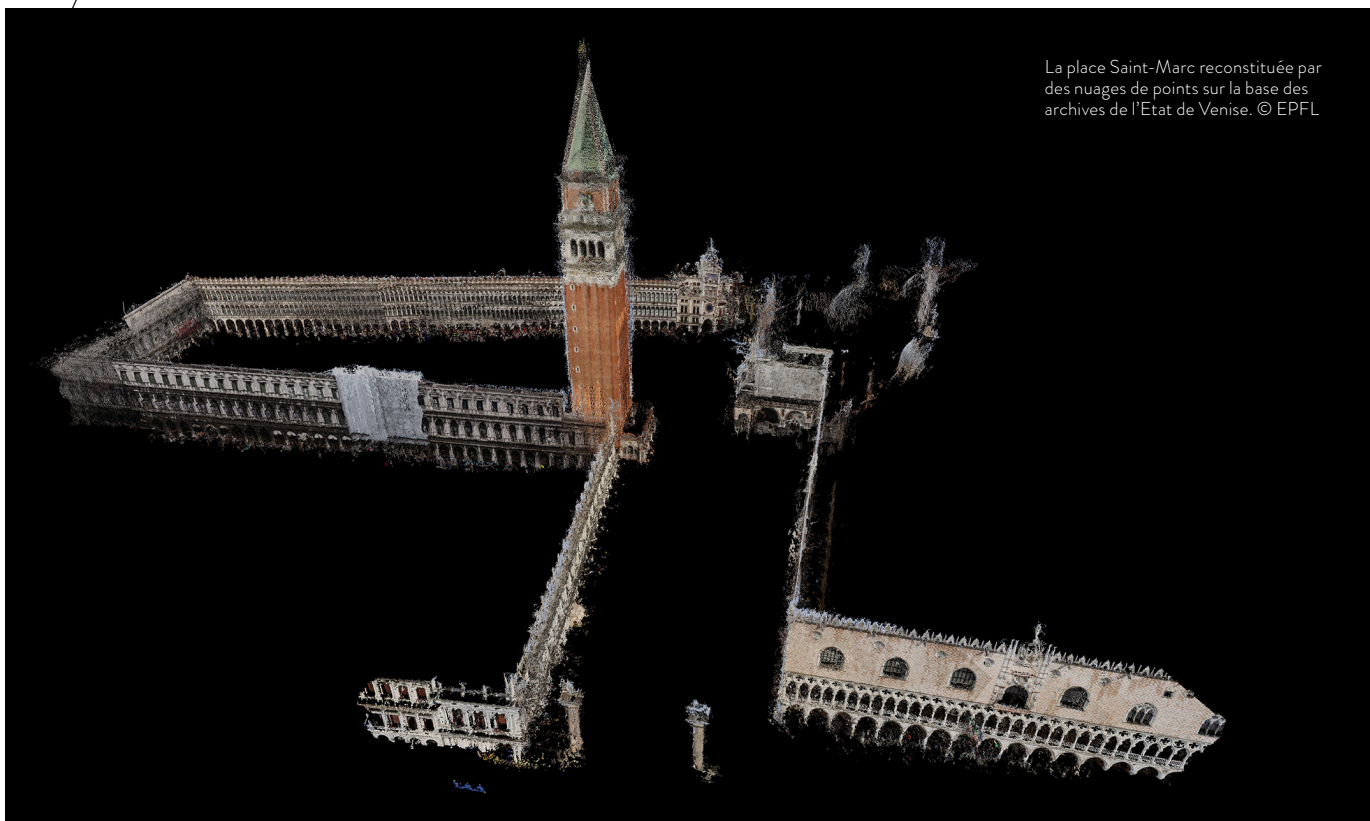


Après Venise, des villes européennes désirent leur Time Machine

2018 est une année charnière pour le projet Venice Time Machine. Lancé en 2012, il arrive à mi-parcours, marqué par deux étapes majeures. D'une part, la volonté de donner une dimension européenne au concept de Time Machine. D'autre part, l'ouverture au public et aux historiens des plus de 2 millions de documents déjà numérisés. Que recèlent ces trésors vieux de plus de 500 ans ? Réponses dans notre dossier.

