



HAL
open science

La carte, de l'analogique au numérique. Modes de traitement et éléments de méthode

Jean-Luc Arnaud, Romain Suarez

► To cite this version:

Jean-Luc Arnaud, Romain Suarez. La carte, de l'analogique au numérique. Modes de traitement et éléments de méthode. Humanités numériques, 2023, 7. halshs-04210469

HAL Id: halshs-04210469

<https://shs.hal.science/halshs-04210469>

Submitted on 18 Sep 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 - ShareAlike 4.0 International License



La carte, de l’analogique au numérique. Modes de traitement et éléments de méthode

D’après : Jean-Luc Arnaud et Romain Suarez, « La carte, de l’analogique au numérique. Modes de traitement et éléments de méthode », *Humanités numériques* [en ligne], 7-2023.

Résumé

Dans le cadre d’un programme de recherche dédié à la vectorisation automatique des documents cartographiques anciens, les auteurs examinent les cartes comme des assemblages de signes conventionnels qui figurent le réel de manière à la fois partielle et partielle. Après un examen détaillé des difficultés que soulève l’interprétation des documents et des ambiguïtés qu’ils présentent, ils considèrent la catégorisation des données comme un point crucial et proposent une approche à la fois multi-scalaire et inter-scalaire pour faciliter leur investigation.

Abstract

In the context of a research program dedicated to the automatic vectorisation of ancient cartographic documents, the authors examine maps as combinations of conventional symbols representing reality in a partial and biased way. After a detailed examination of the problems raised by the interpretation of documents and of the ambiguities they might contain, the authors consider the categorisation of the data as a crucial point and propose a multi-scalar and inter-scalar approach to facilitate their investigation.



Ce texte et ces dessins sont sous licence creative common : [Attribution – ShareAlike 4.0. \(CC-BY-SA\)](#)

This text and theses drawings are under creative common license: [Attribution – ShareAlike 4.0 \(CC-BY-SA\)](#)

Jean-luc.arnaud@univ-amu.fr – romain.suarez@univ-amu.fr

La carte, de l’analogique au numérique. Modes de traitement et éléments de méthode

Les publications et les travaux relatifs à l’automatisation des processus de vectorisation des documents cartographiques se sont multipliés au cours des dernières années¹. Pour leur plus grande part, ces travaux trouvent leur origine dans des laboratoires de traitement informatique de données². En 2020, sous la responsabilité d’un cartographe, Dominique Andrieu, la maison des sciences de l’homme (MSH) Val de Loire a ouvert un nouveau chantier sur cette question – programme Veccar – avec la collaboration de plusieurs membres d’autres MSH³.

Chaque groupe de participants a proposé de traiter les documents qui l’intéressent le plus. Ainsi, entre des plans cadastraux manuscrits, des cartes topographiques gravées en taille douce et des cartes routières imprimées en couleurs, les premiers documents candidats à la vectorisation automatique étaient très divers à la fois par leur facture, par leur contenu thématique ou encore par leur mode de catégorisation du réel. Face à cette diversité, les informaticiens ont proposé de concentrer les premiers travaux sur deux documents : une feuille de la carte de France à l’échelle 1:100 000 publiées au milieu des années 1950 (Arnaud 2022, 372 et suiv.) et une feuille manuscrite du plan cadastral napoléonien. Parmi les exemples proposés, ces cartes sont celles qui présentent le moins d’ambiguïtés. La carte au 1:100 000, imprimée en huit couleurs a bénéficié d’une rédaction particulièrement soignée, sa légende, ne compte pas moins de 90 entrées. Pour leur part, les plans cadastraux figurent la propriété foncière suivant une catégorisation réduite à sa plus simple expression : espace public, propriété privée, construit, non construit. Au contraire des autres documents proposés, dont une grande part des légendes est souvent implicite, ces documents laissent peu de place à l’interprétation de leurs codes graphiques. Les échanges relatifs à cette sélection suggéraient que certaines cartes sont plus aisées à traiter que d’autres dans la mesure où plus l’objectif assigné aux documents est peu déterminé ou bien multiple, plus ils figurent des données de diverses natures et plus leur traitement semblent complexe en ce sens que leur restitution sous forme vectorielle nécessite plus de références.

¹ . Les progrès réalisés depuis une quarantaine d’années dans les techniques de traitement des images, les politiques de numérisation des établissements et le développement concomitant des systèmes d’information géographique (SIG) ont permis aux équipes de recherche de traiter des « contenus » géographiques anciens avec des méthodes d’analyse et des traitements informatiques de plus en plus performants. Voir la synthèse proposée par Chiang, Leik et Knoblock (2014), et, plus récemment *Automatic Vectorisation...* (2020), Chen *et al.* (2021). Voir également le résumé des contributions au workshop *Automatic Vectorisation of Historical Maps*, organisé par la ICA Commission on Cartographic Heritage into the Digital, à Budapest, le 13 mars 2020 : <http://lazarus.elte.hu/avhm/abstracts.html>

² . Faisant exception, le projet Alpage, qui porte sur un plan très détaillé de Paris au début du XIXe siècle, résulte de la collaboration entre plusieurs laboratoires en sciences sociales et un laboratoire d’informatique. <https://alpage.huma-num.fr/>.

³ . Veccar – Vectorisation de cartes anciennes – est porté par la MSH Val-de-Loire et le LIFAT – Laboratoire d’Informatique fondamentale et appliquée de Tours. Les partenaires du projet sont la Maison de l’Orient et de la Méditerranée (Lyon), la Maison méditerranéenne des sciences de l’homme (Aix-en-Provence), la maison des sciences de l’homme et de l’environnement (Besançon), la Maison des sciences de l’homme et de la société Sud-Est (Nice), et le Centre d’expertise et de transfert universitaire de l’université de Tours dans le domaine de l’Innovation logicielle en traitement d’images & acquisition de données 3D (Illiad3).

Au-delà des premiers résultats obtenus et qui semblent assez prometteurs, le traitement de ces documents a montré qu'une mise en ordre des difficultés relatives à l'interprétation et à la catégorisation des multiples codes graphiques est nécessaire. Cette contribution en rassemble les résultats. Les auteurs ont participé au programme Veccar en raison de leur connaissance des documents anciens et de leur expérience de traitement des cartes dans le cadre d'une plateforme à vocation documentaire⁴. Ainsi, leur approche s'appuie sur leur compétence en histoire de la cartographie et sur leur expérience en matière de vectorisation manuelle.

