



Genesis

Manuscrits – Recherche – Invention

30 | 2010

Théorie : état des lieux

Qu'est-ce qu'une édition génétique numérique ?

Paolo D'Iorio



Édition électronique

URL : <http://journals.openedition.org/genesis/116>

DOI : 10.4000/genesis.116

ISSN : 2268-1590

Éditeur :

Presses universitaires de Paris Sorbonne (PUPS), Société internationale de génétique artistique littéraire et scientifique (SIGALES)

Édition imprimée

Date de publication : 20 juin 2010

Pagination : 49-53

ISBN : 978-2-84050-697-3

ISSN : 1167-5101

Référence électronique

Paolo D'Iorio, « Qu'est-ce qu'une édition génétique numérique ? », *Genesis* [En ligne], 30 | 2010, mis en ligne le 30 mai 2012, consulté le 07 décembre 2020. URL : <http://journals.openedition.org/genesis/116> ; DOI : <https://doi.org/10.4000/genesis.116>

Tous droits réservés

Qu'est-ce qu'une édition génétique numérique ?

Paolo D'Iorio

L'ambition d'une édition génétique est de publier l'œuvre d'un auteur de manière à représenter et à rendre intelligible la genèse de ses projets d'écriture ou de création artistique. La plupart des spécialistes s'accordent à reconnaître que seuls les supports et les technologies numériques peuvent permettre la réalisation d'une véritable édition génétique et garantir sa diffusion auprès du public¹. En effet, la publication sur papier est souvent frustrante ou simplement impossible quand il s'agit d'éditer des dossiers génétiques concernant les arts de la performance : œuvres musicales, théâtrales, ou cinématographiques. Ce court article propose une mise au point théorique de la notion d'édition génétique numérique en décrivant ses cinq parties constitutives.

1. Dossiers génétiques

Les éditions critiques traditionnelles publient la production d'un auteur en la divisant selon la typologie des documents : d'un côté les œuvres, de l'autre côté les écrits posthumes, ensuite la correspondance (divisée souvent entre lettres envoyées et reçues), éventuellement des témoignages et des documents biographiques, et ainsi de suite. L'édition génétique, en revanche, rassemble et organise ces mêmes documents suivant la notion de dossier génétique. Un dossier génétique réunit tous les documents écrits, visuels ou audiovisuels concernant un projet d'écriture ou de création artistique abouti ou inachevé. Par exemple, le dossier génétique d'un ouvrage littéraire sera formé par tous ses manuscrits préparatoires, auxquels s'ajouteront les livres et les matériaux iconographiques que l'auteur a consultés ou annotés pendant la phase de rédaction, ainsi que les lettres comportant des instructions

spécifiques à l'éditeur, les documents biographiques comme des contrats ou des factures donnant des témoignages sur l'avancement du travail d'écriture et puis, finalement la ou les éditions de l'œuvre publiées sous la responsabilité de l'auteur lui-même (sans oublier d'éventuelles copies avec corrections autographes). L'édition génétique complète est ainsi formée par un ensemble de dossiers génétiques concernant aussi bien les ouvrages publiés que les projets inédits ou inachevés. L'architecture d'une édition génétique numérique doit être conçue pour pouvoir accueillir l'ensemble des archives d'un auteur. Si l'architecture est bien pensée, la réalisation pratique pourra ensuite être distribuée dans le temps et les différents dossiers génétiques abordés un par un ou confiés à plusieurs groupes de recherche, comme cela est souvent le cas pour les éditions critiques des œuvres complètes de quelque ampleur. La souplesse de l'édition numérique permet d'ajouter en cours de route de nouveaux matériaux jusqu'alors inconnus et surtout d'assurer la circulation des informations et des renvois à l'intérieur de ces *opera omnia*.

La constitution d'un dossier génétique est une tâche qui fait appel à l'intelligence et à la finesse du chercheur, et qui s'appuie sur le classement génétique décrit au paragraphe 4. À cette phase, l'informatique ne peut contribuer qu'en rendant plus aisés le stockage et la

1. Sur la notion d'édition génétique, voir au moins : Jean-Louis Lebrave, « Hypertextes – Mémoires – Écriture », *Genesis*, n° 5, 1994, p. 9-24 ; Almuth Grésillon, *Éléments de critique génétique*, Paris, PUF, 1994, p. 177-202 ; Hans Zeller et Gunter Martens (dir.), *Textgenetische Edition*, Tübingen, Niemeyer, 1998 ; Pierre-Marc de Biasi, *La Génétique des textes*, Paris, Nathan, 2000, p. 69-83 ; Louis Hay, *La Littérature des écrivains*, Paris, José Corti, 2002, p. 369-392 ; Alfredo Stussi, *Introduzione agli studi di filologia italiana*, Bologna, Il Mulino, 2007, p. 147-248.