Par Anne-Muriel Brouet et Sandy Evangelista



La place Saint-Marc reconstituée par des nuages de points sur la base des archives de l'Etat de Venise. © EPFL

« Tout ce qui est antérieur à l'an 2000 n'existe pas, car il n'y a pas de support pour y accéder. » Le propos de Frédéric Kaplan est radical. Le directeur du Laboratoire d'humanités digitales s'inquiète d'un monde où seules les informations numérisées sont accessibles, jetant aux oubliettes toutes celles qui vieillissent sur d'autres supports. « Il y a urgence à faire un pont entre les deux. Il faut aussi vivre avec le passé. »

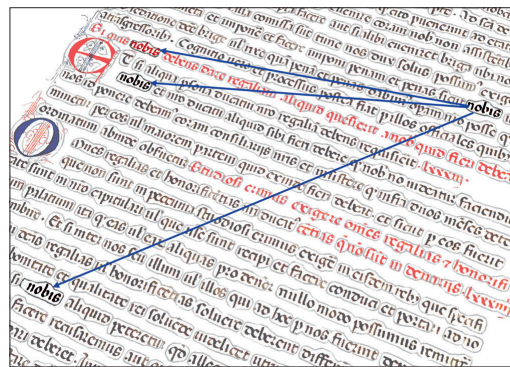
Le projet Venise Time Machine (VTM) n'est autre que l'établissement de ce pont entre passé et présent. Il vise à construire, à partir de millions de documents historiques aux formats hétéroclites, un modèle multidimensionnel de la ville de Venise, dans l'espace et le temps sur les 1000 dernières années. Lancé en 2012, VTM arrive aujourd'hui à mi-parcours. Une échéance marquée par deux étapes cruciales : la mise en ligne d'interfaces de recherche permettant d'explorer les plus de 2 millions de documents déjà scannés et l'extension du projet de Time Machine à l'échelle européenne (voir pages 7-8).

Le chemin parcouru en cinq ans – deux seulement depuis que la numérisation a commencé – se mesure d'abord en chiffres : 190'000 documents d'archives d'Etat, 720'000 documents photographiques provenant de la Fondation Cini, 3000 volumes sur l'histoire de la ville issus des grandes bibliothèques de la cité ont été scannés. « C'est moins de 1% de ce qui existe, c'est infini ! », lâche Frédéric Kaplan, avec encore davantage d'enthousiasme. Mais c'est ainsi qu'à pu être révélé un trafic d'œuvres d'art entre les Pays-Bas et l'Italie (lire page 9) !

Des technologies pour lire dans les livres scellés

Pour parvenir à ce résultat, il a fallu commencer par construire les technologies nécessaires pour scanner et numériser rapidement des kilomètres de documents, parfois vieux d'un demi-millénaire. Un premier scanner semi-automatique et utilisé pour les manuscrits et documents fragiles de grand format a été construit. Un autre, à balayage robotique, tourne les pages automatiquement des journaux et des livres. Un troisième, développé avec Adam Lowe de Factum Arte, est capable de numériser un document photographique recto verso toutes les 4 secondes.

Ce n'est pas fini. Demain, c'est à travers les livres scellés que les chercheurs pourront lire. Développé



Reconnaissance de mots similaires dans une même page. © DHLAB / EPFL

dans le Laboratoire de magnétisme quantique de Giorgio Margaritondo, le tomographe à rayons X permet de scanner un livre en entier sans avoir besoin de l'ouvrir et révéler ainsi les mots inscrits sur le papier. Ceci grâce aux encres à base de fer, employées en Europe durant plus de 1000 ans. Pour l'heure, les physiciens doivent encore extraire manuellement les pages, souvent voilées et abîmées, issues des scans par tomographie à rayons X. Les prochaines étapes incluront la construction d'un algorithme qui saura détecter de manière automatique les différentes pages.

Au-delà de la numérisation proprement dite, plus de 160'000 transcriptions manuelles de noms, lieux, mots clés ont été produites par les historiens, paléographes et archivistes sur les sources documentaires. Grâce à ces annotations sur les documents d'archives, un système de reconnaissance automatique des écritures a permis d'extraire les informations des documents. « Il faut par exemple gérer la diversité des écritures et les multiples abréviations qui n'étaient pas standardisées, explique Sofia Ares Oliveira, chercheuse au Laboratoire d'humanités digitales. Nous avons montré à l'ordinateur des exemples afin qu'il apprenne à apprendre. » Le système est ainsi capable de transcrire des mots non annotés avec un taux d'erreur sur les caractères de 4% sur chaque caractère. « C'est un peu comme lorsqu'on fait une faute de frappe dans un moteur de recherche : c'est suffisant pour que le système de recherche propose des résultats pertinents et on comprend de quoi il s'agit », précise la chercheuse.