La carte, une construction partielle et partielle

La carte n'est pas le territoire, si c'était le cas, chaque territoire serait l'objet d'une unique représentation cartographique à même de répondre à tous les usages. Au contraire, comme de nombreux atlas en témoignent, la description cartographique des différents aspects d'un territoire donne lieu à la production d'une multitude de documents, qui, chacun à son échelle et suivant son thème, rend compte de manière partielle de son organisation. A ce titre, chaque carte constitue une représentation, résultat d'une sélection draconienne des éléments du réel et de leur regroupement en catégories construites en fonction de l'utilisation envisagée (Vallongue 1826). Ainsi, chaque carte n'est pas seulement partielle, elle est aussi partielle, en ce sens, elle porte sur le territoire qu'elle figure un point de vue de connaissance particulier.

Suivant cette proposition, la sélection d'un document à vectoriser est soumise à la correspondance entre la catégorisation proposée par son code graphique, formalisé ou non à travers une légende, et celle requise pour répondre aux questions de recherche envisagées. Ainsi par exemple, pour la carte de France à l'échelle 1:20 000, jusqu'au milieu des années quarante, les voies de circulation sont figurées en fonction de leur classement administratif (national, départemental et chemin vicinal). Or, certaines routes nationales ne sont pas plus praticables que des chemins bien entretenus. A partir de 1945, pour faciliter le choix des itinéraires, on adopte un mode de classement des routes qui intègre leur viabilité. Mais, pour éviter « un déséquilibre et une confusion inadmissibles », la carte figure seulement les déclassements inhérents à la viabilité. En d'autres termes les meilleurs chemins ne sont pas surclassés mais les routes nationales et départementales en piteux états sont déclassées (Couzinet 1948, 11). Lorsqu'elle est énoncée, la prise en compte de cette information semble fondamentale pour, en amont de la vectorisation, sélectionner la version – *ante* ou *post* 1945 – qui correspond le mieux à ses objectifs. La principale difficulté réside dans l'accès à ces informations dans la mesure où elles ont été peu publiées. Lorsque c'est le cas, c'est le plus souvent à travers des documents à usage interne, au sein des agences de production⁵.

Documenter les protocoles de production

Chaque carte est un assemblage de signes qui, sous peine d'illisibilité, ne peuvent pas tous être au même niveau ; les rédacteurs doivent établir des priorités à travers une hiérarchisation de l'information. La hiérarchie entre les éléments d'une carte est toujours l'objet d'âpres discussions entre ses commanditaires, chacun accordant une importance

⁴ . *CartoMundi – Valorisation en ligne du patrimoine cartographique*. <http://cartomundi.fr>.

⁵ . En l'occurrence, l'information donnée ici est tirée d'un opuscule tiré à quelques centaines d'exemplaires seulement (Couzinet 1948).

relative aux éléments figurés sur la carte en fonction de son domaine de compétence. Ces difficultés sont d'autant plus fortes que la carte se propose de répondre aux besoins d'un plus grand nombre d'utilisations. Ainsi, par exemple, sur les cartes routières, la catégorisation des voies en fonction de leur largeur, de leur revêtement, de leur entretien... donne lieu à une multiplication des signes conventionnels tandis que les voies de chemin de fer, larges ou étroites, simples ou doubles, électrifiées ou non, sont figurées par un signe unique – des traits fins – qui constituent seulement des points de repère dans le paysage routier à travers les passages à niveau ou bien les ponts. Au contraire, les cartes topographiques publiées par les agences nationales s'adressent à un public plus large. En France, pour la carte topographique de base à l'échelle 1:20 000 imprimée en trois couleurs (Arnaud 2022, 312 et suiv.), la charte graphique est fondée sur une hiérarchie qui « répond aux désirs et aux besoins de la majorité des usagers ». Ainsi, elle place les voies de communication au plus haut tandis que les bois et les cultures occupent la dernière position (Couzinet 1948, 5). Cette répartition privilégie les flux et les agglomérations au détriment du monde rural et du paysage ; elle n'est pas partagée par tous les producteurs. En Italie par exemple, pour la carte à l'échelle 1:25 000, « les détails représentés ont généralement, dans l'ensemble du dessin une apparence du même ordre que celui dans lequel ils sont perceptibles dans la nature, c'est-à-dire, dans l'ordre décroissant : formes du terrain, habitation, voies de communication, cultures et détails » (Couzinet 1948, 4). Ces différences ne sont pas sans incidence sur la vectorisation dans la mesure où les superpositions entre les signes conventionnels donnent lieu à des représentations différentes en fonction de la hiérarchie adoptée.

Depuis au moins un siècle et demi, les cartographes organisent leurs travaux de rédaction sur la base d'instructions qui, pour chaque document, détermine le code graphique applicable à chaque catégorie d'information. Ces instructions sont bien plus complètes que les légendes livrées aux utilisateurs. Chaque catégorie est l'objet d'une description minutieuse qui en indique les limites, par des seuils de mesures en particulier. Ainsi par exemple, sur la carte de France à l'échelle 1:20 000, on représente une clôture isolée seulement si elle se développe sur une longueur de plus de 50 mètres (*Tableau commenté...* 1960, 61). De manière générale, ces instructions, reproduites en quelques exemplaires seulement et réservées à l'usage des opérateurs, ont été peu diffusées et mal – ou pas – conservées⁶. Or, comme l'ont bien montré Chappart et Reynard (2007) à propos de la figuration des jardins en zone suburbaine, ces informations sont requises pour l'interprétation et l'exploitation des données. Ainsi, pour chaque document – ou chaque catégorie de document –, la connaissance du processus de construction des catégories qui a présidé à son élaboration constitue de toute évidence un préalable nécessaire à sa vectorisation.

Lorsque ce travail préliminaire est négligé ou bien que la documentation nécessaire n'est pas disponible, l'opérateur interprète chaque point, ligne ou surface en mobilisant un savoir souvent peu formalisé qui combine une légende implicite avec ses connaissances relatives à l'organisation de l'espace. Lorsque cette organisation échappe à sa culture et à ses catégories – autour des lieux intermédiaires entre espace public et sphère domestique dans les villes centrafricaines par exemple, dont les catégories européennes sont incapables

⁶ . Ainsi par exemple, le fascicule cité ici semble avoir échappé au dépôt légal ; il n'apparaît pas au catalogue de la Bibliothèque nationale de France.

de rendre compte –, il est susceptible de procéder à des agrégations ou bien à des partages catégoriels aberrants⁷.

Copier, transposer, catégoriser

La transposition d'un document analogique en données numériques n'est pas une simple opération de copie terme à terme. Pour être exploitables, les données doivent être structurées en catégories. Or, malgré leur code graphique, malgré leur légende, les documents analogiques présentent toujours des aspects susceptibles d'interprétations multiples – en milieu urbain dense, la ligne de façade d'un immeuble est aussi la limite d'une rue. Dans la mesure où elle oriente les investigations envisageables, la catégorisation des informations n'est jamais une opération neutre. Elle est d'autant plus délicate à opérer que les similitudes lexicales sont trompeuses.

