

Iconoclastic Textuality

The Ecclesiastical Proust Archive¹

Jeffrey Drouin

ABSTRACT

David Greetham encouraged the theoretical and methodological flexibility toward text that led me down the path of digital humanities. His introduction of archival and textual theories inspired the Ecclesiastical Proust Archive, an open-ended, experimental project investigating the nature of digital textuality as it embodies the massive À la recherche du temps perdu. In a few short examples, this article lays out some of the ways in which the project takes shape, including a multimedia database, semantic taxonomy, network graphing, and topic modeling.

DAVID'S INFLUENCE HAS LENT A FOCUS ON TEXTUAL STUDIES AND archival theory to all of my scholarship. His single course in textual studies introduced me to the examination of manuscripts and typescripts, the vagaries of editorial versioning, and visual reading techniques in periodical studies that, together, formed the methodology of my dissertation. It also prodded me in the direction of scholarly and pedagogical practices that would come to be known as the digital humanities. Today I will focus on the project that bears David's most direct stamp, the Ecclesiastical Proust Archive, which began as an experiment in isolating textual features and grew into a larger project synthesizing editorial, archival, and analytic practices.² The project's examination of textuality might seem at odds with its more recent forays into text analysis and topic modeling. Its evolution is interesting in light of recent trends in digital humanities that attempt to

1. The following talk was delivered during a panel at David Greetham's retirement event at the Graduate Center of the City University of New York on April 11, 2014. A few alumni who had worked with David as graduate students were invited to speak about his influence on their work. I was pleased to return to my alma mater to see some old faces, to meet some new ones, and to be moved by the numerous stories that showed what a prolific and nurturing presence David had been to the program.
2. The archive can be accessed at <http://proustarchive.org>.

balance “big data” analysis and actual humanities interpretation. In what follows I will describe how the project came about as a way of interpreting Proust’s *À la recherche du temps perdu* and then, with David’s nudging, developed into an open-ended digital humanities project that continues to evolve today.

During my first year in our program, AY 2002–2003, I took Eve Sedgwick’s yearlong seminar on Proust and immediately set about writing a long meditative essay on his use of Gothic and Romanesque cathedrals. When I first sat down to read *Swann’s Way*, the philosophical depth of its meditations, the vibrant and palpable descriptions of people and places, the emotional textures of historical and personal memory, and their embodiment in architecture all converged in a manner that was overwhelming. For some reason, the passages having to do with churches were electric to me, and immediately I decided that my project for the course would perform an extended meditation on this rich and complex motif.

My first move was what can only be described as an archival drive to capture the church motif in its entirety. As part of my effort to articulate the wholeness of the motif, which radiated irresistibly from its points in the narrative, I began keeping a spreadsheet that documented every one of its occurrences (fig. 1). I recorded the pagination, transcribed the passage in its entirety, included a note on the narrative context, and added keywords that would help find passages to write about later on (say, by using the Find function). In short, as I soon learned upon reading (in David’s course) Marta Werner and Paul Voss’s introduction to the archival theory issue of *Studies in the Literary Imagination*, I was curating a collection of textual objects and providing an access mechanism that imbued them with an interpretive politics.³

When I ran these ideas by Eve, she simply looked at me and said, in her endearingly awkward way, “Churches?” After all, in an environment where queer theory, the new psychoanalysis, and script theory were all the rage, the subject of churches was hopelessly dowdy—but not too dowdy for David! I showed him the spreadsheet as we began reading archival theory. He immediately urged me to turn it into a database and to include other media, even suggesting a title for the project: *Ecclesiastical Proust*, which ended up becoming *Ecclesiastical Proust Archive*. Once Eve saw what I was up to, she was enthusiastically involved and formed with David a continually supportive mentorship on the theoretical and textual bearing of the project. I would therefore like to say a little bit about this project’s engage-

3. *Studies in the Literary Imagination* 23:1 (Spring 1999).

ment with textuality, since it forms one of the more prominent pieces in my repertoire—and, as a side project, greatly contributed to my “longevity” as a student in the program—all thanks to David’s early guidance!

The *Ecclesiastical Proust Archive* (fig. 2) is a database of text and images that enable a researcher to explore the church motif of *In Search of Lost Time*.⁴ The church motif forms one of the primary recurring elements and acts as the novel’s central metaphor for Lost Time. Churches constitute the orientation points of various settings and narrative phases. Proust at one point considered titling the novel *Le Cathédrale*, with the various parts named after architectural elements of a church. It is unsurprising, then, that Proust’s narrator concludes that books of this magnitude are never complete: “How many great cathedrals remain unfinished!” (VI.508). The narrator’s meditations upon such diverse topics as history, the subject/object distinction, jealousy, and the writing of books are frequently associated with “church”, and in myriad ways.

