Chapter 3

Narratives in Historical Sciences
Remi van Trijp, Inès Blin

Abstract

Historical sciences such as geology, evolutionary biology or history try to offer causal explanations for non-recurrent phenomena (e.g. how the Grand Canyon was formed, how the human eye evolved, or what caused the Second World War), typically using incomplete and fragmentary evidence from the past. Even though these sciences make use of general frameworks such as the Theory of Evolution, they have to work out the specifics of each case and hence they cannot simply apply general laws and deductive reasoning. Instead, causal explanations are expressed in narrative form. While such explanations have long been considered to be "less scientific", there is now a growing awareness that narrative explanations go beyond mere description and also have the potential for empirical testing. This Chapter explores how narratives are used by historical scientists, and how human-centric AI systems may assist scientists in constructing more precise and testable narratives so they can achieve a deeper understanding of society. It presents a first prototype that takes as its case study the French Revolution (1789–1799).

3.1 Introduction

Narrative explanations play a major role in historical sciences such as geology, evolutionary biology, and history that have to provide explanations for non-recurrent phenomena based on evidence that is often incomplete and fragmented. The importance of narratives for scientific explanation has long been downplayed by philosophers of science ((Mary S. Morgan and Norton Wise 2017; Mat and Mary S. Morgan 2019), also see (Carlo R.M.A. Santagiustina 2022) in this volume), but since the second half of the 20th Century there

has been a growing awareness that scientific narratives aren't simple "just-so stories" and that they play a much more important epistemic role than assumed before (see e.g. (Goudge 1958) for one of the earliest influential repositionings of narratives in science philosophy).

This Chapter explores how narratives are used by historical scientists, and how human-centric AI systems may assist scientists in discovering and constructing more precise and testable narratives so they can achieve a deeper understanding of the world and society. More specifically, it explores the two major strategies employed by scientists: particular narratives, which aim to explain particular phenomena and events in a "standalone" fashion; and embedded narratives, which aim to identify a phenomenon or event as an instance of a recurrent pattern. It also defines a narrative as a three-layered structure that includes a fabula, plot and narration. The Chapter concludes with a first prototype for constructing narrative networks inspired by the French Revolution (1789–1799).

3.2 Particular and Embedded Narratives

Just as there exist many different kinds of narrative structures in literature, there exist different types of scientific narratives depending on the *explanandum* (i.e. the phenomenon that needs to be explained). According to (Currie 2014), there are roughly two explanatory strategies that historical scientists use, which he calls *simple* and *complex* narratives. The following two subsections provide an illustration of these two different strategies based on the French Revolution (1789–1799) (Maurois 2017; Lefebvre, Guyot, and Sagnac 1951),[1] but note that we renamed them to *particular* (instead of *complex*) and *embedded* (instead of *simple*) for reasons that we will explain in section 3.2.3, which will also explain how both strategies work.

3.2.1 The French Revolution as a Particular Narrative

In 1789, France was a large and prosperous country and one of the 18th century's world superpowers. Yet it was about to slide into one of the most important revolutions in history, whose impact is still felt today in western civilization. How could this happen?

For starters, there was something rotten in the state of France. While there was sufficient wealth in the country, the *government* was bankrupt, partly because of its expensive wars and support for the American revolution, but mostly because of exorbitant food prices and structural problems with the taxation system. In an attempt to replenish the state coffers, the French king Louis XVI made a decision he would soon come to regret: he called for a general assembly of the so-called *Estates* – the three social orders of French

[1] Both narratives reflect our understanding of the French Revolution from different sources. We underline however that we are not historians, and that scholars of the French Revolution may disagree with our narrative choices.

society of the time, consisting of the clergy (First Estate), nobility (Second Estate), and Commoners (Third Estate) – hoping that they would approve of new taxes and solve his debt crisis.

By asking for new taxes, however, the king had failed to read the proverbial room: a strong class resentment had been stirring in the country, which opposed especially the Third Estate to the first two estates. The representatives of the Third Estate, which included lawyers, local officials, and wealthy land owners, were upset about the growing inequality in French society: since each Estate had one vote, the clergy and nobility could always gang up against the commoners. This meant that the Commoners, which represented by far the biggest slice of the population, had to carry almost all of the economic burden and pay most of the taxes, while getting little political influence and social status in return.

The representatives of the Third Estate decided they would no longer put up with a system that they felt was rigged against them, and created a National Assembly on June 17, 1789 that would be *representative*: one vote per person instead of one per Estate. A clear act of rebellion, but the king was slow to react, and three days later, the representatives took the famous Tennis Court Oath – literally named after the fact that it was pledged on a real tennis court of Louis XVI – in which they swore to remain assembled until a new constitution was drafted.

The Revolution had begun... And what would follow was a period of radical changes between 1789 and 1799, depicted in a timeline in Figure 3.1. In the span of a decade, the regime of the old government (the *Ancient Régime*) would be overthrown, and a Republic would ultimately emerge after a series of violent confrontations between competing factions of the French population.

From Absolute to Constitutional Monarchy

While the first acts of rebellion were led by a relatively wealthy middle class, the general population had grievances of its own: bread prices were soaring due to crop failures and the *gabelle*, a widely hated taxation on salt.[2] By failing to feed the hungry, France was feeding the momentum of the revolution. Especially the city of Paris became a hotbed for riots, and state properties became targets of attacks and plundering. The most iconic outburst of violence happened on the 14th of July 1789, when the Revolutionaries stormed the Bastille, a medieval fortress and prison that represented in their view the monarchy's abuse of power. The fall of the Bastille went down in history as a watershed moment of the French Revolution: the point of no return.

Other major events followed rapidly. On August 4, still in the year of 1789, the National Assembly abolished feudalism, which marked the end of the three-estate system. Later

[2]The *gabelle* was abolished by the French revolutionaries in 1790, only to be reinstated by Napoleon in 1806. Still hated as an unfair taxation on the poor, it would finally be terminated in 1945 after the liberation of France at the end of the Second World War (Chazelas 1968).

that month, August 26, there was the Declaration of Human Rights, which granted human civil rights to a large portion of the population, and which became a major stepping-stone towards more equitable democracies.

A real tipping point in the transfer of power from the monarchy to the revolutionaries was the Women's March on October 5, 1789. The March started with a protest against high bread prices by women who worked at the marketplaces of Paris. A coalition of angry market vendors, revolutionary agitators and other protestors soon formed a mob of thousands of people who marched to the palace of Versailles where the king was residing. The mob forced the king to return with them to Paris and to accept the legitimacy of the National Assembly. Within a matter of a few months, the monarchy had lost almost all of its absolute power to the revolutionaries.

A Game of Thrones

Now that the Absolute Monarchy had been brought down to its knees, the Assembly tried to restore peace in the country. This became apparent at the *Fête de la Fédération* (Festival of the Federation) that was held on the 14th of July 1790, the first birthday of the storming of the Bastille: instead of glorifying the bloodshed of the year before, the festival aimed to foster national unity and even reserved a role for the king, who pledged his allegiance to the new constitution.