Les figurations cartographiques analogiques présentent à l'examen visuel quatre types de signes : des points, des lignes, des surfaces et des écritures. Pour leur part, les systèmes d'information numérique gèrent des unités spatiales qui se partagent entre trois catégories : des points, des lignes et des surfaces. Suivant cette présentation, les deux classifications semblent proches ; il apparaît ainsi que le passage d'un support à l'autre puisse être aisément automatisé. Cependant, la similitude du vocabulaire cache des objets différents à la fois par leur mode de construction et par leur nature. Autrement dit, leur transposition présente des difficultés.

Sur les supports analogiques, les unités ne sont jamais figurées comme de purs objets géométriques sous peine d'être invisibles. Au contraire, pour rendre la carte intelligible, son rédacteur utilise des signes conventionnels qui donnent une visibilité aux objets. Avec ces signes, les points et les lignes prennent de l'épaisseur – à ce titre ils deviennent des surfaces d'encre déposées sur le papier –, des formes ou encore des couleurs, tandis que les surfaces sont figurées par leur périmètre (à travers une ligne fermée) ou bien par un aplat de couleur ou bien un poncif – motif répété. Enfin, les écritures, qui accompagnent la plus grande part des documents cartographiques, sont composées de signes imprimés superposés aux informations géographiques.

Pour leur part, les unités géométriques enregistrées dans les bases de données sont regroupées et structurées en couches d'informations indépendantes les unes des autres. La mise en forme de ces données et le choix de la charte graphique sont déterminés *a posteriori* et peuvent faire l'objet d'autant de versions qu'il existe de besoins. Ces différences entre les cartes sous forme analogique et les bases de données sont assez fortes pour que la vectorisation ne puisse pas être réduite à une simple transcription, sous peine d'obtenir des données vectorielles inexploitables. Au tournant des années 2000, de nombreuses collectivités territoriales mal conseillées ont fait numériser des plans de leur territoire respectif sans spécification, en se basant sur le bon sens et les capacités de discernement des opérateurs. Les bases de données qui résultent de ces opérations se révèlent le plus souvent inutilisables, sinon pour produire de l'imagerie⁸.

⁷. Une part importante des documents cartographiques ne comporte pas de légende. Pour les cartes qui en sont accompagnées, elles sont souvent incomplètes et ne figurent pas d'indications pour les postes considérés comme « allant de soi ». De fait, les auteurs considèrent le plus souvent que les lecteurs partagent avec eux une culture cartographique suffisante pour permettre l'interprétation des documents qu'ils publient.

⁸. C'était le cas par exemple, du SIG exploité par la commune de Martigues (Bouches-du-Rhône) jusqu'à une période récente et qui comportait une multitude de couches d'informations relatives à l'occupation des sols



Figure 1. Chaque ligne sur le papier peut correspondre à plusieurs vecteurs superposés
Sur le papier, chaque ligne de façade sur rue correspond en version numérique à au moins trois tracés superposés : un appartient à la limite du bâtiment (bleu), un autre au polygone de la parcelle (orange), un troisième à la délimitation de l’îlot (rouge). D’autres encore peuvent définir des unités techniques d’intervention ou des ressorts administratifs. Exemple du quartier de la Rotonde à Marseille. Dessin Jean-Luc Arnaud, 2022.

Ainsi, vectoriser à des fins scientifiques ou administratives, ce n’est jamais recopier. Un simple trait sur une carte peut prendre la forme de plusieurs tracés dans la version vectorielle. Par exemple, dans un plan parcellaire, un trait de séparation entre deux unités donne lieu à deux tracés superposés, chacun appartenant au polygone qui figure une des deux unités (figure 1). Vectoriser c’est donc, *de facto*, produire une catégorisation qui, pour être opératoire, est nécessairement construite en relation étroite avec les questions à poser au document. Pour un plan de ville par exemple, le maillage du réseau de voirie ou le découpage de la propriété foncière constituent des catégories différentes et aboutiront *in fine* à la production de couches vectorielles spécifiques. En d’autres termes, vectoriser, c’est porter un point de vue de connaissance particulier sur un document. On ne vectorise pas *a priori*, mais, de la même manière qu’on consulte une carte, on le fait avec des questions.

Construire la vérité-terrain

Cette première conclusion a pour effet de questionner le protocole envisagé dans le cadre de Veccar par les participants chargés du développement informatique. Reprenant des méthodes mises au point en traitement d’images, les informaticiens se sont proposés d’évaluer les résultats de leurs travaux en mesurant les écarts obtenus entre les données construites par vectorisation manuelle, désignées *vérité-terrain* et les données vectorielles obtenues par apprentissage supervisé. Nous avons ainsi été sollicités pour produire le fichier de la *vérité-terrain* de chaque document soumis aux essais de vectorisation automatique. Cependant, au contraire des livres qui sont codés de manière séquentielle, les cartes sont des documents synoptiques dont l’organisation ne comporte pas d’indication sur la manière d’en conduire l’investigation. On ne lit pas une carte *a priori*, on la consulte

présentant la particularité malheureuse de multiples superpositions et de discontinuités totalement incompréhensibles et inexploitable.

pour répondre à des questions. Chaque question guide la consultation en ce sens qu'elle détermine une interprétation particulière de la représentation et des signes conventionnels. Pour chacune, la réponse donne lieu à une *vérité-terrain* spécifique. Il est ainsi apparu que cette notion, mise au point pour le traitement des images, est peu opératoire pour les cartes tant les résultats peuvent être variables et tant elle est déterminée par le point de vue de connaissance porté sur chaque document.