The project’s current instantiation as a database is adept at representing the complexities—and simplicities—of the church motif as extracted from its textual context. The five methods offered by the search page (fig. 3) include a text search (which allows boolean operators and wildcards), a dropdown menu of associations (loosely categorized tags that annotate the church passages), a dropdown menu of narrative context notes, a dropdown menu of image properties, and a pagination delimiter for working with a selected portion of the text. The search results (fig. 4) are displayed in a grid that from left to right displays the pagination info, the passage itself, the associations within the passage (if selected), a note on the passage’s narrative context (if selected), and an image illustrating the passage. The associations appear as links that allow the user to move through the archive by chains of association, as it were. Image captions can be viewed in tooltips on mouseover, and larger versions appear when clicked on. Though the search results are displayed in chronological order, they allow the reader to transcend large gaps in the narrative and behold at once a series of readings that are related, though separated in the original document by textual space.

As an example of the kinds of analysis that the database facilitates, we can examine an association search on the term Love Fantasy. The database returns two records containing the narrator’s memories of childhood love

4. In Eve’s course we read the 1992–1993 Enright revision of the Moncrieff/Kilmartin translation, published by Random House, which was the text used for the database. More recent activities on the project use an electronic version of the 1919 *Nouvelle Revue Française* edition in the original French.

fantasies. The first passage does not refer to any specific church, so the question of how to illustrate it involves a highly interpretive move.

But the interruption and the commentary which a visit from Swann once occasioned in the course of my reading, which had brought me to the work of an author quite new to me, Bergotte, resulted in the consequence that for a long time afterwards it was not against a wall gay with spikes of purple blossom, but against a wholly different background, the porch of a Gothic cathedral, that I saw the figure of one of the women of whom I dreamed. (I.124)

I decided to use my own photograph of the north porch, central portal of Chartres (fig. 5) because it displays one of the quintessential examples of French Gothic, bearing a correspondence with the architectural part and style described in the passage's love fantasy. This particular frame is dramatically overexposed, creating a ghostly effect that corresponds to the emotional tenor of the passage's paradoxically concrete yet hazy imagination of a future romance. However, after viewing it I notice more correspondences that have a larger significance. A tourist in the bottom right corner is looking up at the porch of the giant cathedral, mimicking the position of the narrator and, with him, the implied reader at this stage of the book: both are gazing at the mouth of a seemingly impenetrable archive inscribed with meanings in multiple media that will ultimately be revealed as a function of the love for a woman (fig. 6). Gilberte will be the first of these, while Albertine forms the basis of the narrator's pivotal crisis. Thus, the image of a Gothic porch prefaces and symbolizes the course of the narrative, with all of the concrete inscrutability of life's sensuous experience. This pairing shows the passage to be a kind of synecdoche for the whole novel. An association search on the term Love reveals similar results, with the idea of a future romantic affair taking place on the porch of an unnamed Gothic cathedral.

By now it should be clear that the project venerates the church motif as an icon, despite (or because of?) the contradiction inherent in its radical decontextualization of passages that originate in an organic text. This iconoclasm results precisely from the urge to apprehend and understand the entirety of that object as it manifests in different times and places in the narrative. The self-contradictory iconoclasm of the database text is what marks it as a hybrid between an edition and an archive. Moreover, my recent forays into topic modeling and network analysis have taken the iconoclasm of the archival drive even further.

Network analysis is a technology that takes structured data (in this case the spreadsheet from which the database was made), finds connections

among them, and then draws a graph with edges and nodes to show where the centers of influence reside. In a network graph, the temporal dimension of memory is eliminated: it flattens the chronology of the narrative and its interpretive metadata to make all connections simultaneously present. A visualization created with a program called Organization Risk Analyzer (ORA)⁵ shows the association of Venice as it is networked among church passages and narrative context notes amid the novel's entire network map (fig. 7). When manipulated in real time, the visualization highlights the links to other nodes and their related concepts or passages. What this means for the study of Proust is that we can think of the novel (and the novel *genre*) as a network of nodes consisting of concepts, characters, narrative elements, and any other unit of meaning that might enhance exploration of its text.

