



HAL
open science

Incertitude et histoire numérique

Michael Piotrowski

► **To cite this version:**

Michael Piotrowski. Incertitude et histoire numérique. Humanistica 2023, Association francophone des humanités numériques, Jun 2023, Genève, Suisse. hal-04108242

HAL Id: hal-04108242

<https://hal.science/hal-04108242>

Submitted on 26 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Incertitude et histoire numérique

Michael Piotrowski

Section des sciences du langage et de l'information

Université de Lausanne

michael.piotrowski@unil.ch

Résumé

La recherche historique est toujours confrontée à l'incertitude en raison de l'incomplétude de notre connaissance du passé. Il est pourtant important de distinguer deux types d'incertitude : l'incertitude historique qui concerne les faits du passé et l'incertitude historiographique qui concerne l'interprétation de ces faits. En histoire numérique, les deux types représentent des défis majeurs pour la construction de modèles computationnels, car les cadres de modélisation computationnels actuels ne sont pas bien adaptés à l'incertitude habituellement rencontrée dans la recherche historique. Nous argumentons que la gestion de l'incertitude historiographique, étroitement liée aux récits historiques et aux modèles causaux, est indispensable pour la construction de modèles historiographiques computationnels.

1 Incertitude historique et incertitude historiographique

Pour des raisons évidentes, la recherche historique est toujours confrontée à l'incertitude. Nos connaissances du passé sont toujours lacunaires. Les historiennes et les historiens s'en rendent compte. En étudiant les sources, ils et elles tiennent compte du fait que les mots et les choses ne signifient pas (ou n'ont pas signifié) nécessairement ce qu'ils semblent signifier à première vue, qu'il y a des inexactitudes, des erreurs, des fabrications et des contrefaçons. Même lorsqu'ils sont convaincus d'avoir identifié correctement une personne, un lieu ou la date d'un événement, ils savent qu'il n'y a jamais de certitude absolue.

Cette incertitude, que nous appelons l'*incertitude historique* et qui concerne les faits du passé, est inévitable. Elle donne lieu aux interprétations multiples du passé, à de différents récits. Il y a donc un deuxième niveau d'incertitude que nous appelons l'*incertitude historiographique* : quelle est la bonne interprétation ? Cependant, cette quête pour la vérité est en vain : sous la

condition que les différents récits en question soient tous *viabiles* (au sens de Glasersfeld 2005), ils sont irréductibles, ce que Ricœur (2017) a désigné le « conflit des interprétations ».

2 Incertitude et histoire numérique

Lorsque l'ordinateur n'est utilisé que comme un outil permettant de mécaniser un travail répétitif, l'incertitude n'est pas trop gênante. Par exemple, lorsqu'une base de données remplace les fiches portant des informations extraites des sources pour faciliter l'élaboration du récit historique, il n'y a pas de problème si ces données (qui, comme le souligne Torgerson (2022), proviennent des sources) sont stockées dans une base de données en tant que valeurs « exactes », pour autant que tous les utilisateurs soient conscients de ce fait.

L'histoire numérique n'est pourtant pas simplement l'utilisation d'ordinateurs pour mécaniser le travail répétitif dans la recherche historique. Piotrowski et Xanthos (2020) définissent les humanités numériques comme la construction et l'utilisation de modèles formels, plus précisément computationnels, pour étudier les questions de recherche en sciences humaines. Suivant leur distinction entre *humanités numériques appliquées* et *humanités numériques théoriques*, nous considérons donc l'histoire numérique comme un domaine des humanités numériques appliquées : la recherche historique utilisant des modèles computationnels.

Pour la construction de modèles computationnels, l'incertitude présente un enjeu majeur, indépendamment du type de modèle : le problème se pose lorsque les données sont traitées automatiquement, c'est-à-dire quand les éléments des données sont assignés des significations formelles et peuvent donc être interprétés par la machine. Si l'incertitude des données n'est pas prise en compte, le modèle computationnel est construit sur la base d'informations essentiellement fictives et aura donc une valeur douteuse.