Une expérience qui consisterait à demander à des contributeurs de disciplines différentes de produire, sans spécification posée *a priori*, des versions vectorielles d'un même document permettrait de conforter la variabilité des points de vue. Elle offrirait aussi un moyen de classer les catégories d'informations géographiques entre celles qui ont été retenues par le plus grand nombre de participants à l'expérience et celles qui intéressent des groupes plus restreints, voire des contributeurs uniques. Sur la base d'un tel classement, on pourrait envisager de définir un socle des catégories les plus partagées et, éventuellement, des nébuleuses de catégories qui, chacune, intéresse une discipline, une approche, une théorie, une période... Si tant est qu'il soit envisageable de caractériser un tel socle, l'expérience permettrait de renouer avec la notion de *vérité-terrain*. Il s'agirait alors d'en réduire la portée et les éléments de manière à s'affranchir des particularités des différents points de vue de connaissance. Cet enchaînement d'hypothèses reste à vérifier, en l'absence de socle et, *de facto*, de *vérité-terrain* pertinente, il semble prudent de considérer que le résultat d'une vectorisation est opératoire seulement si les données sont organisées suivant des catégories adossées à un point de vue de connaissance, en fonction de chaque discipline et pour chaque problématique. L'évaluation proprement dite intervient dans un second temps. Imaginer qu'il est concevable de procéder automatiquement à la construction d'un panel de catégories assez complet pour satisfaire tous les usages, c'est faire l'hypothèse qu'il est possible de dresser la liste de toutes les questions envisageables. Or, nul ne sait quelles questions émergeront au cours des prochaines années ni lesquelles tomberont en désuétude. L'histoire des sciences est tout d'abord celle d'un renouvellement des approches et des questions. Combien d'historiens retournent aux sources qu'ils ont consultées quelques décennies plus tôt pour en effectuer des relectures informées à la fois par leurs nouvelles connaissances et par de nouvelles approches.

Ambiguïtés, superpositions et autres difficultés

Pour nos premières tentatives de mise en ordre des documents dans le cadre de Veccar, nous avons considéré chaque carte dans son ensemble. Une telle approche permet effectivement d'opérer un classement initial – monochrome ou polychrome par exemple – mais de nombreuses cartes présentent de fortes variations de leur code graphique suivant les différents milieux qu'elles figurent.

Ainsi, aux échelles moyennes, la figuration des villes constitue le plus souvent des exceptions graphiques étrangères à la représentation du monde rural environnant. Il est d'autant plus difficile de faire passer une ligne de séparation entre ces deux milieux que l'urbanisation présente une faible densité. Ainsi, la forte hétérogénéité de la plus grande part des documents interdit de leur assigner une catégorie d'ensemble. Fort de ce constat, nous avons choisi d'examiner les difficultés rencontrées au cours du processus de vectorisation, quel que soit le type de document, en les regroupant en quatre catégories.

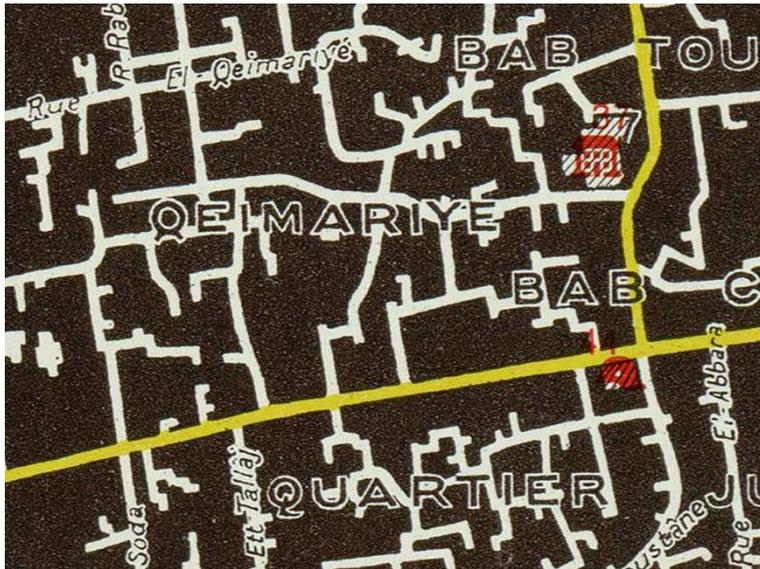


Figure 2. Interférence de la toponymie

Dans certains plans de villes, la toponymie interfère fortement avec le tracé des rues. Extrait de *Stadtplan von Damaskus* [Damas], 1:10 000, Berlin, s.n., 1941. Document Université Paris 8 Vincennes – Saint-Denis.

Lacunes

Sur les documents, les lacunes prennent le plus souvent la forme d'interruption des tracés. Elles peuvent être natives ou bien provenir de la mauvaise conservation du document. Les lacunes peuvent également être d'origine intentionnelle ou bien résulter des instructions relatives à la manière de figurer le signe. Par exemple, pour la carte de France au 1:20 000, les signes qui représentent les limites administratives sont interrompus dans la traversée des lacs, des bras de mer ou des landes lorsque la limite correspondante suit une ligne droite (Couzinet, 1948, 33).

Pour les unités linéaires, les lacunes donnent lieu à des discontinuités d'autant plus aisées à restituer qu'elles sont de faible ampleur. Au contraire, plus les lacunes sont importantes, plus les hypothèses de restitution envisageables sont multiples. La culture cartographique et topographique de l'opérateur s'avère alors déterminante.

Pour les surfaces, les signes linéaires qui les délimitent ne peuvent être enregistrés de manière numérique qu'à la condition d'être fermés. Ainsi, pour ne pas « perdre » l'ensemble d'une surface, les opérateurs sont souvent conduits à fermer des polygones alors que la carte présente des ruptures. Par ailleurs, les périmètres interrompus donnent lieu à des continuités entre des unités surfaciques susceptibles d'appartenir à des catégories différentes. Par exemple, la figuration des espaces publics en milieu urbain est rarement partagée en fonction de la toponymie commune. Une catégorisation qui distinguerait les places et les rues rencontrerait dans de nombreux exemples les plus grandes difficultés à les délimiter.

Les superpositions entre plusieurs signes conventionnels sont également à l'origine de lacunes. Les écritures en fournissent de nombreux exemples et il est assez aisé de les identifier si tant est que l'on sait lire ou, au moins, identifier les lettres de l'alphabet concerné. Dans certains cas, pour ne pas nuire à la lisibilité des écritures, les signes imprimés dans la même couleur et qui les croisent sont interrompus. C'est souvent le cas avec les limites administratives qui doivent constituer des périmètres fermés. Leur restitution est d'autant plus délicate que les écritures sont plus grandes (figure 2).



Figure 3. Difficultés d'interprétation

Les cercles qui ressemblent à des ronds-points figurent des villages. Extrait de la feuille 51, « Limoges », de *Carte de France à 1/200.000*, Paris, Institut géographique national, 1959. Document Institut national de l'information géographique et forestière - IGN.