Trois outils pour plonger dans les données

Grâce à ce travail, les scientifiques mettront en ligne d'ici à l'été un premier moteur de recherche, baptisé Canvas. Il permet d'accéder aux documents



LES ACTEURS

Venice Time Machine est un projet international de l'EPFL, de l'Université Ca'Foscari de Venise et des Archives d'Etat de Venise, en collaboration avec la Bibliothèque nationale Marciana, la Fondation Giorgio Cini, le Système de bibliothèque universitaire Ca'Foscari, l'Institut des sciences, lettres et arts de Venise.

d'archives d'Etat numérisés dans leur version originale, mais surtout d'effectuer des recherches par mot clé (une personne, un lieu, une date...) dans les contenus. L'outil fonctionne comme un wiki, ce qui permet aux utilisateurs autorisés de corriger les erreurs de lecture du système qu'ils rencontrent et d'améliorer ainsi en continu sa performance.

Le second travail de numérisation a été effectué sur la collection iconographique de la fondation Cini. Sur le million de photos d'art que possède la fondation, un premier fonds de 370'000 images a été numérisé. A nouveau, des milliers d'annotations manuelles ont été nécessaires pour entraîner l'ordinateur. Résultat, l'interface Replica proposera une recherche non seulement textuelle, mais surtout visuelle des documents, établissant des connexions entre les images. « La machine a appris à faire des liens entre la composition, la position des personnages, les différents éléments ou motifs afin d'établir des similitudes entre images. Aux historiens de l'art ensuite d'écrire leur histoire », explique Benoît

Seguin qui a conçu ce moteur de recherche dans le cadre de sa thèse qu'il défendra durant l'été.

Enfin, la troisième concrétisation de cette numérisation met les ouvrages et les journaux scientifiques des grandes bibliothèques de la Sérénissime à portée de clics. Près de 3000 ouvrages sur Venise, du XIX^e au XXI^e siècle, et plus de 100 ans de publications scientifiques sur la ville ont été numérisés dans le cadre de Linked Books, un projet soutenu par le FNS et développé par les chercheurs Giovanni Colavizza et Matteo Romanello avec Frédéric Kaplan. Un système ad hoc a été développé pour extraire automatiquement et ensuite chercher les citations. Baptisée Venice Scholar, l'interface permet d'étudier l'historiographie vénitienne à travers ces citations. Comme les deux autres, le système est ouvert et la base de données peut accueillir de nouvelles intégrations. Ces sources secondaires sont également reliées aux sources primaires de Canvas.

« Venice Time Machine a été le premier grand projet à l'EPFL dans le domaine des humanités digitales. Ce projet a, entre autres, amorcé la croissance de l'Institut des humanités digitales et la mise sur pied de la section des humanités digitales. »

Sabine Süsstrunk, directrice de l'Institut des humanités digitales

L'interface Replica établit des connexions entre les images ou les éléments qui la composent. Ici, la même position de la main droite se retrouve dans des œuvres d'auteurs différents : (de g. à dr.) *Leda et le cygne*, Leonardo da Vinci, *Leda et le cygne*, Giampietrino, *Lucretia*, Giampietrino, *Venus et Cupidon*, anonyme. © Web Gallery of Art





Time Machine dans la course au FET Flagship européen

1 65 institutions partenaires dans 32 pays ont soumis en février leur candidature à la Commission européenne dans l'espoir de décrocher le FET Flagship 2020. Un programme qui permet d'assurer le financement de la recherche pendant 10 ans à hauteur d'un milliard d'euros.

Aussi étonnant que cela puisse paraître, l'idée de multiplier les machines à remonter le temps à l'échelle européenne est arrivée presque par hasard. Dans un manifeste rédigé en 2016, Frédéric Kaplan imaginait que toutes les universités se réuniraient et travailleraient ensemble pour partager le patrimoine culturel européen. Le document a reçu un écho inattendu, plusieurs délégués des pays représentés à la Commission européenne ont été séduits et l'ont défendu. C'est ainsi que le manifeste est devenu le premier document concret d'une proposition de Flagship au Conseil européen.