For instance, the network for the Time association (fig. 8) connects various types of recollection to provide insight into the narrator's artistic development. Here we find Time at the center, ringed by "Contemplation sparked by conversation with M. de Cambremer, at Guermites party", "Imagining Florence and Venice (before visit)", "Contemplating experience of Vinteuil's sonata while jealous of Mlle Vinteuil and Albertine", "Contemplating women and the past", "Observations at Guermites party", and "First visit to Balbec". The last in turn connects with Narthex and Carqueville, the site of a Romanesque church in Normandy that the narrator visits with Mme de Villeparisis and Albertine. In other words, Time as a backwards-looking concept is associated with jealousy over women, while the passages about Time as a forward-looking fantasy imagine the reddish domes of Florence and the frescos of Venice. This suggests a deepening of the structure that became apparent in the database searches above, where in early passages the thought of meeting a future lover, though not explicitly concerned with the nature of time, took place on the porch of an unidentified Gothic cathedral. These nodes presented by ORA show that the church passages consciously dealing with the nature of time occur *after* the narrator has experienced being in love with women. And correspondingly, the architectural element of this ring is the narthex, which is the entrance area just indoors or on the threshold to the porch. The narthex was not considered part of the church proper, but was placed close enough so that those deemed unworthy of entry, such as the unbaptized or unconfessed, could still receive instruction from services. Hence, the experience of love has brought the narrator past the porch but, because he is lost through jealousy, he still remains an outsider.

5. The program can be found at <http://www.casos.cs.cmu.edu/projects/ora/>.

Another practice that computes the statistical relationships among tokens is topic modeling, which clusters repeating patterns of single, double, or triple word phrases appearing within a specified span of text such as a paragraph or groups of, say, fifty words. Since the *Recherche* embodies more than one million words, topic modeling can be used to highlight features of the text that are not perceptible during the act of serial reading. I ran an electronic version of volume one of the first French edition, *Du côté de chez Swann*, through Mallet⁶ to show token clusters for ten topics, which reveals some interesting patterns. The command line output shows ten topics, each consisting of the top nineteen recurring words that are statistically significant within the top ten recurring patterns in the text (fig. 9).

Some of the results are unsurprising, such as topic 7, which clearly derives from the many evening scenes at the Verdurins (*soir, chez, maison*) where Swann courted Odette among their coterie (*forcheville, cottard*), often becoming jealously heartbroken (*cœur, désir*) with wondering whether she was seeing other admirers on the sly (*demander, connaissait, amis*). Other topics reveal interesting patterns that fit with scenes across the entire narrative, such as number 10. It emphasizes the use and observation of the eyes (*yeux, vue*) in connection with the Duc and Duchesse de Guermantes, whose mysterious airs and glances are described in the Combray church passage in the database section above, as well as their association with art and symbolism of France (*image, figure*). But what also emerges is the consistency of the preposition before (*devant*), emphasizing the narrator's location not only in front of their paintings and of their glances, but also in front of a church (*église*) in connection to a woman (*dame*), a recurrence that was teased out by reading the database passages from the English translation.

Using a PHP script and MySQL database,⁷ we can extract the tokens, word counts, and their connections from the Mallet topic model files into a graph file that generates edges and nodes, allowing us to view the ten topics as a network model in Gephi (fig. 10).⁸ This entirely computer-generated model of associative networks in *Du côté de chez Swann* is markedly different from the static model created by my particular reading of the church motif above, though it shares some consistencies and interesting disparities.

For instance, when we drill down and filter to look more closely at the terms that join the different topics (fig. 11), we see that the word for noth-

6. The program can be found at <http://mallet.cs.umass.edu/>.

7. The technique used here was supplied by Elijah Meeks, Digital Humanities Specialist at Stanford University.

8. <http://gephi.org>

ing (*rien*) is the one that most frequently connects topics 6 and 9, which respectively center on themes of beautiful bodily gestures in music and domestic relationships, while time (*temps*) joins topic 6 with 3, which is focused on positive terms for love of Gilberte. According to the statistical features of the text, then, the first two parts of *Du côté de chez Swann* associate the expression of romantic love primarily with time, while the memory of familial love is associated primarily with absence. This perhaps comes as no shock to most readers of Proust, but if we compare this model with a search for the term “nothing” in the church motif database, as before, we receive a number of passages associated predominantly with romantic love. These two datasets, then, suggest a reading of the church motif as concerned with concepts of absence in romantic love, somewhat against the grain of the rest of the novel. There is not enough time here to deal with the problematics of translation/tutor text comparisons or the relation of computational algorithms to critical interpretation. But it is clear that domain expertise is just as necessary in digital scholarship as it is in print, as shown by the (illuminating) disparities between a human-reading and machine-reading of the text.

Iconoclasm—a society’s destruction of its own established religious imagery, venerated institutions, or cherished beliefs now regarded as fallacious or superstitious—is an apt word to describe the breaking up of Proust’s studiously organic text (for many years he insisted it be published in a single volume). However, we might see the iconoclasm of “automagically” tokenizing the text through software as, paradoxically, an act of devotion, seeking to find the epiphany in the hidden details. When the text is considered as an archive, computational methods for analysis provide a capacious reading tool for making connections that have not already been made, and finding questions we hadn’t thought to ask. It is to David’s guidance that I owe this peculiarly fulfilling relationship with text.