manipulation des reproductions des matériaux². Dans la phase de publication, en revanche, les technologies numériques sont d'un apport décisif, permettant de marquer l'ensemble des documents de manière à spécifier leur appartenance à un ou, comme c'est souvent le cas, à plusieurs dossiers génétiques. Selon le même procédé, les matériaux qui n'ont pas (encore) été inclus dans un dossier génétique pourront être visualisés selon des critères de foliotage traditionnels, car l'édition génétique numérique comprend en son sein un catalogage complet de l'archive et son édition en fac-similé.

2. Catalogage et édition en fac-similé

Le support numérique permet de reproduire tous les documents qui composent les archives d'un auteur – manuscrits, imprimés, matériaux audiovisuels – et d'en donner un catalogage et une description matérielle unitaires et cohérents. Cela permet de reconstituer l'unité de ces documents, le plus souvent dispersés dans différents lieux de conservation, et de former une véritable édition en fac-similé qui est partie constitutive – nécessaire bien que non suffisante – de toute édition génétique numérique. Un tableau de concordance entre le nouveau catalogage et les catalogages précédents doit assurer la compatibilité avec les éditions et les études existantes et permettre aux chercheurs de retrouver rapidement les originaux dans leurs lieux de conservation. De même, avoir l'espace et la possibilité de publier également les descriptions matérielles précédentes, peut se révéler utile car les documents originaux peuvent s'être altérés dans le temps. S'appuyant sur une telle édition en fac-similé, il est possible de procéder à un classement génétique et d'esquisser une représentation des processus de création, qui sont les composantes spécifiques d'une édition génétique. Mais avant d'aborder ces deux points, disons quelques mots sur la question de la transcription.

3. Transcriptions

D'un point de vue théorique, la transcription n'est pas un des éléments constitutifs de la notion d'édition génétique. En cela l'édition génétique diffère

profondément de l'édition critique. En effet, l'édition critique d'auteurs anciens se propose de reconstruire un texte dont l'original est perdu à partir d'un ensemble de copies contenant des versions discordantes ; dans le cas d'auteurs modernes, elle vise à établir le texte que l'auteur entendait publier en l'expurgeant des fautes intervenues lors de l'impression. Dans les deux cas, la reconstruction d'un original perdu et l'établissement d'un texte imprimé de manière imparfaite ne peuvent être réalisés que par la production d'un nouveau texte. Le produit d'une édition génétique, en revanche, n'est pas un nouveau texte, mais une représentation, une mise en rapport et une explication des documents de genèse existants. L'édition génétique peut donc légitimement se contenter, d'un point de vue théorique, de reproduire en fac-similé le dossier génétique complet, en l'accompagnant d'un classement génétique des documents et de leurs sous-parties (zones des manuscrits ou des imprimés, séquences des documents audiovisuels, détails des œuvres d'art plastiques, etc.) et d'une représentation et explication des processus génétiques sous-jacents. Certes, une édition génétique où les documents de genèse sont uniquement disponibles en fac-similé pourrait être difficilement lisible par les non spécialistes. Mais cela dépend essentiellement du degré de lisibilité de l'écriture de l'auteur, et il ne faut pas oublier qu'à côté des dossiers manuscrits du XIX^e siècle il existe des dossiers génétiques plus récents, souvent presque entièrement dactylographiés ou écrits à l'ordinateur où le problème de la lisibilité va se poser différemment, pour ne pas parler des dossiers de genèse filmique, musicale ou artistique.

Cela dit, nous considérons qu'il est très souhaitable de doter une édition génétique de transcriptions pour les trois raisons suivantes : 1) dans les cas des manuscrits d'écrivains, souvent mal calligraphiés, la transcription diplomatique ou linéarisée est une « stratégie

2. De même qu'en philologie classique l'informatique, après des années de tentatives, s'est finalement révélée de faible recours concernant la formation des stemmas ou pour soutenir les conjectures nécessaires à la reconstruction du texte (voir A. Stussi, *Introduzione, op. cit.*, p. 245-246) ; de même nous ne pouvons pas nous attendre à un apport significatif de l'informatique pour la reconstruction de la genèse des œuvres modernes.