Les humanités numériques sont de plus en plus conscientes de ce problème. Cependant, les cadres de modélisation computationnelle actuels ne sont pas bien adaptés pour gérer l'incertitude généralement rencontrée dans la recherche historique. Même si les généralisations sont possibles en histoire, les phénomènes historiques ne sont pas régis par les lois de la nature ; contrairement à l'incertitude dans la vie de tous les jours, la plupart, voire la totalité, du contexte d'origine fait défaut ; nous ne pouvons généralement pas demander des éclaircissements aux acteurs historiques, et même lorsque nous le pouvons, leurs points de vue sont très subjectifs. De plus, non seulement notre connaissance du passé est toujours fragmentaire, mais comme l'a souligné Thaller (2018), l'histoire s'attache particulièrement à « décoder ce que l'expéditeur n'a pas voulu transmettre », c'est-à-dire que les sources les plus pertinentes utilisées par les historiens n'étaient pas destinées à leur servir de messages.

Par conséquent, l'objectif est généralement de *réduire* l'incertitude historique autant que possible afin d'obtenir des données non ambiguës – comme des coordonnées géographiques – qui se prêtent à un traitement informatique. En d'autres termes, si, par exemple, on sait que les coordonnées sont « bonnes », on peut également être sûr que des coordonnées identiques se réfèrent en fait au même endroit. L'accent a été mis jusqu'à présent principalement sur l'identification des référents des expressions linguistiques – en particulier les entités nommées, les dates et les lieux – ou sur la représentation de différentes lectures de manuscrits.

Puisque l'élimination de l'incertitude n'est souvent pas possible, des travaux ont également été menés pour rendre l'incertitude explicite et lisible par les machines, afin qu'elle puisse être associée aux données – ce qui nécessite de *modéliser* l'incertitude (voir, par exemple, Piotrowski 2019 ; Edmond 2019) – ou pour rendre l'incertitude évidente dans les résultats du traitement informatique, par exemple, dans des visualisations (voir, par exemple, Kräutli et Boyd Davis 2013 ; Therón et al. 2018). La métrologie (la science des mesures) a développé des approches sophistiquées pour travailler avec l'incertitude de mesure et pour éviter ou au moins contrôler la propagation des erreurs, mais puisque l'histoire étudie des phénomènes d'un type différent, ces méthodes ne peuvent pas être transférées, et nous manquons de méthodes rigoureuses pour

traiter l'incertitude dans l'histoire numérique et les humanités numériques en général.

3 Positivisme numérique

Une autre approche est ce que l'on pourrait appeler le « positivisme (historique) numérique », qui vise à surmonter l'incertitude grâce aux *big data*. Le projet *Venice Time Machine* (Kaplan et Lenardo 2020) en est un excellent exemple, mais cette idée sous-tend implicitement une grande partie de la recherche en humanités numériques. Le principe du positivisme historique formulé par Langlois et Seignobos (1898) s'applique à la fois à ses formes du XIX^e et du XXI^e siècle :

On peut penser qu'un jour viendra où, grâce à l'organisation du travail, tous les documents auront été découverts, purifiés et mis en ordre, et tous les faits dont la trace n'a pas été effacée, établis. – Ce jour-là l'histoire sera constituée [...]. (Langlois et Seignobos 1898)

Il s'agit toutefois d'un grave malentendu : reconstruire le passé n'est pas la même chose qu'écrire l'histoire, même si la reconstruction était « complète ». L'analogie de Winks de « l'historien en tant que détective » (1969) est très appropriée : les historiens sont confrontés à une « scène de crime », l'état actuel ou passé du monde. Leur tâche consiste à construire un « cas », un récit qui explique comment, à travers quels événements, cet état du monde est apparu ; dans le récit, les traces sélectionnées deviennent des preuves. Tout récit est donc étayé par un modèle causal, mais qui n'est traditionnellement pas explicite.

4 Conclusion

Si nous considérons l'histoire numérique comme la construction de modèles computationnels pour étudier les questions de recherche historique, une tâche essentielle est de rendre explicites et de formaliser ces modèles causaux. Le but ici n'est pas de reconstruire numériquement le passé, mais plutôt de formaliser les modèles causaux des historiens, ou, comme Gardin (1991) l'a appelé, *la formalisation du discours savant*.