However, under these still waters, danger was lurking below. Different factions with literal cutthroat politics emerged that started to compete with each other for power. The first majority of the National Assembly was a coalition of moderate deputies known

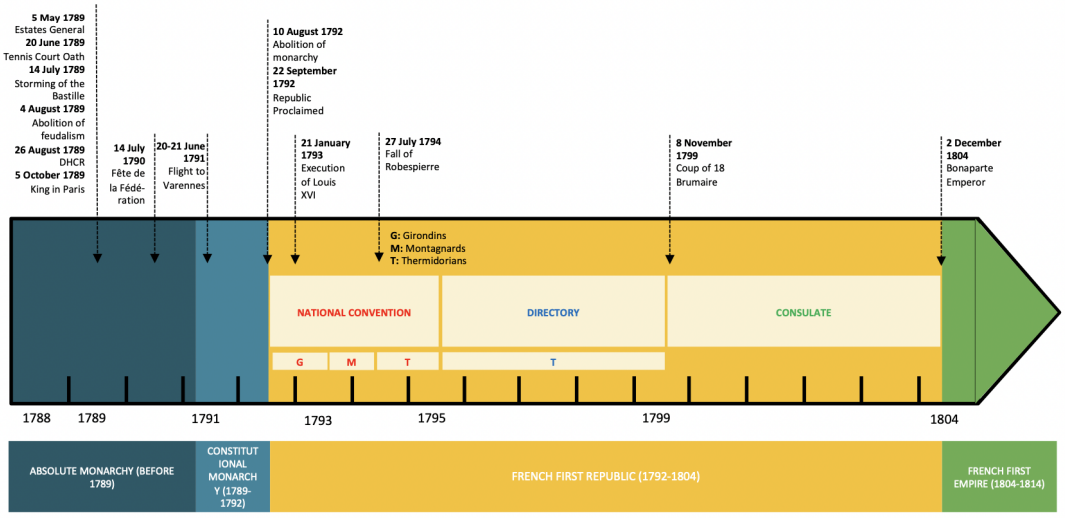


Figure 3.1: Timeline of the main events and political regimes or governments during the French Revolution (1789–1799). Translated from French and modified from: "Les temps forts de la révolution française", https://yann-bouvier.jimdofree.com/ressources/histoire/chronologie-revolution-francaise/ .

as the *Feuillants Club* who wanted to preserve a role for the king in a constitutional monarchy, but they were under attack from both reactionaries who wanted to restore the Ancien Régime as well as from a heterogeneous group of revolutionaries and anti-royalists, especially the infamous *Jacobin Club*.

Tensions quickly rose, both within and outside of France. Other European monarchies such as Austria and Prussia were getting increasingly worried that the revolution would spread to their country, and contemplated attacking France to defend the Ancien Régime or simply to take advantage of the country's inner chaos. Moreover, in June 1791, the king and his family tried to flee from Paris, presumably to start a counter-revolution. Rather than intimidating the French, these events only fuelled the flames of the revolution and public opinion radicalized even further, which made the Jacobin Club the post powerful political faction in France.

Some Jacobin deputies soon made calls for a revolutionary war against Austria and Prussia, and forced out the Feuillants from the Assembly – having hundreds of their rivals arrested and tried for treason. The ousting of the Feuillants meant that the Assembly was now more than ever prey to the power of the political clubs instead of being an autonomous governing body. However, the Jacobin Club was not a coherent group and was undergoing a power struggle itself. Especially Maximilien Robespierre rose to prominence by denouncing the war plans of the members he mockingly called the *Girondins* (named after their home department in the Southwest of France). Even though the Girondins still had the upper hand and managed to push their war through the Assembly, they would start a fierce rivalry with the dissenting Robespierrist movement, which became known as *La Montagne* (the mountain) because its members (the *Montagnards*) sat on the highest benches of the Assembly.

The war declared against Austria in April 1792 started off disastrously, however, and the Girondins were blamed for the losses suffered during the first battles. Moreover, anger had been boiling up about foreign counter-revolutionary threats and the king's failed flight attempt, which led to the insurrection of August 10, 1792, in which armed revolutionaries stormed the Tuileries palace in Paris where the king and his family were staying. In the meantime, many volunteers driven by nationalism strengthened France's army, leading to a stunning victory against invading forces on 20 September 1792. On the same day, a new Assembly called the National Convention was formed, which abolished the monarchy and proclaimed the First Republic of France.

Amidst all these events, the Montagnards seized the opportunity to take control of the country. While the Girondins hesitated about what to do with the king, the more radical Montagnards took a hard stance against Louis XVI, which gained them support from the lower class commoners (known as the *sans-cullottes*), who felt betrayed because the Girondins would not establish universal rights extending to all citizens. Accused of being royalists, the Girondins were purged from the National Convention in 1793, giving full reign to the Montagnards and its leader Robespierre.

However, as members of the bourgeoisie, the Montagnards were under pressure to satisfy

the needs of the radical sans-cullottes, which resulted in several policies enacted in favor of helping lower-class citizens and the poor, but also in extremely violent persecution of anyone who was accused of being an anti-revolutionary. Conspiracies were formed against the Montagnard's "Reign of Terror", however, which ended in a coup d'état on 9 Thermidor, year II (27 July 1794 on the French revolutionary calendar). Robespierre, the man who used to be called the "Incorruptible", was guillotined the next day. This *Thermidorian Reaction* would be marked by persecution of former Jacobins and other people who were associated with Robespierre.

The new regime, called the Directory or the Thermidorian Convention, tried to stabilize the Republic but was facing many struggles such as rebellions from royalists, former Jacobins, and the perpetual wars between revolutionaries and counter-revolutionaries across Europe. These struggles ended when a strong leader took over: general Napoleon Bonaparte. Bonaparte had led many victorious military campaigns, and ended the Directory when he came back to France through the coups of 9–10 November 1799. A new government, called the Consulate, was founded, which many historians consider to be the end of the French Revolution.

3.2.2 The French Revolution as an Embedded Narrative

Why do revolutions occur? Are there regularities or patterns in the dynamics of a revolution? One attempt at answering these questions is the influential book *The Anatomy of Revolution* by Crane Brinton (1965). First published in 1938, the book compares the outbreak and progress of revolutions to the outbreak and progress of fever, as summarized in Figure 3.2. Brinton examined and compared four revolutions: the British revolution of 1677, the American revolution (1776–1784), the French Revolution (1789–1799), and the Russian revolution of 1917.

Preliminary Stage Symptoms

Before the outbreak of the "disease", there are usually already some symptoms present. In all four case studies, Brinton noticed that the government was experiencing financial troubles, even though the society itself was prosperous. In France, the government went bankrupt not because of a lack of wealth, but because of structural problems with its tax collection system, and because attempts to reform taxes were blocked by the ruling elite. Just like the other three cases studied by Brinton, the French government was weak and inefficient.