de facilitation³ », un instrument qui permet un accès plus confortable au texte, même s'il ne dispense pas du recours au fac-similé pour appréhender les éléments graphiques (tracé, disposition de l'écriture sur la page, dessins, etc.) qui peuvent contribuer à expliquer la genèse ; 2) l'existence d'une transcription linéarisée permet d'exploiter toutes les possibilités de recherche textuelle, depuis la simple recherche de mots jusqu'aux recherches sémantiques de type linguistique, génétique ou autre, surtout dans le cas où le texte transcrit a été balisé de manière appropriée⁴ ; 3) d'un point de vue théorique il me semble souhaitable que, en plus du fac-similé et d'une transcription diplomatique, l'édition génétique offre également non seulement une transcription linéarisée, mais un véritable texte critique. En effet, après avoir accompli un travail considérable d'analyse génétique, on dispose de tous les éléments requis pour l'établissement d'un texte critique et il serait dommage de s'en priver, de s'arrêter en laissant cette tâche à une autre édition, critique, à venir. En outre, de cette sorte l'édition génétique comprendrait en son sein non seulement une édition en fac-similé dont nous avons parlé au paragraphe 2, mais également une édition critique et pourrait ainsi se présenter comme la forme éditoriale la plus complète et aboutie pour la publication de l'œuvre d'un auteur.

Que transcrire ? Pour éviter des « répétitions » et économiser du papier, les éditions imprimées ne publient habituellement que le dernier état du texte et reproduisent dans l'appareil critique les variantes qui se trouvent dans les brouillons. Avec un peu d'efforts, suivant scrupuleusement les indications souvent codées de l'appareil critique, le lecteur devrait en principe être capable de reconstruire le contenu de toutes les versions préparatoires dans leur intégralité. Grâce au support numérique, il est maintenant possible de faire l'économie de cette gymnastique mentale qui souvent désespère les lecteurs et rend illisibles les éditions. L'édition génétique numérique peut et doit transcrire intégralement tous les documents qui composent son dossier génétique. Des collations entre les différentes versions, voire de véritables appareils de variantes, pourront être générés automatiquement, au cas par cas, par des logiciels appropriés⁵. Mais il ne s'agit pas seulement d'une question pratique. À la limite, le changement même d'un

seul mot peut produire une réécriture et un glissement de sens très important. Décider si le nouveau mot réalise la version finale d'un texte ou représente le commencement d'une phase rédactionnelle complètement nouvelle où ce terme va jouer un rôle central et va transformer profondément le contenu, le style ou la trame de l'écrit, relève d'une certaine analyse génétique. Mais étant donné que plusieurs analyses et parcours génétiques peuvent intéresser simultanément le même segment, il n'est pas possible, en génétique, d'assumer une version comme texte de base absolu à partir duquel définir les variantes. Pour être en mesure de reconstruire très précisément tous les glissements de sens dans leur contexte et de construire des séquences de segments textuels qui correspondent à différents processus génétiques, le généticien ne saurait se satisfaire d'un appareil de variantes : il a besoin de disposer de la transcription intégrale de tous les documents. De surcroît : il doit disposer d'une transcription intégrale non seulement de toutes les pages, mais également de toutes les strates de réécriture qui se trouvent sur la même page, pour autant qu'il soit possible de les cerner. Pour le dire avec les mots de Cesare Segre : « À la rigueur, on pourrait affirmer que dans un texte avec des corrections nous sommes également en présence, en termes linguistiques, de textes successifs superposés dans le même espace et définissables, par abstraction, comme des couches successives⁶. »

Comment transcrire ? De même que, du point de vue du contenu, les transcriptions d'une édition génétique doivent reproduire tous les segments textuels et toutes leurs réécritures explicites et implicites, de même, d'un point de vue graphique, elles devraient s'efforcer à la

3. Almut Grésillon, *Éléments de critique génétique*, op. cit., p. 129.

4. Cela vaut également pour le cas des genèses audiovisuelles, où la production d'une transcription de ce qui a été prononcé ou joué par les artistes facilite la détection d'éventuelles différences entre le script, la partition, le scénario et les performances effectivement réalisées.