Le but est de construire un simulateur à grande échelle pour cartographier 5000 ans d'histoire. À ce jour, 165 institutions partenaires issues de 32 pays ont constitué un consortium dans le but de développer de nouvelles technologies de numérisation, d'analyse, d'accès, de préservation et de communication du patrimoine culturel à grande échelle. Le dossier a été soumis ce mois. La course au FET Flagship comprend plusieurs étapes et le verdict sera rendu l'an prochain.

Plusieurs Time Machine déjà en marche

Certaines villes n'ont toutefois pas attendu l'échéance européenne pour lancer leur Time Machine. Amsterdam, Nuremberg, Paris, Jérusalem, Budapest et Naples ont décidé de plonger dans leurs archives afin de créer une immense banque de données du passé reliée à l'échelle nationale et européenne. Pourquoi ? Julia Noordegraaf, professeure du patrimoine numérique à l'Université d'Amsterdam, et Sander Münster, qui dirige le Département de la conception et de la production des médias à l'Université technique de Dresde, nous ont répondu.

Pourquoi ce projet vous intéresse-t-il ?

JN : En tant que professeur du patrimoine numérique, mon intérêt est de rassembler et promouvoir la recherche sur la réutilisation et la signification du patrimoine numérique, ainsi que son impact sur la perception et l'appréciation du patrimoine culturel.

SM : L'application Time Machine offre une grande opportunité d'intensifier la recherche sur le patrimoine culturel et la collaboration entre les chercheurs, les entreprises, les autorités publiques et les autres parties prenantes dans ce domaine.

Vous avez adhéré au projet rapidement, était-ce une évidence ?

SM : Je suis les activités de la Venice Time Machine depuis longtemps, j'étais donc très heureux de rencontrer le cerveau, Frederic Kaplan, lors d'une conférence au Canada à la mi-2017. Nous avons rapidement découvert que nous partagions un intérêt commun, il m'a invité à monter dans la machine à voyager dans le temps, ce que j'ai fait avec plaisir.

Quand Amsterdam a-t-elle voulu sa Time Machine ?

JN : Tout a commencé avec le programme de recherche Creative Amsterdam : une reconstruction de la vie culturelle d'Amsterdam en numérisant et en reliant les données sur la culture dans des secteurs tels que les arts du spectacle et les arts visuels. Nous avons rapidement réalisé deux choses : premièrement que les représentations spatiales seraient un excellent moyen d'organiser et de présenter nos données ; et, deuxièmement, que le potentiel de recherche de nos ensembles de données serait beaucoup plus grand si nous pouvions également les relier à d'autres données sur l'histoire urbaine. Cela permettrait de reconstruire le changement culturel à la lumière des développements urbains, économiques, sociaux et politiques plus larges. Lorsque nous avons découvert la Time Machine de Venise en 2016, nous avons réalisé que c'était exactement le genre d'infrastructure et de vision dont nous avons besoin pour rassembler et connecter ces sources et données hétérogènes et fragmentées.

Quels genres de documents ont été numérisés ?

JN : Les archives de la ville d'Amsterdam ont été numérisées à une échelle massive. L'une des principales sources en cours de numérisation est l'Archive des notaires d'Amsterdam 1578-1915, qui couvre 30'000 grands volumes et qui a été récemment ajoutée au registre de la Mémoire du monde de l'Unesco.

Avez-vous le droit d'accéder à tous les documents du passé à Dresde ?

SM : Non, ils sont classés, parfois ils ont un accès limité en raison des droits d'auteur. Nous avons récem-



LE DROIT DE SAVOIR VERSUS LE DROIT À L'OUBLI

Si tous les fantômes du passé marchent à vos côtés, est-ce terrible ou extraordinaire ? L'exploration des archives et leur mise à disposition au grand public posent un certain nombre de questions éthiques. Le droit à l'oubli est-il plus important que le droit à la connaissance ? Ce point est essentiel dans un projet aussi important que Time Machine, qui va explorer des milliers d'années d'archives. L'Europe reconnaît le droit à l'oubli, les Etats-Unis valorisent la liberté d'expression. Après que la Cour européenne a fait jurisprudence en 2014, en acceptant la requête d'un homme d'affaires espagnol qui demandait d'effacer les articles en ligne consacrés à sa faillite, certains se sont indignés. L'Europe n'était-elle pas en train de créer une version censurée de l'histoire ? C'est un autre combat que devront livrer les chercheurs.



