft a summary of their responses. Then, move to the next speaker, and repeat until you have finished. Then, repeat the whole process for desires, and then for intentions.

Then, consolidate the summaries of each speaker that is linked to an agent into a short text that outlines the beliefs, desires and intentions for the agent as a whole. Construct this from the associated speakers.

The outcome of this task is a short summary for each speaker and agent in terms of beliefs, desires and intentions.

5. Revise the model

This task aims to iterate the set of responses and summaries that have been obtained, make sure that there is continuity and integrity across them, and make the necessary changes so that they are complete and consistent.

The outcome of this task is an improved model. You can revise the model again if you feel the need, until it eventually stabilises.

You may also want to revise the context that you defined during project initiation, and update its positions and agents according to the information revealed during agency analysis.

Analytics

During this phase, you apply automated analytical algorithms to the discourse models from the previous one, and obtain quantitative and visual results.

You can carry out many analytics processes, depending on the needs of your project. Some projects do not need any analytics at all, whereas others rely strongly on the quantitative and visual results that they can provide.

Analytics processes are implemented by computers, and each one has its own characteristics in terms of what input model and parameters they need, what kind of results they produce, and how these results are presented, either as lists, tables or charts. The online *LogosLink User's Manual* provides a detailed description of the analytics available from this tool. You can also export analytics data and process them further or produce additional visualisations by using other tools.

Results

During this phase, you elaborate products that you or a third party can apply to address the target situation of your project.

Typical results include documentation, diagnosis and action planning documents, but many other types are possible. Since different projects may have very different requirements in relation to what outputs are expected, you can "plug in" your own results process at this point to obtain a customised output.

In any case, the products of results processes are usually custom-made for a particular recipient, and their contents are tailored to their specific roles and capacities. It is usually this party who uses the results produced in this phase to act or intervene in relation to the situation or problem being addressed. This allows you to select the most appropriate third party to apply the results of the methodology depending on the specific domain area and needs.

Documentation

The purpose of this process is to describe and characterise in detail the target situation of your project in terms of themes, positions and agents. This usually involves the following tasks:

1. Describe the situation being addressed and the associated issues
2. Briefly describe the methodology as enacted
3. Outline the context, in terms of themes, positions and agents
4. Develop a detailed characterisation of the themes, positions and agents in the context, any linked elements in other analysis areas (ontological, argumentation an agency), and connections between them
5. Revise and refine

Little or no interpretation should be necessary for documentation. You are supposed to describe your findings as objectively as possible.

1. Describe the situation and issues

This task aims to produce a brief characterisation of the situation under study, together with a list of the relevant issues that have been identified

The outcome of this task is a short text.

2. Describe the methodology

This task aims to describe the IAT/ML methodology as enacted in your project. That is, don't describe the abstract methodology; rather, explain how you applied it to your project, what decisions you made, and what interim products you obtained.

For example, describe the corpus criteria that you developed, plus the number and size of the documents that you gathered and analysed. Explain also what kinds of analysis (ontological, argumentation and/or agency) you carried out, and why you decided to use or not use each of them.

The outcome of this task is a short text.

3. Outline the context

This task aims to describe the context of the work, that is, the themes, positions and agents that have been defined and studied.

The outcome of this task is a structured text.

4. Develop characterisation of themes, positions and agents

This task aims to describe the different context elements (themes, positions and agents) in full detail.

To do this, collect analysis and analytics results throughout the project, and present them so that they can be easily understood and integrated. In addition, consider linked elements in each analysis realm. For example, speakers in argumentation models may be linked to agents, in the sense that each speaker is considered to be representative of one or more agents. Look at what these speakers are saying and how they are arguing, and present it in relation to their linked agents.

Integrate the result so obtained with the texts produced by the previous tasks in this process.

The outcome of this task is a draft document.

5. Revise and refine

The outcome of this task is an improved document. You can revise it as many times as necessary if you feel the need, until it becomes stable.

Diagnosis

The purpose of this process is to describe and characterise in detail the target situation of your project in terms of themes, positions and agents. This usually involves the following tasks:

1. Document the situation
2. Add an interpretation or discussion of the previous
3. Provide information on how this diagnosis differs from previous diagnoses of the same situation or problem, if there are any
4. Describe any future perspectives
5. Revise and refine

A diagnosis involves developing a documentation first, and then elaborating an interpretation of it that related themes, positions and agents among themselves and in relation to the social situation being addressed. A significant amount of interpretation is often necessary for diagnosis.

1. Document the situation

This task aims to produce a detailed characterisation of the situation under study.

Please see the *Documentation* results process, page 34, for detailed instructions on how to do this.

The outcome of this task is a diagnosis report.

2. Add an interpretation or discussion

This task aims to provide an expert interpretation or discussion of the previous characterisation in relation to selected relevant issues.

It is not necessary that you address all the issues that you identified. You can select a few and focus on them, leaving the others for future projects. Determine which ones to address according to their relevance, prominence within your corpus, analysis results, and potential contributions to the project.