5. Par exemple TUSTEP, MEDITE ou Juxta.

6. Cesare Segre, « Critiques des variantes et critique génétique », *Genesis*, n° 7 (1995), p. 37, qui développe une indication de Gianfranco Contini publiée dans « La critica degli scartafacci », dans *Rassegna d'Italia*, 1948, p. 1048-1056. Voir également l'intéressante discussion de ce concept et de ses implications éditoriales dans Alfredo Stussi, *Introduzione*, op. cit., p. 158-160 et 162-163.

fois de reproduire au plus près toutes les caractéristiques de la page manuscrite et de rester aisément lisibles. Ce sont naturellement des exigences contradictoires qui ne peuvent être satisfaites qu'en publiant, pour chaque unité textuelle, plusieurs transcriptions de type différent en fonction du contenu et de l'aspect graphique de la page ; ce qui peut se faire plus aisément sur support numérique. Les transcriptions peuvent se regrouper en trois grandes familles : 1) *les transcriptions linéarisées* se conforment au format typographique standard dans la publication des livres et renoncent à donner une représentation de la disposition des signes graphiques sur la page. Elles sont notamment utiles pour pouvoir effectuer des recherches textuelles dans le corpus ; 2) *les transcriptions diplomatiques*, en revanche, s'efforcent de suivre fidèlement l'aspect de la page manuscrite. Cette fidélité est plus ou moins prononcée en fonction des éléments graphiques qu'on entend reproduire : taille des caractères et leur position sur la page, type et couleur des encres, orientation de l'écriture, etc. Selon la manière de reproduire ces éléments on peut diviser les transcriptions diplomatiques en mimétiques, qui essaient de les reproduire graphiquement à l'identique, ou symboliques, qui les représentent avec des signes diacritiques ou d'autres conventions ; 3) *les transcriptions ultradiplomatiques* se situent à la limite entre le fac-similé et la transcription, car elles reproduisent en fac-similé la page manuscrite en substituant les signes de l'écriture avec des caractères typographiques. Les transcriptions interactives sont un type particulier de transcriptions ultradiplomatiques, qui laissent apparaître les caractères d'imprimerie à la demande du lecteur. Pour produire plusieurs transcriptions de la même page, par exemple une version linéarisée et une version diplomatique pour chaque strate d'écriture, il est conseillé de réaliser une seule transcription et de l'encoder avec un langage d'encodage adapté, comme le Langage d'encodage génétique⁷. En revanche, la conception d'un langage d'encodage capable d'automatiser la production des transcriptions ultradiplomatiques apparaît une tâche beaucoup plus complexe en raison du grand nombre de variables qui rentre en jeu ; cela reste un travail à réaliser par un graphiste avec des logiciels de dessin vectoriel.

4. Classements

Comme nous l'avons vu plus haut, une première forme de classement d'une archive est le catalogage, qui vise à repérer et décrire les documents dans leur matérialité en les ordonnant par types et en les numérotant selon l'ordre normal de lecture. Les éditions en fac-similé suivent habituellement cette forme d'ordonnement. Un classement strictement chronologique, en revanche, devrait faire abstraction de la typologie des documents et disposer chaque segment sur une ligne temporelle selon sa date de création, pour autant qu'il soit possible de l'établir avec précision. C'est-à-dire que si on peut démontrer qu'un écrivain a d'abord écrit une note dans un cahier, ensuite une lettre à un de ses correspondants et finalement a corrigé un passage dans un cahier appartenant à un autre projet d'écriture, ces trois rédactions doivent

7. Le « Langage d'encodage génétique » (LEG, anciennement « HyperNietzsche Markup Language », HNML) est une déclinaison du langage XML conçue pour encoder des manuscrits. Il permet d'encoder les caractéristiques matérielles de l'écrit, comme la couleur, le type d'instrument d'écriture utilisé, le type d'alphabet, ainsi que les processus génétiques de l'écriture, comme l'ajout, la rature, la surcharge. Il dispose en outre d'un ensemble de balises pour marquer les interventions de l'éditeur scientifique dans le cas d'une édition critique des manuscrits, comme la résolution des abréviations, les corrections orthographiques, les commentaires. Les caractéristiques principales du LEG sont : la simplicité, qui permet de maîtriser l'extrême complexité des phénomènes génétiques, où chaque mot ou même chaque lettre est susceptible d'être encodé ; la capacité de gérer des structures imbriquées, ce qui est indispensable pour encoder les procédés de révision présents dans les manuscrits des écrivains quand, par exemple, un mot souligné a été ensuite remplacé par un autre mot non souligné et écrit avec une encre différente ; et la possibilité d'encoder les strates d'écritures, c'est-à-dire de marquer un ensemble de modifications génétiquement liées qui appartiennent à la même phase de révision ou à la même campagne d'écriture. Le LEG a été conçu par Paolo D'Iorio et par son équipe en 2003 dans le cadre du projet HyperNietzsche (voir Harald Saller : « HNML – HyperNietzsche Markup Language », dans *Jahrbuch für Computerphilologie*, 5, Paderborn, Mentis Verlag, 2003, p. 185-192). À partir du LEG, un groupe de travail de la *Text Encoding Initiative* est en train d'écrire un nouveau chapitre de ses *Principes directeurs pour l'encodage et l'échange de textes électroniques* consacré aux éditions génétiques. Parmi les événements liés à ce travail en cours, citons l'atelier « Genetic Editions in a Digital Framework » organisé à l'ITEM en mai 2009 par P. D'Iorio (CNRS), F. Jannidis (univ. Würzburg), Elena Pierazzo (King's College London), Malte Rehbein (univ. Würzburg).