Julia Noordegraaf (JN).
© Monique Kooijmans

ment mené une enquête mondiale auprès des spécialistes du patrimoine culturel sur l'accessibilité des données en termes juridiques, de propriété commerciale et de perte de documents due à un mauvais étiquetage ou au stockage local. Cela reste pour tous des obstacles majeurs.

Avons-nous besoin de connaître le passé pour construire l'avenir ?

JN : A une époque de mondialisation et de numérisation sans précédent, les gens ressentent de plus en plus le besoin d'ancrer leur identité individuelle et collective dans l'histoire. Le patrimoine est notre « ADN culturel ». L'Europe n'a plus de culture commune, alors qu'historiquement les cultures nationales, régionales et locales se sont construites les unes au travers des autres. Les jeunes ont oublié cette histoire commune. L'accès de la Time Machine au patrimoine culturel contribuera de manière significative à résoudre ce problème.

Quels outils utilisez-vous ? Allez-vous en créer de nouveaux ?

SM : Nous sommes des chercheurs, donc désireux de faire avancer les frontières de ce qui est techniquement possible. Les outils que nous développons sont dédiés à la création des modèles de villes historiques à partir de vieilles photos par photogrammétrie ou d'interfaces centrées sur l'utilisateur, basées sur un navigateur 4D. Mais ce n'est pas seulement une question d'outils. Les systèmes d'information sont des paramètres sociotechniques complexes qui interagissent entre la technologie, le contenu et les besoins des utilisateurs. Il est de notre intérêt de couvrir et d'étudier l'entier du spectre dans notre recherche.



The Western Church in Amsterdam. A still image from the video Expansion of Amsterdam in the Seventeenth Century of the Amsterdam City Archives. © DR

Demain le projet pourrait prendre une dimension européenne grâce au FET Flagship, pourquoi est-ce important ?

JN : L'histoire des villes comme Amsterdam, Venise ou Paris ne peut être reconstruite sans prendre en compte la dimension internationale. Le lieu et l'espace sont un excellent moyen de représenter l'histoire locale. Bien entendu, les gens, les idées et les produits n'étaient pas liés à des lieux spécifiques. La mobilité est donc au cœur de l'histoire hollandaise et européenne et devrait être représentée dans l'infrastructure européenne des données historiques. Pour cela, nous avons besoin d'avancées technologiques telles que proposées dans le FET Flagship.

SM : L'Europe peine encore à trouver sa propre voie dans l'ère de l'information. Etant donné que le patrimoine culturel est l'un des piliers de l'identité et de la réputation européennes, il peut être intéressant de le renforcer et d'enrichir par des moyens numériques.



> SUR INTERNET :
TIMEMACHINE.EU

« J'ai toujours eu plusieurs morts dans ma vie »

Isabella di Lenardo est historienne de l'art et de l'architecture, spécialisée dans la circulation des savoirs dans l'Europe du XVI^e au XVIII^e siècle. Grâce aux documents numérisés des archives d'Etat de Venise, elle suit pas à pas certains personnages de la Renaissance, tisse leur réseau social, décèle leurs petits problèmes voire leurs travers d'êtres humains.

« Le projet Venice Time Machine a réveillé en moi un intérêt immense pour l'analyse de grandes quantités de données. » Isabella Di Lenardo veut profiter des millions d'informations disponibles pour satisfaire sa curiosité. Comment la vision historique change-t-elle avec l'analyse non pas du rôle d'une

personne ou d'une famille, mais de 28'000 propriétaires, de toutes les veuves de Venise qui avaient des appartements, de toutes les corporations religieuses, de toutes les églises ? Qu'apprend-on ? L'histoire est-elle correcte ou à réécrire ? Les faits sont-ils remis en question ?

La première lecture des transactions immobilières balaie déjà quelques certitudes notamment sur le rôle de la bourgeoisie dans la République. « On a toujours pensé que Venise était une ville de la noblesse et de l'aristocratie, c'est faux !, insiste l'historienne qui travaille au Laboratoire des humanités digitales. Les actes notariés montrent que dès la moitié du XVIII^e siècle, la bourgeoisie acquiert de nombreux biens. » Après l'âge d'or vénitien au XVI^e siècle, le XVII^e relativement calme contribue