At the same time, a prosperous middle class emerges with grievances about its socio-economic status and growing inequality. In France, these were "commoners" that included lawyers, businessmen and land owners. The Commoners were the so-called Third Estate who paid a disproportionate amount of taxes and who lacked political power because the other two estates – the clergy and nobility – would always side against them. In America

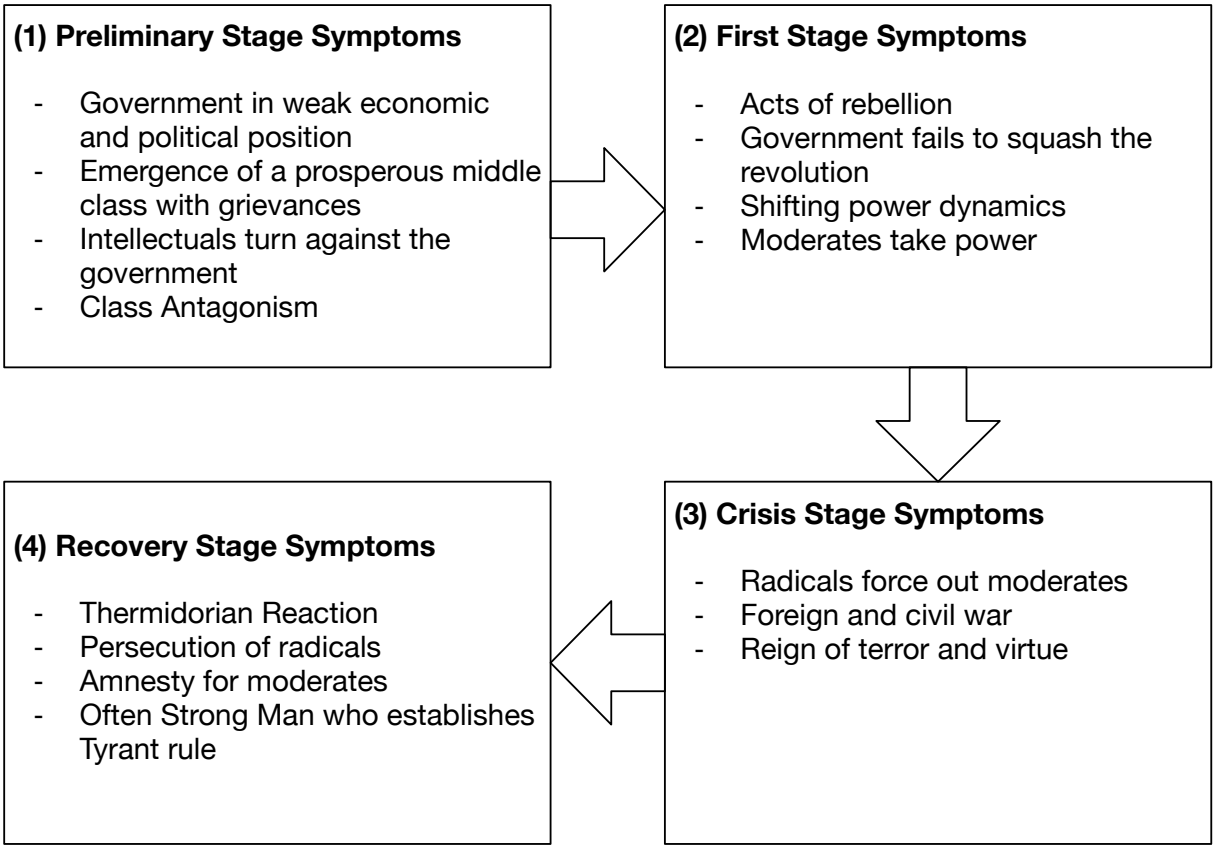


Figure 3.2: This diagram summarizes the scientific narrrative of (Brinton 1965), who compared the dynamics of a revolution to fever symptoms. All case studies are presented as narratives that are embedded in this overall structure.

these were merchants who resented the British ruling class for similar reasons. In each case, this middle class also found support from intellectuals, who increasingly turned away from the governments to demand far-reaching reforms.

First Stage Symptoms

Brinton then states that revolutions go through different stages similar to the progression of fever, with moderate symptoms at first but soon leading to a state of delirium. In the first stage, first actions are taken against an unpopular economic situation (particularly taxes), and at least two groups become opposed to each other. In France, the Third Estate of Commoners rebelled by forming a representative Assembly that would protect them from being outvoted by the First and Second Estates, which created an opposition between royalists and revolutionaries. In each case study, the government was too slow to react or failed to suppress the opposition by force, and the power dynamics shift to the new group, of which the most moderate ones take up leadership first. In the French Revolution, the Absolute Monarchy was abolished in favor of a constitutional monarchy.

Crisis Stage Symptoms

In the next stage, revolutions reach peak fever in which increasingly radical factions force out the more moderate groups through coups d'état, culminating in a reign of terror where any (perceived) opposition to the revolution is violently squashed. The Jacobins in France, the Bolsheviks in Russia, and the Sons of Liberty in America were all well-organized and disciplined but radical factions that did not hesitate to use violence to get their ways. Once in power, leaders such as Robespierre in France or Lenin in Russia prove to be authoritarians, and install a reign of terror in which all (perceived) opposition to the revolution and the new government is violently squashed. Often these swift changes in government happen against a backdrop of foreign and civil war, which pressures the governments into rapid centralization and zero-tolerance policies.

Thermidorian Reaction

According to Brinton, there is only so much that a society can take, so reigns of terror are met with a "Thermidorian Reaction", which is like a recovery from the fever. In the French Revolution, the Reaction stared with the arrest and execution of Maximilien Robespierre, which ended the Reign of Terror of the Montagnards. During this period, there is usually a prosecution of the most radical revolutionaries, while moderates receive amnesty. Often, however, such post-revolutionary societies remain unstable until a strong man such as Napoleon Bonaparte seizes control. Brinton's conclusion is that the French, English and Russian revolutions "began in hope and moderation, all reach a crisis in a reign of terror, and all end in something like a dictatorship – Cromwell, Bonaparte, Stalin" (Brinton 1965, p. 23).

3.2.3 Discussion of the Strategies

Section 3.2.1 illustrates what is arguably the most common scientific narrative in the historical sciences: the particular narrative. A particular narrative is an explanation for what the scientist identifies as a single event without necessarily trying to find out what is general or universal about the event. (Currie 2014) calls these narratives *complex* because they are typically high in *detail*: the explanation requires *specificity* and a high *diffusion* of information in order to be adequate. Indeed, one can fill a whole library section with studies that focus entirely on the events that led to the French Revolution and how the revolution unfolded over time without *embedding* these studies into a larger theory about what "revolutions" are.

Consider now section 3.2.2 on the other hand. In this narrative, the focus is not on the details or specifics of the French revolution, but rather how it illustrates a regular pattern of what constitutes a revolution (or at least those studied by (Brinton 1965)). In other words, the narrative is *embedded* in a larger framework so the specifics can be left out.

This is why (Currie 2014) calls these narratives *simple*: the information presented is less detailed, less specific and less diffuse than for the particular narrative.