C'est ici que la distinction entre les deux types d'incertitude que nous avons introduits ci-dessus devient cruciale : l'incertitude *historique* concerne les faits du passé ; l'incertitude *historiographique* concerne la construction des modèles causaux qui relient ces faits entre eux.

Nous soutenons qu'en conséquence, si nous voulons construire des modèles historiographiques computationnels, la voie à suivre n'est pas simplement d'accepter l'incertitude en histoire comme un mal inévitable, causé par le fait malheureux que notre connaissance du passé est incomplète. Bien qu'elle joue un rôle, l'incertitude historiographique n'est pas simplement un effet secondaire de l'incertitude historique : oui, l'incertitude est inévitable, mais c'est cette incertitude qui *permet* l'écriture de l'histoire, et en conséquence nous devons la traiter comme partie intégrale de celle-là.

Aujourd'hui, il nous manque encore les moyens pour accomplir cette tâche. Cependant, il nous semble qu'elle ne peut pas être accomplie ni par la quantification ni par le *big data*. Comme Gilles-Gaston Granger a remarqué, la pensée formelle à l'œuvre dans les sciences humaines ne doit pas se limiter à « la réduction des phénomènes aux calculs, mais aussi comme invention de structures nouvelles, voire même d'une mathématique originale » (Granger 1967).

Bibliographie

- Jennifer Edmond. 2019. [Strategies and Recommendations for the Management of Uncertainty in Research Tools and Environments for Digital History](#). *Informatics*, 6(3).
- Jean-Claude Gardin. 1991. *Le calcul et la raison : essais sur la formalisation du discours savant*. Éd. de l'École des hautes études en sciences sociales, Paris.
- Ernst von Glasersfeld. 2005. Konstruktion der Wirklichkeit und des Begriffs der Objektivität. In Heinz Gumin et Heinrich Meier, éditeurs, *Einführung in den Konstruktivismus*, Veröffentlichungen der Carl Friedrich von Siemens Stiftung, pages 9-39. Piper, München.
- Gilles-Gaston Granger. 1967. *Pensée formelle et sciences de l'homme*. Aubier-Montaigne, Paris, Nouvelle éd. augmentée d'une préface.
- Frédéric Kaplan et Isabella di Lenardo. 2020. Building a Mirror World for Venice. In *The Aura in the Age of Digital Materiality : Rethinking Preservation in the Shadow of an Uncertain Future*, pages 197-201. SilvanaEditoriale, Milan.
- Florian Kräutli et Stephen Boyd Davis. 2013. [Known Unknowns: representing uncertainty in historical time](#). In *Proceedings of Electronic Visualisation and the Arts 2013*, pages 61-68, London. British Computer Society.
- Charles-Victor Langlois et Charles Seignobos. 1898. *Introduction aux études historiques*. Hachette, Paris.
- Michael Piotrowski. 2019. [Accepting and Modeling Uncertainty](#). *Zeitschrift für digitale Geisteswissenschaften*.
- Michael Piotrowski et Aris Xanthos. 2020. [Décomposer les humanités numériques](#). *Humanités numériques*(1).
- Paul Ricoeur. 2017. *Le conflit des interprétations : Essais d'herméneutique*. Seuil, Paris.
- Manfred Thaller. 2018. [Decoding what the sender did not want to transmit. Information technology and historical data; or something](#). In Michael Piotrowski, éditeur, *Proceedings of the Workshop on Computational Methods in the Humanities (COMHUM 2018)*, Lausanne. CEUR Workshop Proceedings.
- Roberto Therón, Antonio G. Losada, Alejandro Benito, et Rodrigo Santamaría. 2018. [Toward supporting decision-making under uncertainty in digital humanities with progressive visualization](#). In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'18)*, pages 826-832, Salamanca, Spain. ACM.
- Jesse W. Torgerson. 2022. [Historical Practice in the Era of Digital History](#). *History and Theory*, 61(4) :37-63.
- Robin W. Winks, éditeur. 1969. *The Historian as Detective : Essays on Evidence*. Harper & Row, New York, NY.