Lorsque les signes à l'origine d'une lacune ne présentent pas d'interférence graphique, comme c'est le cas lorsque deux lignes de couleurs différentes sont superposées, il est assez aisé de restituer la partie manquante. Cette situation est très fréquente, dans les croisements entre les réseaux routier, ferré et hydrographique. Mais le nombre de couleurs employées pour l'impression des cartes est souvent limité. Ainsi, la différence entre les signes linéaires peut être portée par leur largeur ou par leur composition. Par exemple, en Allemagne, pour les cartes au 1:25 000 des années quarante imprimées en noir seul, les rédacteurs ont démultiplié les signes linéaires composés par l'association de traits continus, de pointillés et de tirets d'amplitudes différentes (Couzinet 1947). Au contraire, pour la carte de France au 1:200 000 de la fin du XIXe siècle, le réseau routier, dont la plus grande part est figurée en rouge, est ponctué par de nombreux cercles, aussi imprimés en rouge, qui interrompent le tracé des routes comme autant d'immenses ronds-points. Un grand nombre de ces cercles raccordent seulement deux voies face à face et d'autres forment des culs-de-sac (figure 3). De toute évidence, il ne s'agit pas de ronds-points. Effectivement, à chaque cercle correspond le nom d'un village. Une vectorisation qui ne tiendrait pas compte de cette particularité aurait pour effet soit de réduire le réseau routier à une suite de segments sans continuité, soit de figurer une rocade autour de chaque village (ce type d'équipement est alors totalement absent du réseau routier). Cet exemple montre que les clés d'interprétation d'un signe lui sont souvent externes. L'examen du signe seul ne donne pas de réponse. Ce sont les signes environnants qui en portent le sens.

Bifurcations

Les tracés portés sur les cartes comportent de nombreuses bifurcations ou bien des figures en T, qui proposent plusieurs solutions de continuité. Dans le cadre de leur vectorisation, l'opérateur a le choix entre trois options :

- * poursuivre la ligne d'un côté
- * poursuivre la ligne de l'autre côté
- * interrompre les 3 lignes au point de jonction



Figure 4. Bifurcations

A. Le fait que les surfaces figurant les bâtiments soient agrégées sans tenir compte des unités de construction interdit de restituer les limites parcellaires. Les figures en T ne sont pas interprétables. Extrait de *Niort*, 1:5000, manuscrit, 1842 (document IGN). B. La numérotation des routes – en rouge – permet d'identifier chaque unité sans ambiguïté. La route n° 39 traverse Broyes d'où débute la route n° 45. Extrait de La Barre Duparcq ; Charles-Henry (de), *Département de la Marne, Atlas cantonal, carte générale révisée en 1911-1912*, 1:200 000, Paris, Erhard, s.d. Document IGN.

La dernière option, la plus simple, est rarement la plus pertinente. Pour une rivière par exemple, elle donnerait lieu à son partage en biefs séparés par chaque confluent. Pour les routes, le résultat serait semblable entre chaque carrefour. Cependant, si les rivières et leurs affluents sont généralement figurés par le même signe conventionnel, les voies de circulation routières sont l'objet de classements qui donnent lieu à la multiplication des signes. Ainsi, les différences graphiques entre les lignes se rencontrant à chaque carrefour offrent à l'opérateur des continuités ou bien des ruptures qui facilitent sa tâche (figure 4).

Lorsque les lignes en question délimitent des surfaces, la troisième option n'est pas envisageable, l'opérateur doit choisir un type de prolongement. Dans les plans cadastraux ou parcellaires qui figurent à la fois les périmètres des unités de propriété et ceux des constructions par des lignes de même qualité, le choix à opérer à chaque bifurcation est déterminé à la fois par la topologie – de manière générale les unités de construction ne chevauchent pas plusieurs parcelles – et par des indicateurs externes, tels que des numéros ou encore des flèches qui associent des unités contiguës.

Les plans très détaillés, présentent aussi des figures en T dans la représentation des escaliers. Dans certains cas, il n'est pas envisageable de définir s'ils appartiennent à l'espace public ou bien à la parcelle qu'ils bordent. Par ailleurs, chaque emmarchement apparaît comme une unité indépendante – un polygone – alors que chaque escalier constitue une entité. Les escaliers sont identifiables à la fois par la taille des unités qui les composent (elles sont bien plus petites que les plus petites constructions) et par leur topologie, des ensembles fermés composés de lignes parallèles et équidistantes. Ces critères d'identification ne sont pas applicables aux emmarchements isolés, plus difficiles à caractériser.

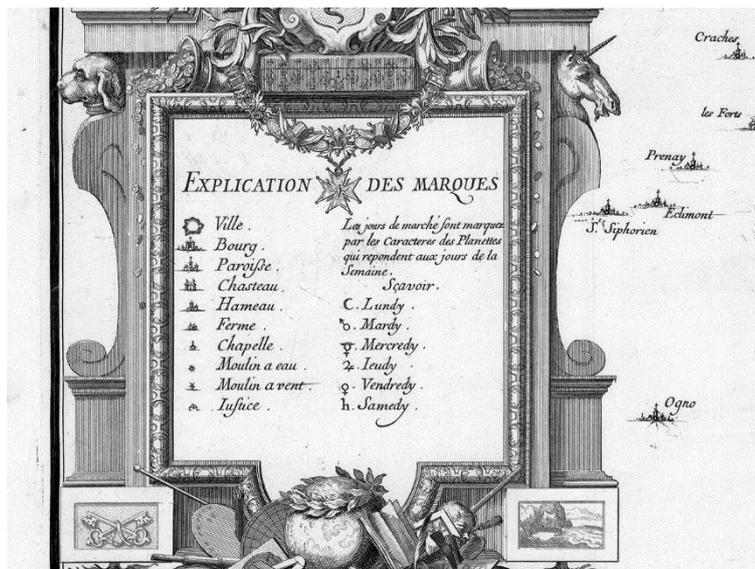


Figure 5. Chacun sa légende
Exemple d'une légende unique en son genre pour indiquer les jours de marché dans les villages des environs de Paris à la fin du XVIIe siècle. Extrait de la feuille 7 de : *Carte particulière des environs de Paris*, Paris, Académie royale des Sciences, 1678. Document IGN.

Signes conventionnels à géométrie variable

Les cartes sont des ensembles de signes qui n'ont pas toujours été conventionnels dans la mesure où, jusqu'à une période récente, ils étaient peu partagés. Jusqu'à la fin du XVIIIe siècle, chaque rédacteur établissait la légende (souvent implicite) de sa carte en fonction des objectifs qui lui étaient assignés (figure 5), des moyens techniques à sa disposition et de son sens de l'esthétique. François de Dainville (1964) a patiemment collecté et classé les signes utilisés pendant la période moderne ; son travail témoigne à la fois de leur variabilité et de l'inventivité des auteurs. Ces signes ne sont pas toujours aisés à interpréter, ils le sont d'autant moins lorsque les catégories – réalités - dont ils rendent compte ne sont plus d'actualité tels les fours à chaux, les commanderies ou encore les moulins à papier. Pour les documents plus récents, postérieurs à 1802, les signes présentent plus d'uniformité⁹, ils restent cependant différents entre les pays et, au sein de chaque pays, entre les établissements producteurs de cartes. Pendant la Première Guerre mondiale, les pratiques cartographiques françaises et anglaises sont assez différentes pour que les cartes échangées entre les Alliés soient accompagnées de calques de transcription (*Rapport du service...* 1924, 181).