We prefer the terms *particular* and *embedded* narratives because embedded narratives can become quite complex and detailed in their own right ((see e.g. Palmer and Armitage 2014, for a complex narrative that relates several revolutions to each other)). Another way to look at the two strategies is to draw analogies to *literary narratives* such as novels on the one hand (particular narratives), and *literary criticism* (that is, the comparison, analysis and evaluation of literary works) on the other (embedded narrative). Indeed, a work of fiction may contain intertextual references to other art, but must essentially be able "to stand on its own feet." Likewise, a linguist who studies the French language may refer to syntactic rules in related languages, but must ultimately provide a self-contained explanation for the grammar of French. A literary critic, then, finds motifs and themes across texts. And likewise, linguistic typologists may develop a comparative theory about the world's language structures.

If we want to develop human-centric AI systems that can assist historical scientists in constructing their narratives, we therefore need to be aware of which narrative strategy the scientist wishes to employ: a particular narrative, which offers a full and detailed explanation of a single event; or an embedded narrative, in which the focus lies on the identification of recurrent patterns.

3.3 Fabula, Plot, and Narrative

> *"The king died and then the queen died is a story. The king died, and then queen died of grief is a plot."* (Edward Morgan Forster)

So far we have used the term *narrative* in an intuitive way, but if we wish to design human-centric AI systems that make use of this concept, we need to provide a more concrete definition. This is a non-trivial task, since the western tradition of narratology goes all the way back to Ancient Greece, which has led to often conflicting views of what constitutes a narrative. We will therefore offer here some tentative definitions that serve as the foundations of our MUHAI research, but which will be further fleshed out throughout the project. More concretely, we will distinguish three concepts that together constitute a narrative:

- The *fabula* (often called the *story*) is a collection of actions, events, or facts.
- A *plot* (also called the *syuzhet*) is a structure that arranges the elements of the fabula in a causal chain or causal network.
- A *narration* (also called the *discourse* or *narrative presentation*) concerns *how* the narrative is presented.

3.3.1 The Fabula and the Dimension of Veracity

As explained in the introductory chapter of this volume (Luc Steels 2022), one of the most important dimensions in scientific narratives concerns the *veracity* of the narrative. Indeed, the primary concern of any scientific discipline is getting its facts straight. The foundational layer for every scientific narrative is therefore what is called the *fabula* in narratology, which roughly means "the story as it actually happened."

The goal of the fabula is to have a collection of facts and descriptions of events that are as objective and close to reality as possible, without trying to relate those facts with each other. For example, the Florentine *Catasto* is a historical record that offers historians raw data about the tax assessment of households in Florence and its surrounding territories between 1427 and 1429 (Herlihy and Klapisch-Zuber 1985). The Catasto includes, among others, the assets and debts of Florentine households, which helps to estimate how wealthy these households were.

In order to obtain a high degree of veracity, historical scientists must have a good estimate of the *reliability* of their evidence. For example, they must verify whether the Florentine Catasto indeed includes all or at least most of the Florentine households. When compared to historical population estimates, it turns out that the Catasto probably underreports the number of inhabitants, but that the difference is small enough to remain representative (*The Florentine Catasto of 1427* n.d.).

3.3.2 The Plot: Plausibility and Meaningfulness

The study of narratology has distinguished the fabula from the *plot* ever since its introduction by the Russian formalist school of literary criticism (Erlich 1973). The plot provides structure and coherence to the elements of the fabula through causal links, so that previously unconnected facts can now travel together as a group.[3] The simplest plot is a chain of causal links that ends with the historical event that a researcher wants to explain, though some events may be so complex that it is more appropriate to represent the plot as a *causal network*.

For example, suppose we want to explain why France celebrates its national holiday on the 14th of July. The fabula contains several facts, such as the fact that there was an attack on the Bastille on the 14th of July 1789, that there was a *Fête de la Fédération* (Festival of the Federation) on the same day in 1790, and that the 14th of July became the national holiday in 1880 – almost a century later. The plot, then, is a *plausible* causal chain that leads from the first event in 1789 all the way to the establishment of the current national holiday.

We emphasize the word *plausible* here because establishing a causal link between two events is a non-trivial task. Often the historical record consists of fragmented and incomplete pieces of evidence, so the scientist is forced to posit conjectures and fill in the gaps.

[3] The original notion of *syuzhet* conflated what we call plot and narration in this paper.

Even when there is abundant data available, causality remains tricky. A classic example is establishing the cause of death of a person, which is why physicians and medical examiners receive explicit training in verbalizing their uncertainty when signing a person's death certificate (Hanzlick 1997). The more uncertainty there is, the more hypotheses (or plotlines) may emerge in the scientific literature.

As explained in the introductory chapter of this volume (Luc Steels 2022), we consider something *meaningful* if it is *relevant* for a particular task. Constructing a plot is therefore essentially a *meaning-making* process because the historical scientist needs to decide which facts from the fabula are relevant for their scientific explanation. For example, on the 1st of July 1989, just after the start of the French Revolution but two weeks before the storming of the Bastille, the ballet *La fille mal gardée* (The Wayward's Daughter) of Jean Dauberval premiered at the Grand Théatre de Bordeaux in France (Guest 1960). Even though this piece is important for modern ballet, it is irrelevant for understanding what happened in the French revolution and is therefore meaningless in this context.[4]

3.3.3 The Narration: Structure, Style and Narrative Positioning

The third layer of a narrative concerns the *narration* or the way in which a narrative is presented to its audience. We can distinguish three dimensions in this layer: structure, style and narrative positioning.

Structure and Style of the Narration

The structure of the narration is the order in which the events of the plot are exposed. The most straightforward structure is to present the events in chronological order, in which case the structure of the narrative follows the direction of the plot. Indeed, the plot and the structure of the narrative have often been conflated, but the difference between the two have become more clear as artists have expanded their storytelling devices, such as the use of flashbacks or flashforwards in movies.

One good example is the 1994 movie *Immortal Beloved*, written and directed by Bernard Rose, which starts with the death of Ludwig Van Beethoven (played by Gary Oldman), whose testament states that all of his assets will be left to his "immortal beloved." But who is this person? The audience then accompanies Beethoven's assistant and friend Anton Schindler (played by Jeroen Krabbé), who tries to solve this mystery by visiting all of the women who played a role in Beethoven's life. During each visit, we get to see a

[4] This is not to say that a performance can never be meaningful in a historical narrative. For example, the opera *La Muette de Portici* (The Mute Girl of Portici) by Daniel Auber is often said to have played a role in creating unrest that started the Belgian revolution of 1830, though historians have downplayed its actual importance (Slatin 1979).

flashback from Beethoven's life, each time providing a piece of the puzzle until the whole plot is revealed and we figure out the identity of the maestro's immortal beloved.

Changing the narrative structure from chronological reporting to a more complex narration can be very effective for creating a compelling scientific narrative as well. One excellent example is Chapter 7 of *Life Ascending: The Ten Great Inventions of Evolution* by the evolutionary biochemist Nick Lane (2009). Early on in the Chapter, which explains the evolution of sight, Lane uses foreshadowing to whet the reader's appetite as follows:

> *"[T]he rise of molecular genetics in the last decades furnishes us with a wealth of detail, giving very particular answers to very particular questions. When these answers are all threaded together, a compelling view emerges of how the eye evolved, and from where – a surprisingly remote and green ancestor. In this chapter, we'll follow this thread to see exactly what use is half an eye, how lenses evolved, and where the light-sensitive cells of the retina came from. And in piecing together this story, we'll see that the invention of eyes really did alter the pace and flow of evolution..."* (Lane 2009, p. 175)

In the remainder of the chapter, rather than strictly adhering to the chronological order of the plot, Lane follows the chronology of the scientific discoveries that each provided pieces of the puzzle, much like a detective novel follows clues until the reader can reconstruct the causal chain of events.