The University of Tulsa

Works Cited

- PROUST, Marcel. 1919. *Du côté de chez Swann*. Paris: Éditions de la Nouvelle Revue Française.
- . 1992–1993. *In Search of Lost Time*. 6 vols., translated by D. J. ENRIGHT, Terence KILMARTIN, Andreas MAYOR, C. K. SCOTT MONCRIEFF. New York: Random House.
- Voss, Paul J. and Marta WERNER. 1999. “Toward a Poetics of the Archive: Introduction”. *Studies in the Literary Imagination* 32:1: i–vii.

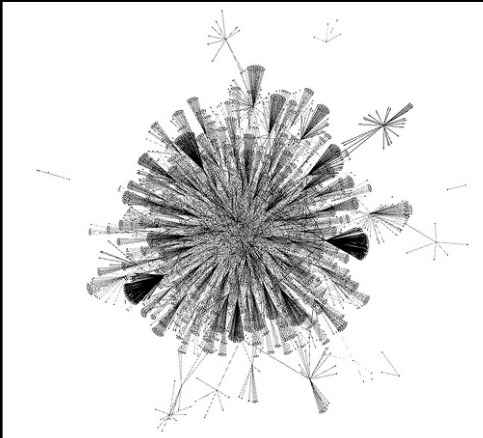
Entry	Vl/Pt/Ch/Pg	Passage	Narr. Context	Associations	Description
1	I/I/1.1	And half an hour later the thought that it was time to look for sleep would awaken me, I would make as if to put away the book which I imagined was still in my hands, and to blow out the light; I had gone on thinking, while I was asleep, about what I had just been reading, but these thoughts had taken a rather peculiar turn, it seemed to me that I myself was the immediate subject of my book: a church, a quarter, the rivalry between François I and Charles V. This impression would persist; for some moments after I awoke, it did not offend my reason, but lay like scales upon my eyes and prevented them from registering the fact that the candle was no longer burning. Then it would begin to seem unintelligible, as the thoughts of a previous existence must be after reincarnation; the subject of my book would separate itself from me, leaving me free to apply myself to it or not, and at the same time my sight would return and I would be astonished to find myself in a state of darkness, pleasant and restful enough for my eyes, but even more, perhaps, for my mind, to which it appeared incomprehensible, without a cause, something dark indeed.	Dreaming Reading	Sleep Imagination Subject/Object Kings Music Darkness Dreaming Reading Reason	Beginning of novel, dreams self as church
2	I/I/153-4	She attempted by a subterfuge, if not to eliminate altogether this commercial banality, at least to minimise it, to supplant it to a certain extent with what was still art, to introduce, as it were, several "thickesses" of art: instead of photographs of Chartres Cathedral, of the Fountains of Saint-Cloud, or of Vesuvius, she would inquire of Swann whether some great painter had not depicted them, and preferred to give me photographs of "Chartres Cathedral" after Corot, of the "Fountains of Saint-Cloud" after Hubert Robert, and of "Vesuvius" after Turner, which were a stage higher in the scale of art.	Bedtime reading	Art Layers of representation Simulacra Authenticity Grandmother Chartres	Grandmother's gifts, for relief of anxiety
3	I/I/164	And as in the game wherein the Japanese amuse themselves by filling a porcelain bowl with water and steeping in it little pieces of paper which until then are without character or form, but, the moment they become wet, stretch and twist and take on colour and distinctive shape, become flowers or houses or people, solid and recognisable, so in that moment all the flowers in our garden and in M. Swann's park, and the water-lilies on the <u>Yvonne</u> and the good folk of the village and their little dwellings and the parish church and the whole of <u>Combray</u> and its surroundings, taking shape and solidity, sprang into being, town and gardens alike, from my cup of tea.	Teatime memory	Locus of memory Combray Landscape Folk Combray	End of the Madeleine passage
4	I/I/165	<u>Combray</u> at a distance, from a twenty-mile radius, as we used to see it from the railway when we arrived there in the week before Easter, was no more than a church epitomising the town, representing it, speaking of it and for it to the horizon, and as one drew near, gathering close about its long, dark cloak, sheltering from the wind, on the open plain, as a shepherdess gathers her sheep, the woolly grey backs of its huddled houses, which the remains of its mediaeval ramparts enclosed, here and there, in an outline as scrupulously circular as that of a little town in a primitive painting.	Combray landscape	Landscape Representation Name/Naming Mediaeval Primitive Painting Rustic Combray	Beginning of Part I, Section II
5	I/I/171	It was Françoise, motionless and erect, framed in the small doorway of the corridor like the statue of a saint in its niche. When we had grown more accustomed to this religious darkness we could discern in her features the disinterested love of humanity, the tender respect for the gentry, which the hope of receiving New Year bounty intensified in the nobler regions of her heart.	Memory of Aunt Léonie	Statues Saints Darkness Love Humanity Reverence Aristocracy Niche	Transition of memory of Aunt Léonie to memory of Françoise

Figure 1. Spreadsheet of the church passages with interpretive keys and contextual information.