De manière générale, plus les cartes sont récentes, plus la codification de leurs signes est forte. Cependant, pour l'occupation du sol en particulier – figurée par des motifs répétés juxtaposés –, les rédacteurs prennent des libertés et s'autorisent à suivre un modèle sans pour autant le recopier terme à terme. Les exemples relatifs au couvert végétal, aux zones entre deux eaux – marais, salines, récifs – et aux rochers sont multiples ; celui du cimetière musulman de Smyrne (Izmir) est plus rare (figure 6). Ce mode de rédaction des signes facilite leur adaptation aux contours des zones qu'ils qualifient. Cependant, lorsque ces signes portent des lignes, il n'est pas toujours aisé de distinguer les figurations topographiques du remplissage symbolique des surfaces.

⁹ . En France, pays précurseur en la matière, une commission composée de représentants de multiples administrations qui produisent des cartes est chargée en 1802 de définir des règles partagées de rédaction et les signes conventionnels (*Mémorial an XI*).

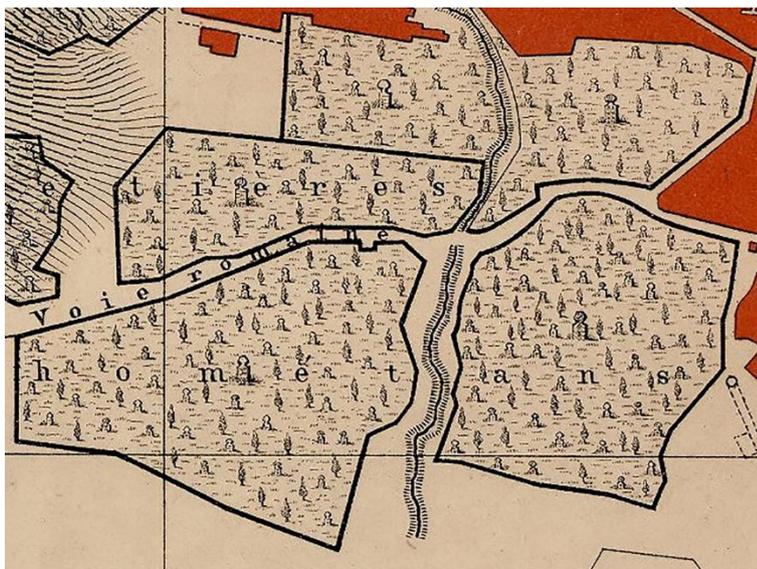


Figure 6. Signe conventionnel à géométrie variable
Pour figurer le cimetière musulman de Smyrne [Izmir], l'auteur a choisi de dessiner une alternance d'arbres et de stèles dont la forme est variable. Extrait de Lamec Saad, *Plan de Smyrne*, 1:5000, Leipzig, Wagner & Debes, 1876. Document IGN.

En d'autres termes, le seuil qui sépare ces deux catégories n'est pas facile à déterminer. La figuration des escaliers dans les plans détaillés illustre bien cette difficulté. Le plan cadastral du Caire à l'échelle 1:500 présente la particularité de figurer les escaliers situés dans les jardins ou dans les cours des propriétés privées (figure 7). Leur représentation semble sans ambiguïté, l'emprise au sol de chaque escalier est découpée par des lignes qui semblent figurer les emmarchements ; comme il se doit, elles sont parallèles dans les escaliers droits et elles rayonnent dans les quartiers tournants. Cependant, l'examen du dessin montre que, suivant cette interprétation, chaque marche mesurerait 40 à 80 centimètres de largeur. Or, dans un escalier normal, le giron se développe sur une trentaine de centimètres. De fait, compte tenu de l'épaisseur des lignes dans ce plan, il n'a pas été possible de tracer toutes les marches mais environ deux sur trois seulement. Autrement dit, si le périmètre de chaque escalier constitue bien une information topographique, son occupation par des pseudos emmarchements est un signe conventionnel à géométrie variable. Ce signe est très suggestif et il permet effectivement de reconnaître les escaliers, le mimétisme a cependant ses limites.

Pour leur part, les encyclies – lignes parallèles aux rivages qui se dégradent en allant vers le large – sont constituées de tracés semblables à ceux qui figurent les détails de la topographie. On les identifie comme encyclies par leur continuité, par le fait qu'elles ne croisent pas d'autres signes (sinon des écritures) et par l'ensemble qu'elles constituent. Ainsi, une ligne d'encyclie isolée de son contexte ne serait pas interprétable (figure 8).

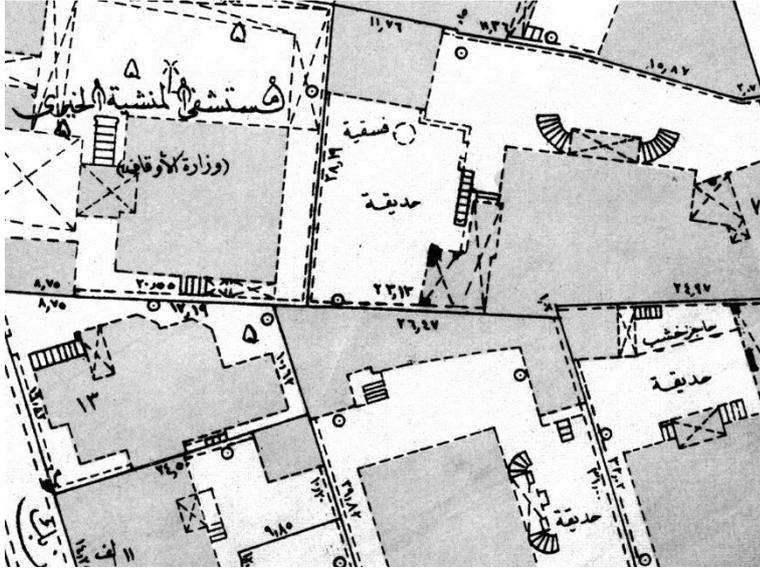


Figure 7. Signe conventionnel et topographie

L'emprise des escaliers est topographique mais la figuration des emmarchements est un signe conventionnel à géométrie variable. Extrait du plan cadastral du Caire au 1:500, feuille 217, 1937. Collection Jean-Luc Arnaud.



Figure 8. Les encyclies et leur environnement

Encyclies et régions de marais dans la carte d'état-major. Extrait de la feuille n° 152, *La Rochelle*, 1:80 000, Paris, Dépôt de la Guerre, 1850. Document IGN.