But there is more to the structure of narration than simply changing the order of information. In her 1998 book, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (1998), Digital Media professor Janet Murray predicted that the rise of interactive media would drastically impact the structure of narration because the audience is invited to play a role in creating that structure. One simple example is a website with hyperlinks: the order in which information is presented to the reader is changed whenever they click on one of the hyperlinks.

In scientific narratives as well, new forms of digital storytelling allow for the audience to co-create the structure of the narration, and in some cases even the narrative itself. One example are the CLARIAH data stories[5], which are historical narratives (e.g. about the wealth of Florence during the Renaissance) that include interactive code blocks that contain queries on a knowledge graph. Another example is the Google N-gram Viewer that allows users to type in phrases and then see how often those phrases occurred over time in a corpus of books. If the graph shows that one phrase suddenly becomes popular while another one declines, the user may become intrigued and may want to figure out what caused this shift to happen.

Whereas the structure of the narration concerns high-level choices, the *style* of the narration focuses on more concrete considerations such as register (i.e. the degree of formality of a text), verbosity, lexical diversity, grammatical complexity, and so on. Style choices

[5] https://stories.datalegend.net/

mostly depend on the goal (e.g. informing, educating, awareness-raising, promoting, and so on), modality (e.g. presentation, research paper, social media posts and blog posts, interviews, and so on) and intended audience of the narrative (e.g. children, general public, experts, policy makers, and so on).

Narrative Positioning

Stereotypically, a scientific narrative is meant to answer a question, such as how the human eye evolved. In reality, however, researchers and research groups engage in an activity that can be called *narrative positioning* (Berry 2021). Related to the concept of *framing*, narrative positioning is a technique for situating research with respect to other research endeavours using a narrative form.

The current volume of articles, especially its introductory chapter by Luc Steels (2022), is a prime example of narrative positioning by explaining how meaningful and human-centric AI compares to data-driven AI. Narrative positioning thus has a major impact on every part of a research programme, starting with which questions it is interested in answering, what kind of experiments need to be designed, which kind of technologies are needed, what kind of measures and evaluation criteria can help to track the progress of the research programme, and how results should be interpreted.

The importance and impact of narrative positioning shows that narratives not only have a major epistemic role to play in scientific explanation, but even affects and steers scientific practice itself.

3.3.4 Towards Meaningful and Human-Centric AI

Now that we have identified three important layers that make up a narrative (the fabula, plot, and narration), we can also identify how narratives can play a role in the design of meaningful and human-centric AI. In this paper, we are mainly concerned with AI systems that assist historical scientists in constructing scientific narratives in a more efficient and reliable way and to provide them with tools for better understanding complex issues in society, but the same principles could be applied to other professions as well, such as journalists who need to report on complex issues such as the COVID-19 pandemic with rapidly changing insights and a flood of information and disinformation to go through.

- *Data Veracity*: Human-centric AI systems must provide adequate knowledge representations and tools that facilitate the verification of data veracity and the assessment of reliability of pieces of evidence. This is crucial for constructing the fabula of both historical events (in which data is fragmented, incomplete and sparse) as well as contemporary events (in which knowledge is still evolving, and in which it is often difficult to distinguish facts from unfounded claims and misinformation).

- *Narrative Networks*: Meaningful AI systems must provide tools to assist humans in constructing a plot or narrative network. Examples include, but are not limited to, the extraction of causal relations from natural language texts ("understanding"), selection of relevant (and therefore meaningful) pieces of knowledge representations (e.g. through heuristic search on ontologies and knowledge graphs) for explaining a particular topic ("production"), and so on.
- *Narrative Matching*: AI systems may also assist human researchers in detecting regularities and patterns across narratives, allowing them to discover embedded narrative structures that are otherwise difficult to find because of the explosion of data and human confirmation bias.
- *Narration*: Human-centric AI systems may provide adequate tools for efficiently mediating between a human communicator and their intended audience. These may include interactive web interfaces and searchable graph visualizations ("production") and awareness-raising visualizations about information and opinion spaces ("comprehension").
- *Narrative Positioning*: A human-centric AI system must be able to explain its objectives and decisions in narrative form, including an assessment about uncertainty or decisions made based on reactive intelligence (see (see Luc Steels 2022, in this volume)).

3.4 Case Study on the French Revolution

This section presents a first prototype that explores some of the issues in the design of meaningful and human-centric AI systems that were discussed in the previous sections. More specifically, we focus on a *particular narrative* using the French Revolution as a case study, similar to the one presented in section 3.2.1. Our prototype includes the three layers of a narrative in the following ways:

1. The fabula is represented as a *knowledge graph*, taken from Wikidata in English, which is assumed to be high in veracity.
2. The plot is operationalized as a *narrative network* in which meaningful/relevant events are ordered in a chronological sequence.
3. The narration follows the sequence of the plot, and includes an interactive visualization of the narrative network.

3.4.1 The Fabula as a Knowledge Graph

The first layer of a narrative is the fabula, which can be thought of as a factbase. In our prototype, we used a *knowledge graph* (more precisely, the English Wikidata knowledge

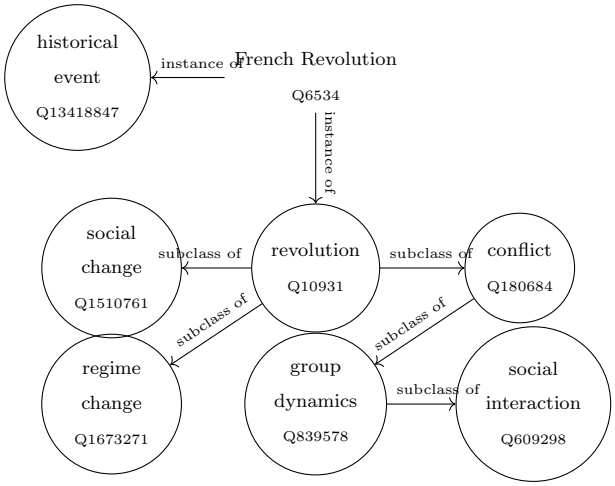


Figure 3.3: A subgraph of Wikidata centered around the search phrase "French Revolution". The nodes were manually selected to make sense of the "social and political revolution" description.

base) to represent the fabula of the French Revolution. A knowledge graph is a semantic network that may contain information about real-world entities (i.e. objects, events, situations, concepts, and so on) and relations between them. The content is encoded with the help of appropriate standardized knowledge representations, also called ontologies.

While it is impossible to have a totally objective representation of reality, we can assume that the fabula represents the ground truth and therefore scores high in the dimension of veracity (see section 3.3.1) if there is sufficient consensus about its facts. Since Wikidata is a collaborative and open endeavour, we will assume here that all of its statements are objectively true. Secondly, we assume that the fabula only contains statements that concern the four WH-questions about an event: who, what, when and where.