Ecclesiastical Proust Archive
A Textual / Visual Resource

Home Search Engine Image Galleries Discussion Forum

Gorgeous Network Visualizations with Gephi
I just started playing around with [Gephi](#), which produces really amazing interactive graphs for intuitive data analysis. Thought I would post one here. Not sure how to read or use it yet, but it shows the entire church motif in the *Recherche*.



Post Category: [visualization](#)
[Add comment](#) December 3rd, 2010 at 12:11am Jeff Drouin

A Little Close Reading with Network Analysis Software
I thought I would do some closer "distant-reading" of the *Recherche*. When using ORA to look at a metanetwork, the visualization can be manipulated in real time to highlight the links among nodes and their related concepts or passages. What this means for the study of Proust is that we can think of the novel (and the novel genre) as a network of nodes consisting of concepts, characters, narrative elements, and any other unit of meaning that might enhance exploration of its text.

about
rationale
links
bibliography
contributors
contact

recent posts:

Gorgeous Network Visualizations with Gephi

A Little Close Reading with Network Analysis Software

Tome I Word Cloud

Upcoming Lecture -- Digital Methods for Literary Criticism: Proust, Illustration, and the Archive

Proust LOLCat

Social Network Analysis of the *Recherche* (and The Novel)

An Old Essay Fragment

Project Bamboo Workshop 4 Summary

Graphs, Maps, Trees

Archive Theory: S/Z

post history:

Figure 2. Ecclesiastical Proust Archive homepage.

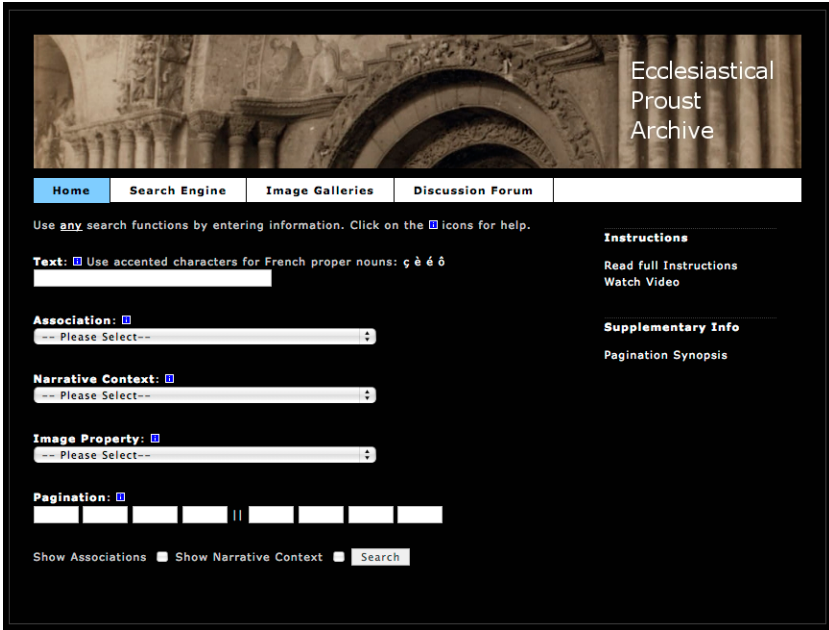


Figure 3. The search page.


Pagination	Passage	Associations	Context	Image
1 1 124-124	<p>But the interruption and the commentary which a visit from Swann once occasioned in the course of my reading, which had brought me to the work of an author quite new to me, Bergotte, resulted in the consequence that for a long time afterwards it was against a wall gay with spikes of purple blossom, but against a wholly different background, the porch of a Gothic cathedral, that I saw the figure of one of the women of whom I dreamed.</p>	<p>Bergotte Books Gothic Love fantasy Porch</p>	<p>Reading</p>	
1 1 250-250	<p>And remembering the glance which she had let fall upon me during mass, blue as a ray of sunlight that had penetrated Gilbert the Bad, I said to myself: "She must have taken notice of me." I fancied that I had found favour in her eyes, that she would continue to think of me after she had left the church, and would perhaps feel sad that evening, at Guermentes, because of me.</p>	<p>Aristocracy Chromatics Sombray Gaze Light Love fantasy Marriage Windows</p>	<p>Fantasy after first sighting of Mme de Guermentes; Mass</p>	

Figure 4. Search results for the association "Love Fantasy".