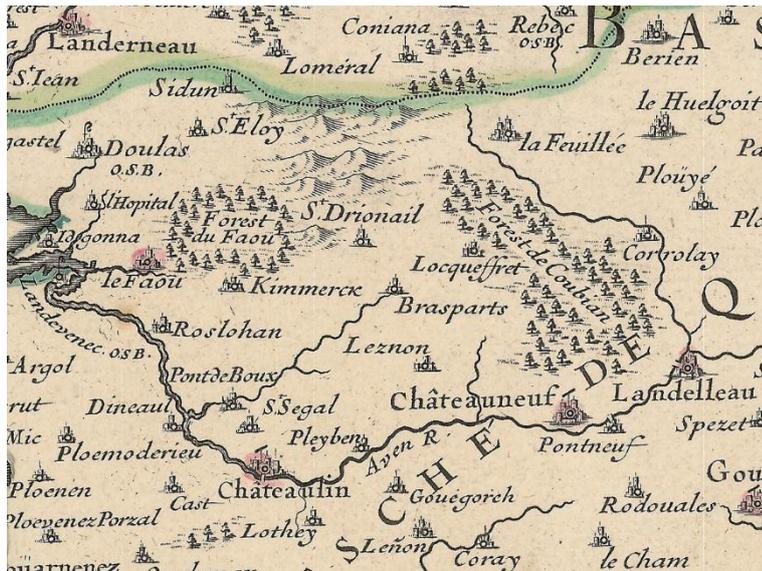


Figure 9. Hiérarchie / hiatus

Le poncif est interrompu pour laisser place au nom de la forêt. Extrait de : Jaillot, Hubert, *La Bretagne divisée en ses neuf évêchés...*, env. 1:500 000, Paris, s.n., 1706. Document IGN.

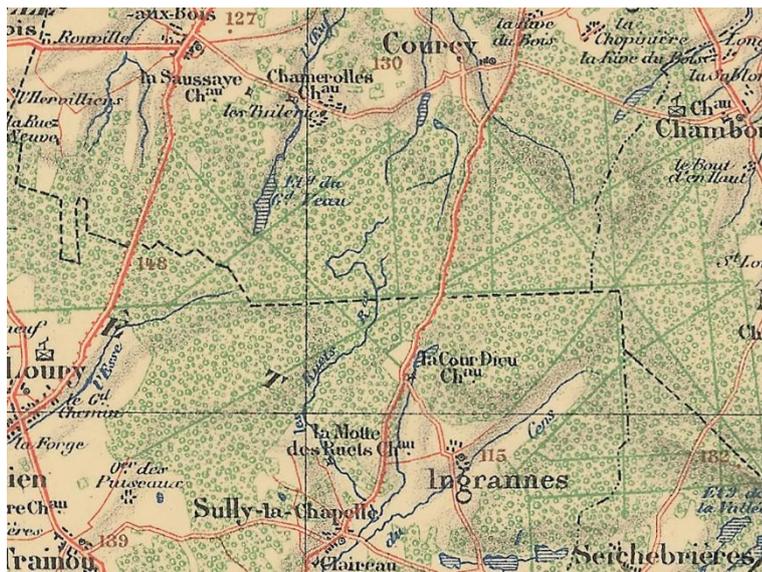


Figure 10. Poncif à géométrie variable

La forêt, traversée de routes, de chemins et de limites administratives est figurée par un poncif à géométrie variable. Extrait de « Département du Loiret », 1:200 000, Orléans, H. Herluison, 1899. Document IGN.



Figure 11. La forêt sur un terrain mouvementé
Un autre mode de représentation, en noir, pour figurer la forêt. Extrait de *Carte des chasses du roi, à Fontainebleau*. 1:34 560, Paris, Dépôt de la guerre, s.d. [1815]. Document IGN.

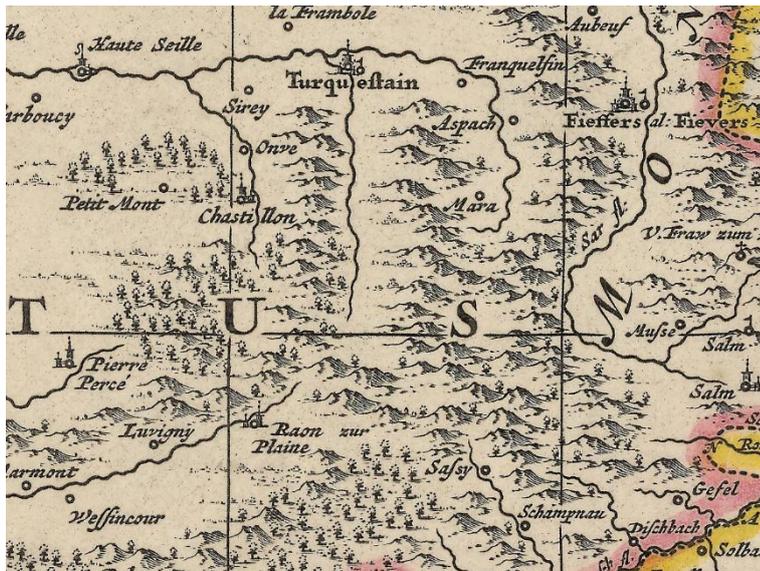


Figure 12. Un signe « montagne » étalé sans correspondance avec le relief
Le poncif qui figure le relief est étalé sur la carte de la même manière que celui chargé de représenter la forêt. Extrait de *Visscher, Nicolaum, Landgraviatus Alsatiae inferioris*, env. 1:280 000, s.l.n.d. [première moitié du XVIIIe siècle]. Document IGN.



Figure 15. Une seule voie de circulation, deux figurations

Le signe qui figure la route en milieu naturel (double trait noir) est modifié dans la traversée des zones urbanisées (zone blanche en creux). Extrait de : *Environs de Nice*, 1:50 000, Paris, IGN, 1964. Document IGN.