à générer une classe dirigeante qui vit sur ses acquis. Début du XVIII^e siècle, les caisses de l'Etat sont vides, les Vénitiens doivent opter pour une stratégie d'ouverture aux bourgeois. Un certain nombre de familles bien plus riches que les nobles qui vivaient dans les terres à Vérone, Vincenza, Brescia investissent la ville. « C'est très intéressant en termes de dynamique, si on regarde la microdimension de la noblesse à cette époque, on ne remarque pas ce transfert, d'ailleurs toutes les études faites jusqu'ici ne l'ont pas vu. »

Et que diront les déclarations d'impôts ? Les chercheurs viennent de terminer la grande phase de numérisation de 300'000 documents recto verso, qui comprennent chacune des centaines de mentions, les noms des personnes, des achats, des transactions. Cela multiplie d'autant la somme d'informations à analyser. Si le travail ne fait commencer, les premières constatations montrent que l'histoire n'est pas tout à fait celle que l'on raconte et qu'il y a des nuances que seule l'analyse des archives pourra révéler.

Des tableaux de maître cachés dans la viande de porc

Face à la masse gigantesque de données, les chercheurs tentent de démêler l'écheveau en suivant les traces laissées par un personnage particulier. « J'ai ainsi plusieurs morts dans ma vie », résume la chercheuse qui a accompagné pas à pas Joris Hoefnagel dans ses pérégrinations Venitiennes au XVI^e siècle. Ce cartographe flamand a, notamment, travaillé avec un de ses concitoyens résidant à Venise. Grâce à lui, Lodewijck Toeput a conçu la première représentation de la place Saint-Marc. Hoefnagel collabora aussi avec Abraham Ortelius, géographe et homme de lettres, qui traversa toute l'Europe en dessinant les cartes des villes où il s'arrêtait. Pour Isabella, ces artistes incarnent déjà une certaine idée

de la mondialisation et de l'identité européenne. « Hoefnagel, je l'ai rencontré pour la première fois à Venise dans un acte notarié qui donnait procuration à son frère pour retirer de l'argent. C'était très émouvant, car j'ai découvert qu'il avait un frère qu'on ne connaissait pas. »

Petit à petit, au fil des lectures, elle met à jour le réseau de personnes qui l'ont accueilli à Venise en 1560. Parmi elles, des marchands de pierres précieuses d'Anvers, des Flamands très probablement à l'origine de l'importation de nombreux tableaux en Italie – un sujet qu'elle a exploré il y a plusieurs années pour préparer sa thèse. Ces tableaux ont été enregistrés et taxés en Flandre, mais l'historienne n'en trouve trace à Venise. « J'ai cherché les impôts d'entrée de ces tableaux... aucun. Je me demandais où ils étaient passés. Et j'ai découvert que, déjà à l'époque, il y avait un marché noir : les peintres ne voulaient pas payer d'impôts et cachaient leurs œuvres dans des caisses transportant de la viande de porc ou des tapis. » Un exemple qui montre qu'on ne peut pas s'appuyer seulement sur les archives d'un pays; les histoires sont connectées et on ne découvre les réponses qu'en croisant les informations, toutes les réponses ne sont pas à Venise.

Même les dessins pouvaient mentir

Il s'est avéré aussi que Jacopo de' Barbari, peintre et graveur du XV^e siècle, a pris des libertés dans ses cartes représentant Venise. Les erreurs ont été mises à jour grâce aux données cadastrales. En reconstituant la ville de cette époque, il est apparu que deux quartiers importants de Venise dessinés par de' Barbari, le Rialto, centre des affaires, et l'Arsenale, où l'on construisait les galères de la République, ne correspondaient pas à la réalité. Le peintre avait omis certains détails. « Je pense que ces omissions étaient dictées par une politique de sécurité. Il ne fallait pas montrer aux ennemis l'endroit stratégique où ils pouvaient potentiellement frapper. »

L'historienne vient de mettre un point final à un article consacré au peintre allemand Albrecht Dürer lorsqu'il séjournait à Venise. « Ça fait un an que je m'occupe de lui. J'ai reconstitué son réseau de connaissances, déterminé qui il a rencontré. En histoire on n'a jamais tous les documents, il faut donc faire des simulations, essayer d'interpréter les lettres, pourquoi a-t-il écrit à untel, qui essayait-il de rencontrer ? » Les réponses se trouvent peut-être dans les 80 km d'archives vénitiennes en voie de numérisation.



Isabella di Lenardo.
© Hillary Sanctuary

Piazzetta e Palazzo ducale.
© Hoefnagel