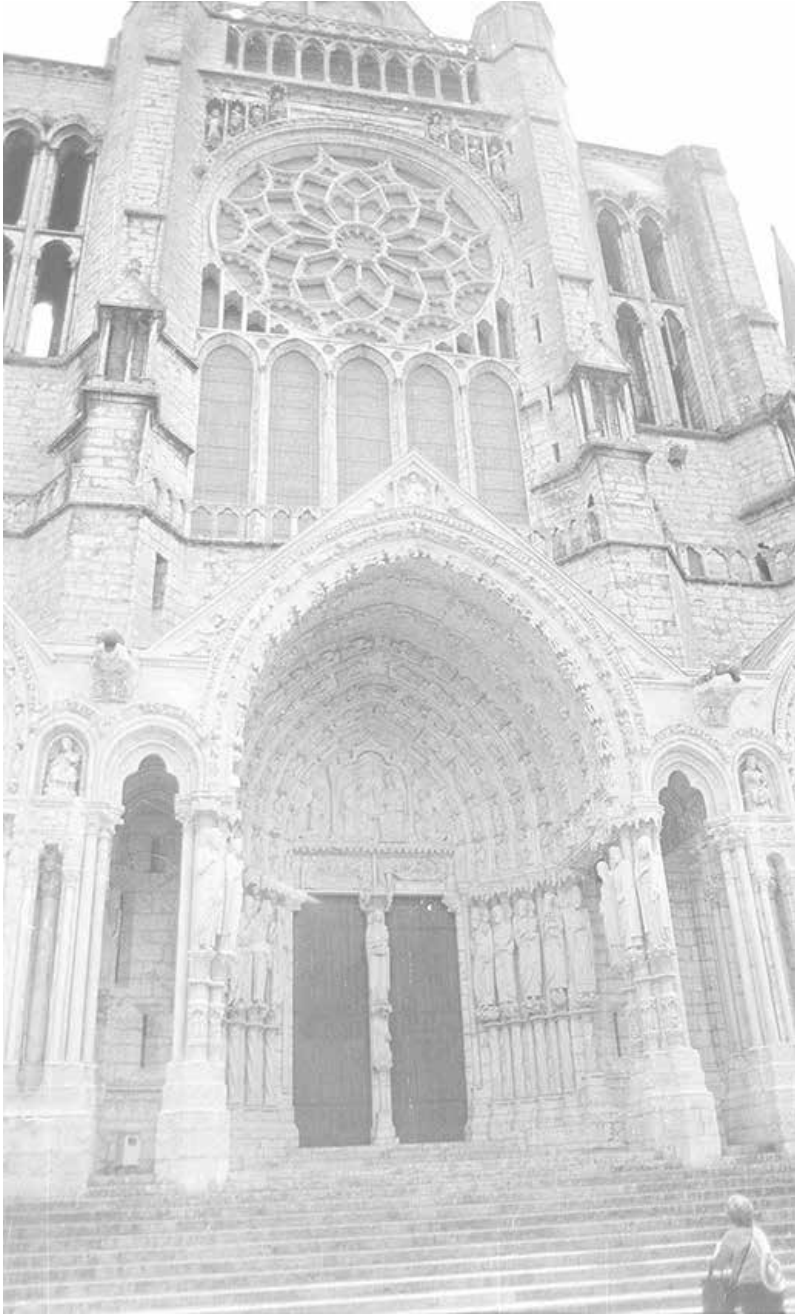


Figure 5. Chartres Cathedral, north porch. Photographer: Jeffrey Drouin.



Figure 6. Chartres jamb statues, west porch. Photographer: Jeffrey Drouin.