3.4.2 The Plot as a Narrative Network

A historical scientist who wants to construct a plot that connects all of the relevant events of the French Revolution through causal relations could in principle explore the knowledge graph manually. The French Revolution[6] is described in Wikidata as a "social and political revolution." Which facts or knowledge can corroborate this description? Figure 3.3 shows a subgraph in Wikidata centered around the search phrase "French Revolution". Exploring the graph manually from this starting point, one can see that the French Revolution is an instance of a revolution, which has interesting properties such as "social change", "regime change", or "conflict", which matches the description of the revolution as a socio-political event.

However, these links do not yet offer a narrative explanation of which events occurred during the French Revolution and how they relate to each other. The historical scientist

[6] https://www.wikidata.org/wiki/Q6534

therefore must construct a plot to make sense of what happened, starting by collecting all of the facts that are meaningful or relevant for explaining the French Revolution. Here, we manually explored which paths in the knowledge graph lead to the relevant events of the revolution.

In total we retrieved 48 events for which Wikidata often offers adequate information about the what, when and where questions (i.e. the type of events, their time and place), but little about the who question (i.e. the participants in those events). Sparsity of data is a common problem with structured data that any AI-system must anticipate, and to which we will return later in this paper. For our case study, we solved this problem of missing information by leveraging *Infoboxes* from Wikipedia.[7] A Wikipedia Infobox is a table with textual properties and attributes that summarizes the most important content about the current page. Moreover, these infoboxes also contain URL links to other Wikipedia pages. For example, if an Infobox about an event X contains a cell named "combattant" that links to a page about person A, then we can assume that there is a latent link between A and X that is missing from the knowledge graph. Using this information from Wikipedia, we were able to enrich the fabula with information about participants in the French Revolution (e.g. person A participated in event X), and to establish causal links in the plot (e.g. event X caused event Y).

Besides the selection of meaningful facts, constructing the plot of the narrative network requires the development of timelines that put events in their chronological order, and ultimately establishing a causal explanation link between events.

Figure 3.4 and Figure 3.5 display timelines that could be manually constructed from the data collected for the whole French Revolution and the French First Republic respectively. Above the blue arrow are the political regimes (first row) and governments (second row) during each period. Below the blue arrow is the timeline built with the events collected in Wikidata and Wikipedia. Events are ordered chronologically. *pre_event* and *post_event* indicate that one event was before of after another one respectively. The difference with *event_start* and *event_end* is that in the latter, one event is precisely started by another one, i.e. there are timestamps overlaps. Finally, numbered events are main events chronologically ordered identified during this period. As for the colours, a quadrille background is an event that was not retrieved originally in Wikidata, whereas a full background was retrieved in Wikidata. Dashed surroundings indicate that no corresponding Wikidata entity was found, whereas full surroundings indicate that it was.

If Figure 3.4 and Figure 3.5 permit to identify the main events during each period of the French Revolution, there are still some links missing to understand the whole picture. In particular, some links are missing or the outcomes of the events are unknown. For instance when there is a coup d'état, was it successful?

The causal line depicted in Figure 3.7 and manually built displays events together with causal links. The legends are the same than in Figures 3.4 and 3.5. A full square on the left of a node is a victory, whereas a quadrille square is a defeat.

[7] https://en.wikipedia.org/wiki/Help:Infobox

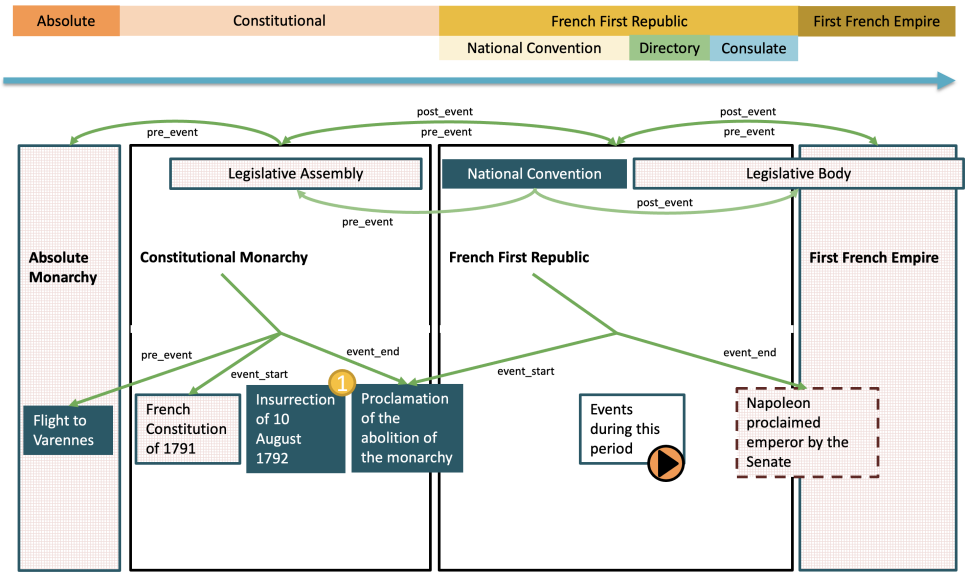


Figure 3.4: Temporal timeline extracted for the French Revolution. One can identify important events in relation to bigger events or at the intersection of other events. For instance, the event "Proclamation of the abolition of the monarchy" happens in-between the "Constitutional Monarchy" and the "French First Republic" events, and given the name of the event it is plausible that it played an important role in the transion of the two bigger events. Likewise for "Napoleon proclaimed emperor by the Senate". One can also understand how legislatures, political regimes follow each other.

3.4.3 The Narration of the Plot

Once the historical scientist has constructed a plot, they need to narrate the plot in an adequate form to their audience. For example, they can write a short essay in the style of section 3.2.1 of this paper, illustrated with timelines such as the ones of Figures 3.4, 3.5 and 3.7. They could also use novel digital tools such as an interactive web demonstration to involve the audience more intimately into the process of narrative construction, or to allow a critical reader to verify a closer inspection.

A small web demonstrator was implemented for the case study on the French Revolution. Several pages and options are available: i) collecting events about the French Revolution. The user can select paths in the graph from which to extract the events. The user can then select the type of data that should be extracted: Wikidata only or also Wikipedia, text content from Wikipedia. ii) extract infoboxes from Wikipedia. For each Wikipedia page identified for one Wikidata page, the infobox (if any) is extracted, and url links are additionally stored. Some processing options are also available. iii) build a first network. This first version would not be a finalised narrative networks since it is triples directly extracted from Wikidata, and thus do not contain events and their description (who, what, when, where). The timelines and cause lines described in Section 3.4.2 were manually built and not automatically derived from these triples. iv) visualise the network. Events were manually ordered and it is possible to slide over events over time, seeing new entities appearing over time.