figurer dans la même séquence chronologique tout en n'étant liées ni typologiquement ni génétiquement. Les éditions critiques disposent habituellement les textes par ordre chronologique, mais à l'intérieur des différentes typologies de documents. Le classement génétique quant à lui entend établir, pour chaque segment textuel, la séquence de toutes ses réécritures jusqu'à la version publiée, si elle existe. Les séquences des réécritures sont normalement appelées « parcours génétiques » et le classement génétique coïncide donc avec l'ensemble des parcours génétiques.

Grâce à la souplesse du support, l'édition génétique numérique fait coexister différents systèmes de classement en redéployant les documents selon les intérêts du lecteur qui peut choisir de feuilleter les textes selon la succession des pages ou dans l'ordre chronologique, ou alors de suivre les parcours génétiques. En effet, chaque unité textuelle y est systématiquement liée à ses états génétiques précédents et suivants et le lecteur peut accéder à tout moment au parcours génétique complet, feuilleter le fac-similé et la transcription qui sont liés à chaque étape et suivre l'évolution de l'écriture et le développement de la pensée.

5. Représentation des processus génétiques

En traitant du classement, nous sommes arrivés tout près de l'élément le plus caractéristique d'une édition génétique, qui correspond au moment où, en utilisant tous les éléments précédemment cités, le fac-similé, la transcription, le classement, elle s'efforce de représenter la genèse de l'œuvre. Les parcours, en effet, tout en faisant partie du classement génétique, constituent aussi une première manière d'appréhender les processus de création. Ils peuvent être complétés par des tableaux ou des diagrammes qui en rassemblent les données concernant plusieurs parcours, donnent une vision d'ensemble des mouvements à l'œuvre dans un dossier génétique.

Un diagramme génétique peut nous montrer, par exemple, quelles pages de quels manuscrits contribuent à l'écriture d'un segment textuel particulier, d'un chapitre, voire d'un livre entier⁸. La visualisation des modifications textuelles est un deuxième moyen de saisir la genèse. Si le texte a été opportunément balisé et si les programmes de collation sont performants, il est possible d'en tirer, au-delà de la visualisation des simples modifications textuelles, des leçons plus générales sur les opérations génétiques comme l'ajout, le déplacement, etc.⁹. Malgré leur efficacité indéniable, les parcours, les diagrammes et la collation sont loin de nous faire percevoir les raisons qui ont produit le passage d'une version à l'autre. Certes, nous n'avons pas encore exploité toutes les possibilités de simulation offertes par les technologies disponibles et il est toujours utile d'expérimenter dans cette direction. Il nous semble toutefois que dans ce domaine on ne peut pas faire l'économie de la prose du généticien et que l'histoire de la production de l'œuvre peut finalement être plus aisément racontée que montrée. D'ailleurs, dans cette cinquième partie, l'éditeur doit désormais se préparer à laisser la parole à l'interprète, le récit de la naissance du texte se situant au milieu du chemin entre l'édition génétique et la critique génétique¹⁰.

8. Voir Pierre-Marc de Biasi, *La Génétique des textes*, op. cit., p. 80 ; Inga Gerike, « Les manuscrits et les chemins génétiques du *Voyageur et son ombre* », dans P. D'Iorio, *HyperNietzsche*, Paris, PUF, 2000, p. 159 ; Daniel Ferrer, « La représentation hypertextuelle des manuscrits : quelques leçons de douze années d'expériences », dans Aurèle Crasson (dir.), *L'Édition du manuscrit. De l'archive de création au scriptorium électronique*, Louvain-la-Neuve, Academia-Bruylant, 2008, p. 189-208.

9. La collation automatisée des caractéristiques graphiques des documents, en revanche, semble demander un langage formel de description extrêmement complexe à réaliser, voir Jean-Louis Lebrave, « Du visible au lisible : comment représenter la genèse ? », *Genesis*, n° 27, 2006, p. 11-18.

10. Voir Pierre-Marc de Biasi, « Les six grandes étapes de la recherche en génétique des textes », dans Aurèle Crasson (dir.), *L'Édition du manuscrit*, op. cit., p. 40 sq.