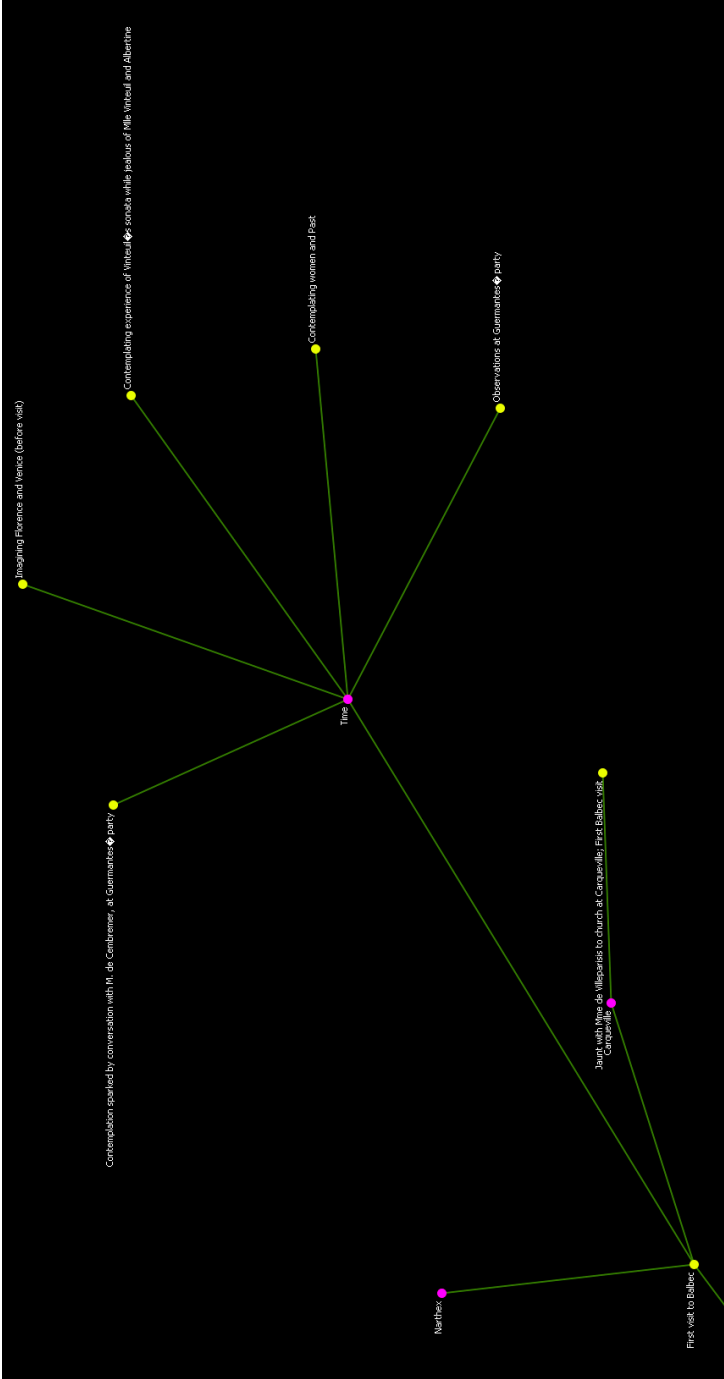


Figure 8. Time network, as visualized in ORA.

- ① chose moment pouvait jamais puis rien esprit pourtant visage savait voulait dire
savoir mal trouvait première devait autres instant
- ② dit bien dire air jamais beaucoup tête toujours princesse ami docteur reste choses.
sais enfin regard répondit jeune entendu
- ③ vie amour plaisir souvent celle ainsi gilberte pu pensée besoin donnait tant sorte
milieu cause femmes étais connaître joie
- ④ après temps jusqu'heure pendant allait presque chambre longtemps près seul
passer heures penser jour tard souvenir chercher toute
- ⑤ combrav côté déjà rue soleil semblait fleurs saint bois place eau ciel petits vers jardin
matin champs dessus autour
- ⑥ faisait toutes petite peine seule beau toute sourire donner phrase quelques trouver
parfois contraire nature suite musique croire corps
- ⑦ swann odette chez verdurin monde disait gens femme forcheville homme soir effet
amis connaissait demander personne cœur cottard
- ⑧ voir faire aller autres jours jour toujours maison venait venir désir grande contre dès
autant paris rien lequel bien
- ⑨ grand tante mère père françoise faire bien fille disait parents maman voix partie
personne bonne petit mort famille laisser
- ⑩ devant guermantes yeux nom air petit surtout or doute mieux église image fit vue
dame tant aussitôt figure lesquelles

Figure 9. Ten-topic model of *Du côté de chez Swann*, produced with Mallet.