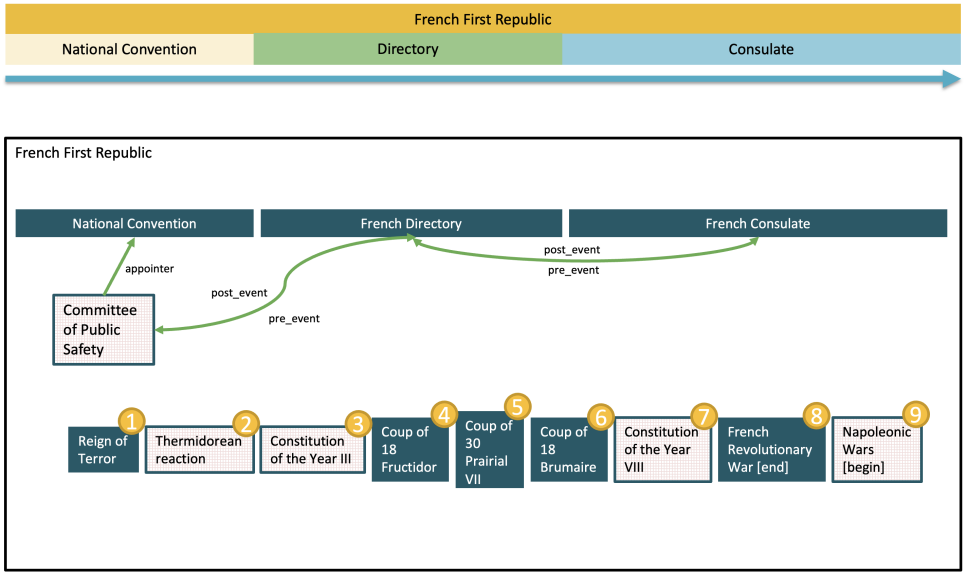
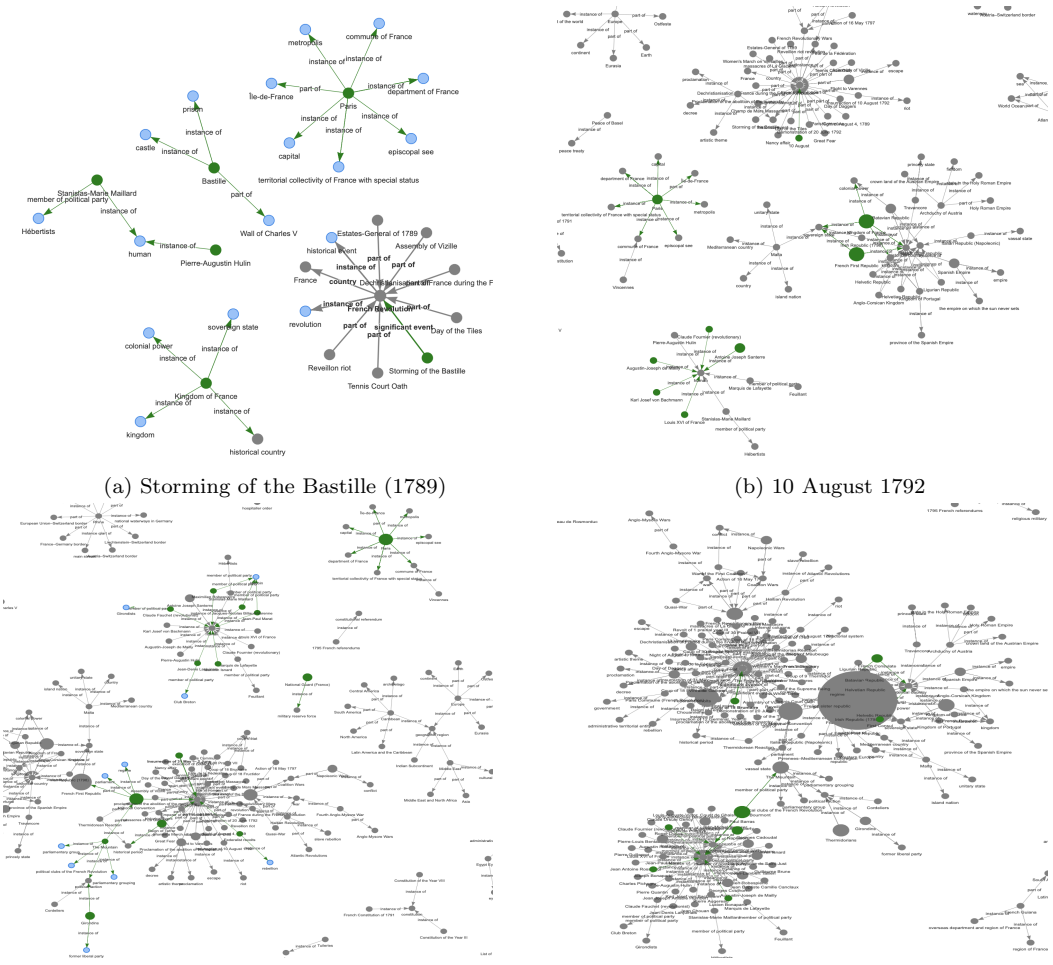


Figure 3.5: Temporal timeline extracted for the specific period of the French First Republic. One can identify the three Conventions that were described in Figure 3.2.1. One can also see the main events identified for this period.

Figure 3.6 displays a visualisation of the graph at different steps of the construction. For purposes of visualisation, the size of a node was increased if it was encountered multiple times across all infoboxes. Therefore, visually bigger nodes represent entities that are likely to have a more important role in the French Revolution. At each step for a given event, green nodes represent entities that were found in the event infobox, whereas blue nodes represent additional information. Grey nodes are nodes that are already part of the graph.

Figure 3.6a shows the entities related to the Storming of the Bastille. It is one of the first events identified during the French Revolution, where the insurgents took the Bastille as a sign of protestation of the royalist power. Different types of entities are added: the event itself, but also other types of categories like people and location (both geographical and historical). One interpretation of this experience could be that Stanislas-Marie Maillard and Pierre-Augustin Hulin both participated in the Storming of the Bastille, that happened at the Bastille prison in Paris, during the historical period of the Kingdom of France. Across the four steps displayed in Figure 3.6a, categories identified remain similar. For instance, Figure 3.6b shows the entities identified for the event of 10 August 1792. During that insurrection, Republicans were against Royalists and ended victorious, foretelling the end of the constitutional monarchy. Likewise, Figure 3.6c shows the graph built at the moment of the Insurrection of 31 May - 2 June 1793, that resulted in the fall of one political faction of the National Convention during the First French Republic, the Girondins, and the rise of another one, the Montagnards. Lastly, Figure 3.6d shows the final graph, after adding the entities from the Coup of 18 Brumaire in 1799 that brought General Napoléon Bonaparte to power, and that in the view of many ended the French



(a) Storming of the Bastille (1789)

(b) 10 August 1792

(c) Insurrection of 31 May - 2 June 1793

(d) Coup of 18 Brumaire (1799)

Figure 3.6: Graph construction at different expansion steps.

Revolution.

3.4.4 Discussion of the Case Study

In this section we presented a case study in which we emulated how a historical scientist could use present-day web technologies and knowledge bases for constructing a narrative network, including a representation of the fabula as a knowledge graph, the construction of a plot as a causal network, and a way to involve the reader in the narration through a custom-built interactive web demonstration.