You can do this in many non-exclusive ways, such as the following:

- Focus on especially relevant issues and develop additional insights.
- Establish comparisons or contrasts between different themes, positions or agents in the context.
- Compare the findings with common assumptions or previous information.

In any case, make sure that your interpretation is clearly anchored on the characterisations developed during the previous task.

Also, it is often a good idea to involve third parties in the development of the interpretation. As described at the beginning of this section, the diagnosis documents is usually targeted to a particular recipient. You can involve them in the discussion and interpretation of the results at this stage.

The outcome of this task is an expanded diagnosis document.

3. Provide comparative information

This task aims to produce a comparison of the diagnosis with previous versions of similar documents.

It may be the case that a diagnosis of the same situation or problem is carried out repeatedly over time, in order to assess how the situation changes. If this is so, you should describe how the current diagnosis matches or differs from previous versions.

The outcome of this task is an expanded diagnosis document.

4. Describe perspectives

This task aims to provide a brief discussion about future improvements to the diagnosis.

You can focus on issues that were identified but which you decided not to address at the moment. You can also mention any other aspects of the situation under study that you think should be addressed in the future.

The outcome of this task is an expanded diagnosis document.

5. Revise and refine

The outcome of this task is an improved diagnosis document. You can revise it as many times as necessary if you feel the need, until it becomes stable.

Action Planning

The purpose of this process is to provide advice to the recipient on how to mitigate or solve the problems associated to the target situation of your project.

Action planning does not recommend a particular course of action. Rather, it lists possible action options and describes the likely consequences of each one. In this manner, the recipient of the action plan can decide what is the best course of action by assessing the context at any particular moment and looking at the expected consequences of each possible action option.

Action planning usually involves the following tasks:

1. Develop diagnosis of the situation
2. Establish goals
3. Develop action options and consequences
4. Revise and refine

1. Develop diagnosis

This task aims to produce a detailed characterisation of the situation under study.

Please see the *Diagnosis* results process, page 35, for detailed instructions on how to do this.

The outcome of this task is a diagnosis report.

2. Establish goals

This task aims to produce a list of goals to be addressed by the action plan.

Goals are usually influenced by the issues that you have identified. For example, the issue of "Why are immigrant kinds doing worse at school?" can suggest a goal such as "Reduce the school drop-out rates of immigrant kids by 20% in the next 5 years".

It is often unrealistic to address all the issues that you have identified; rather, you can focus on addressing one or a few of them. To decide what goals to address, look at the diagnosis report and the most relevant issues. For each of them, consider its feasibility and urgency, as well as any potential dependencies on other problems. Also consider the expected ability of the intended recipient.

Bear in mind that most goals that you can realistically address by using IAT/ML are related to discourses. It is unlikely that you will be able to produce significant social change that is unrelated to discourses. In the example about reducing the school drop-out rates of immigrant kids, you may be able to influence the prevalent discourses of kids, teachers and parents about immigrant students, but not other aspects of immigration such as poverty or refugee crises.

No matter how many issues you have identified, be clear about which ones you are addressing via goals and which ones you are leaving out for the future.

The outcome of this task is a list of goals to be addressed by the action plan.

3. Develop action options and consequences

This task aims to produce a detailed list of action options for each goal, plus foreseeable consequences for each option.

Action options may be formulated in many ways, e.g. "when you speak to agent X about position Y, use this type of arguments and address themes A and B"; "raise theme X with agents A and B, but do it separately"; or "talk to agent X and tell them about this and that in order to make them aware of the consequences of their position". Many other formulas may be useful.

When developing action options, consider the different context aspects:

- **Issues**. Consider what issues you are dealing with, so that you keep the focus on it.
- **Themes**. Consider what themes may be raised with whom, and what consequences should be expected from doing it.
- **Positions**. Consider what positions may be promoted or adopted, with whom, and what consequences should be expected from doing it.
- **Agents**. Consider what agents may be engaged, in relation to which themes and positions, and what consequences should be expected from doing it.

For each action option, explain:

- How useful it may be, in terms of results.
- How easy or difficult it should be to implement, in terms of expected barriers.
- What factors exist that contribute to making the option attractive.
- What factors exist that contribute to making the option unattractive.
- What associated risks of taking vs. not taking it exist.

Consequences, in turn, should be expressed in terms of the impacts (positive and negative) that each option is expected to have in relation to each agent. It is likely that some option will benefit some agents but harm others.

The outcome of this task is a structured text document, the action plan, that includes the diagnosis of the situation, lists the established goals, and describes the associated action options and consequences.

4. Revise and refine

This task aims to iterate the action plan so far developed, revise the goals, options and consequences, and modify them as needed to fix mistakes, complete missing details, or capture additional nuances.

The outcome of this task is an improved action plan. You can revise it as many times as necessary if you feel the need, until it stabilises.

References

[1]   CBS News, "Transcript: President Bill Clinton on 'Face the Nation,' September 20, 2020," *CBS News*, Sep. 20, 2020. https://www.cbsnews.com/news/transcript-president-bill-clinton-on-face-the-nation-september-20-2020/ (accessed Apr. 06, 2022).