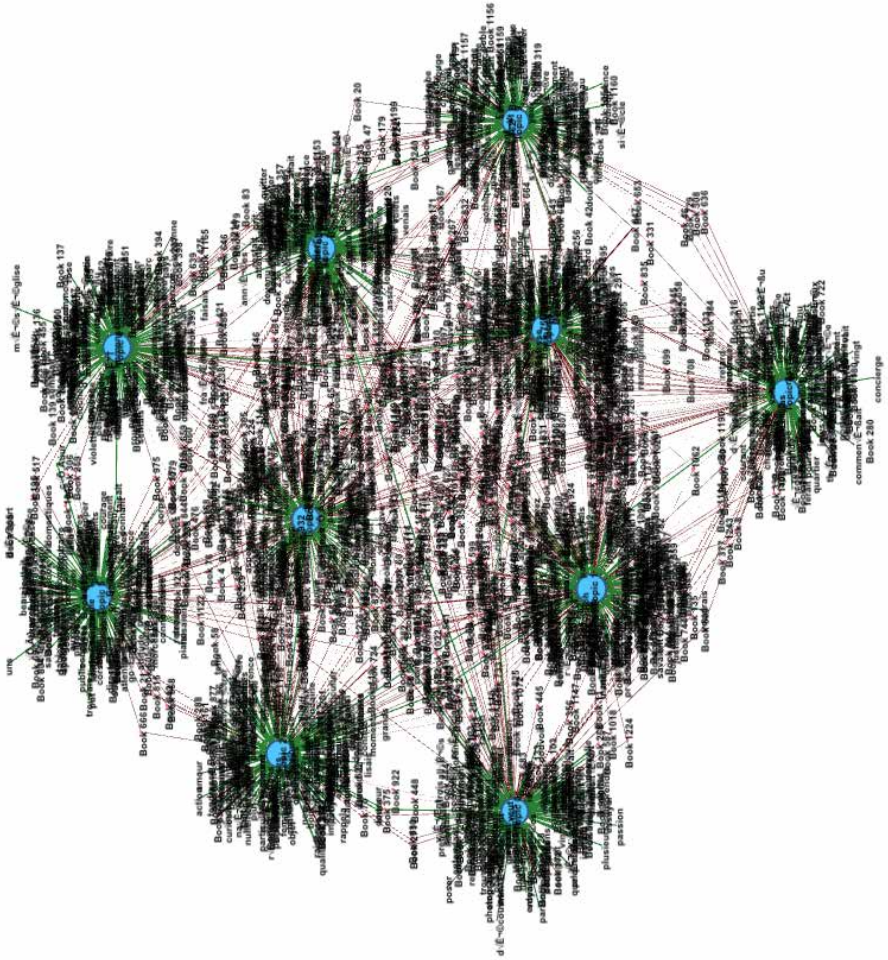


Figure 10. Network graph, made with Gephi, of a ten-topic model of *Du côté de chez Swann* produced with Mallet.

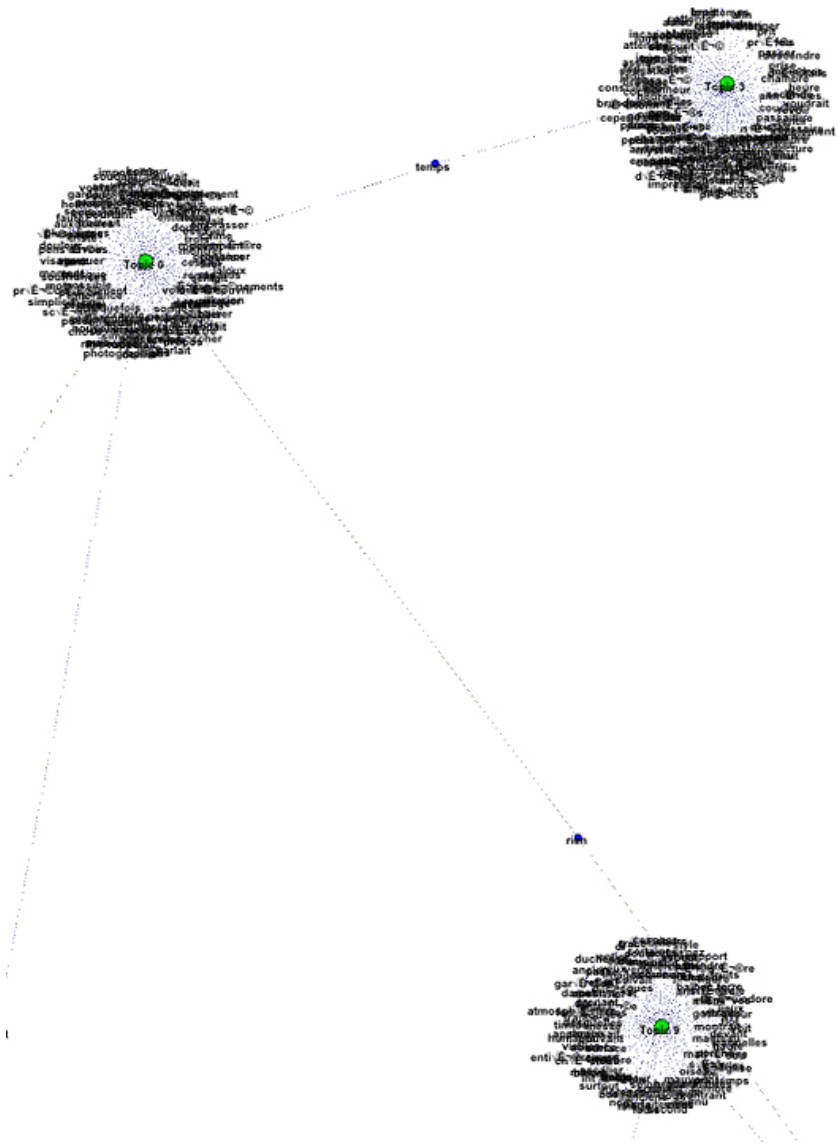


Figure 11. Close-up of topics 3, 6, and 9. The word *rien* connects 6 and 9, while *temps* connects 6 and 3.