This workflow has the potential to lift scientific narratives to a whole new level of scientific rigour because the research community can more reliably verify which parts of the narratives are grounded in facts (represented as the linked data of a knowledge graph) and whether those facts are high in terms of veracity and reliability; and verify which causal

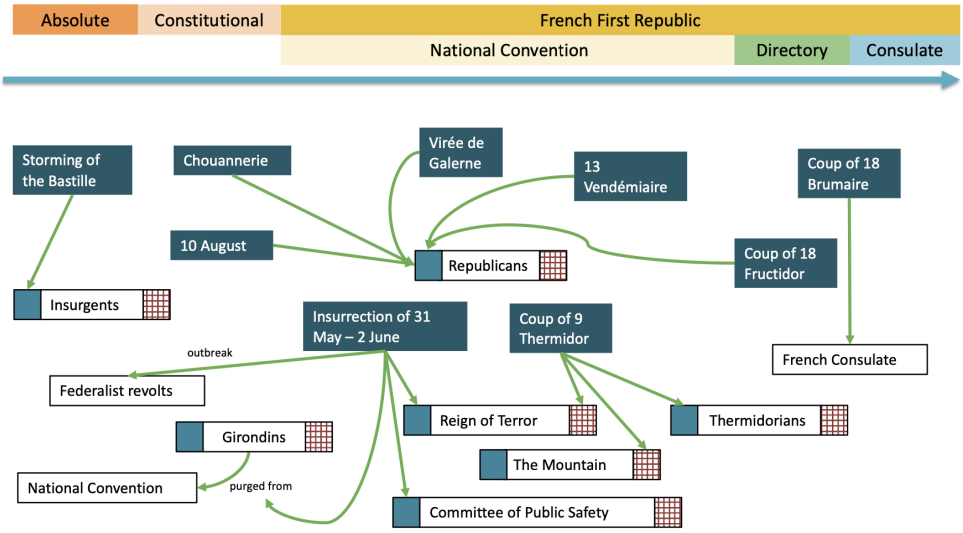


Figure 3.7: Causeline extracted for the French Revolution. The main focus is on the different political parties, and especially during the National Convention. One can see the three conventions on the causeline: the "Insurrection of 31 May - 2 June" is the transition between the Girondins and Montagnards Convention, whereas the "Coup of 9 Thermidor" is the one between the Montagnards and the Thermidorians one. The outcomes are mostly in terms of defeat or victory of politicals parties. One can also see numerous events involving the victory of Republicans: those events were indeed conflicts between Republicans and Royalists.

relations and conjectures were contributed by the scientist, and how plausible these relations are. If the fabula is undisputed, alternative plots (or scientific explanations) could be devised and grounded in the same factbase, which allows for more objective comparison of scientific hypotheses.

Of course, such a workflow would also immensely increase the burden and required skillset of the scientist to accomplish their work. Human-centric AI systems are therefore needed to assist humans in this gargantuan task. In the case of particular narratives, some important AI services could include:

- Retrieval of meaningful events from the factbase for constructing a plot (e.g. using search heuristics), to order those events chronologically, and to suggest potential causal links to the human scientist.
- Repair mechanisms to recover from data sparsity, e.g. through information-retrieval or text-to-knowledgebase parsing. Most information is still expressed as unstructured data such as natural language texts, whose volume is increasing at a much faster rate than that of structured data.
- Authoring tools that automatically ground a piece of text or a web demonstration to the fabula, and that analyze whether the narration covers the plot adequately through measures of comprehension and understanding (see in this volume for more relevant measures (Luc Steels 2022)).

Human-centered AI systems could also provide tools for helping scientists to construct embedded narratives. For example, once a narrative network of the French Revolution has been developed, an AI system could use graph-matching techniques for detecting whether that particular network can be generalized to find regular patterns such as the one presented in section 3.2.2. Not only would an AI assistant be able to find relevant matches at a faster rate than a human researcher could, it could also find generalizations that the researcher did not foresee and thereby overcome confirmation bias.

Alternatively, the human researcher could develop a general narrative pattern (such as the one for revolutions illustrated in Figure 3.2), and use an AI assistant to empirically test the adequacy of that pattern by searching for matches and mismatches in the fabula. Such applications would not only be useful for understanding the past, but also to understand present-day society. One example that is very actual at the moment of writing this article is the question how we will know when the COVID-19 pandemic has ended or at least become endemic. One piece of evidence may come from computational simulations (Lavine, Bjornsyad, and Antia 2021), while more empirical evidence could be mined through particular narratives about how past pandemic ended (e.g. the Plague, smallpox, the great influenza pandemic, SARS and others (Charters and Heitman 2021)). An AI-assistant that understands embedded narratives could help researchers to identify which biological, social, economical and political markers are worth observing, and to rapidly set up observatories that could empirically test the relevance of those markers as the pandemic unfolds over time. Human-centric AI systems could also help policy makers and journalists to better make sense of the numbers, which minimizes the danger of misinterpretation, miscommunication, and subsequent polarization about which safety measures should or should not be taken.

Finally, human-centric AI systems may also help scientists to overcome their own blind spots. As science becomes increasingly specialized, one danger is that different fields become isolated from each other so that a researcher may overlook relevant facts that are outside of their field of observation. To take an example from COVID-19 again, a virologist or epidemiologist can make recommendations for safety measures from their particular viewpoint, but lack the expertise to appreciate what the impact of those measures would be in terms of mental health, economic inequality, and so on. AI systems could help scientists to examine different plots and viewpoints to make better recommendations.

3.5 Conclusion

This paper explored the role of explanatory narratives in the historical sciences and how this could inform human-centric AI systems. More specifically, it showed that the long tradition of scientific narratives is increasingly appreciated by philosophers of science as playing an important epistemic role for understanding the world and society.

This paper then discussed two major narrative strategies – particular and embedded – and illustrated each strategy using the French Revolution. It then defined the structure

of a narrative as containing three layers: the fabula, the plot, and the narration. Through a first prototype, in which we emulated a historical scientist who wants to develop a particular narrative for explaining the French Revolution, we explored both how the notion of a narrative could be incorporated in the design of human-centric AI systems, as well as how such AI-systems may help researchers to reach new levels of scientific rigour and explanatory power in their work.

Acknowledgements

The authors wish to thank all of their MUHAI partners for their invaluable feedback on this work: the team of Luc Steels at the Venice International University, the team of Frank Van Harmelen, Annette ten Teije and Ilaria Tiddi at the Free University of Amsterdam, the team of Rainer Malaka and Robert Porzel at the University of Bremen, the team of Katrien Beuls and Paul Van Eecke at the Free University of Brussels, and Pieter Wellens at Apicbase Antwerp. We also wish to thank our colleagues Martina Galletti, Michael Anslow and David Colliaux for the many fruitful discussions; and Hiroaki Kitano (President and CEO of the Sony Computer Science Laboratories, Inc) and Vittorio Loreto (Director of the Sony Computer Science Laboratories Paris) for creating such a superb and stimulating research environment.