Changement de milieu

L'organisation et l'occupation de l'espace ne sont pas isotropes, c'est pour cette raison qu'on en fait des cartes. Il est au contraire composé par l'assemblage de milieux différents. Suivant la nature et les spécificités de chaque milieu, les voies de communication qui les traversent peuvent présenter des particularités dimensionnelles ou constructives. Ainsi, par exemple, lorsqu'une route entre en milieu urbain, elle devient une rue ; lorsqu'une voie ferrée traverse des régions accidentées, elle est posée sur des remblais, des ponts ou encore à l'intérieur de tunnels. Il ne s'agit pas seulement d'un changement de vocabulaire, ces lieux présentent aussi des caractéristiques morphologiques particulières. En ville, la largeur des voies diminue, les fossés disparaissent au profit de drains souterrains tandis que les trottoirs se substituent aux bordures herbues ou gravillonnées, parfois plantées d'arbres. Au-delà de la voie proprement dite, la route est bordée de haies, de champs ou bien de bois, la rue est délimitée par des murs ou bien des constructions. En fonction du milieu traversé ou bien de particularités morphologiques, la figuration des voies peut être soumise à variation (figure 15). Qu'elle devienne une rue en traversant une ville, qu'elle soit bordée d'arbres ou d'une ligne de tramway, qu'elle soit posée sur un remblai, sur un pont, le long d'un mur de soutènement ou bien sous un tunnel, une route constitue une unité spatiale unique dans le cadre d'une étude du réseau de circulation alors que son morcellement en sections de qualités différentes est susceptible d'intéresser d'autres travaux.

Or, si la restitution d'une route tronçon par tronçon semble assez aisément envisageable, il reste dans un second temps à raccorder les sections pour construire l'unité spatiale « route ». On rencontrera les mêmes difficultés avec les voies ferrées et, dans une moindre mesure, les canaux (en remblais, en déblais, sous des tunnels...). Après l'identification des composantes natives de base – susceptibles d'être exploitées suivant des approches à grande échelle –, on cherchera à construire des unités d'autant plus agrégées que les analyses seront plus macroscopiques. En d'autres termes, le processus de restitution de chaque document procède nécessairement par étapes successives. Au cours

de ce processus, il n'est pas de bonnes ou de mauvaises échelles d'observation des phénomènes, c'est en testant et en confrontant plusieurs échelles, en faisant varier les seuils, les modes d'agrégation... qu'apparaissent souvent les résultats les plus féconds (Lepetit 1993, 121).

Contexte

On l'a examiné à travers plusieurs exemples, la compréhension de chaque objet qui compose la carte gagne à être envisagée dans son contexte. Il s'organise suivant plusieurs échelles depuis les éléments valables pour l'ensemble de chaque document (date, échelle, éventuellement thème, pays de production, langue, organisme de production, commanditaire...), jusqu'aux éléments les plus fins (point, lettre, tiret...). Cette approche se révèle fondamentale dans la mesure où elle contribue au décodage de la légende implicite : la liste des mots et des concepts qui qualifient chaque signe conventionnel. Suivant ce processus d'investigation, chaque échelle d'observation apporte son lot d'informations. La prise en compte du contexte est d'autant plus importante que la carte compte moins de couleurs différentes. Cependant, les documents les plus polychromes présentent aussi des ambiguïtés que seule l'analyse du contexte permet de lever.

L'unité la plus élémentaire, le point, peut correspondre au signe diacritique d'un i ou d'un j, au point final d'une abréviation, à un arbre, à un point coté, à un élément d'une ligne en pointillés ou mixte, à un élément de poncif... Pour affecter un point à une de ces catégories, il est strictement nécessaire d'examiner l'environnement dans lequel il s'inscrit. S'il est voisin d'écritures, sa position par rapport aux lettres et à la ligne d'assise du texte permet de caractériser sa valeur ; s'il est voisin d'une valeur chiffrée (dans des limites à encadrer en fonction des unités de mesure en vigueur), c'est sans doute un point coté ; s'il est aligné avec d'autres points de même valeur, il peut correspondre soit à un élément de pointillé, soit à un arbre d'alignement, l'espacement entre les points et le contexte (plus large) dans lequel s'inscrivent les lignes en pointillés aidant à trancher entre ces deux interprétations ; s'il est répété à intervalles réguliers, c'est sans doute un élément de poncif ; enfin, s'il n'est voisin d'aucune indication ni rattachable à un ensemble, ce n'est pas un arbre isolé, ni une tour, ni une cheminée (en position isolée, ces signes sont toujours qualifiés par une mention écrite), c'est peut-être un défaut du papier, une tache d'encre ou encore une trace de moisissure... Les exemples de cette dernière catégorie sont abondants.

Pour les unités linéaires, leur affectation à une catégorie est également facilitée par l'examen du contexte. Cependant, il est souvent moins proche que celui exploité pour les points. Par exemple, les mots *gare* ou *station* situés à proximité d'un rectangle, lui-même placé le long d'une ligne permettent d'affecter cette ligne à la catégorie *voie ferrée*. L'information est ponctuelle mais par la continuité de la ligne et de ses embranchements, elle est « extensible ». Une seule gare peut permettre d'identifier l'ensemble d'un réseau. Il reste à suivre la ligne, cette opération est facilitée par sa configuration générale suivant un examen à échelle moyenne puisque les rayons de courbure des lignes de chemin de fer ne sont pas inférieurs à 300 mètres, ce qui n'est jamais le cas pour les routes ou les rivières. Les embranchements en biseau constituent aussi une des caractéristiques des voies ferrées (figure 16).

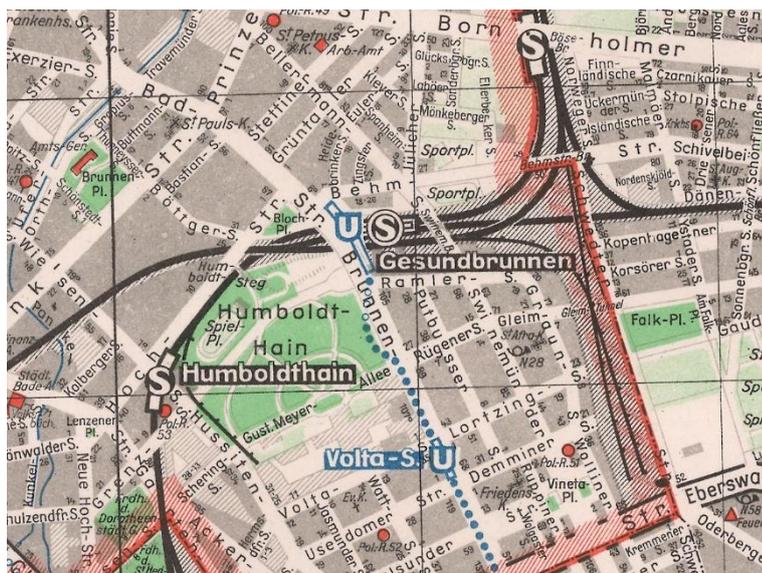


Figure 16. La géométrie comme indicateur
 Les embranchements en biseau constituent une des caractéristiques des voies ferrées. Extrait de *Berlin and Vicinity*, feuille 1, 1:25 000, Washington, Army Map Service, 1953. Document IGN.

A travers cet exemple, la notion de contexte se complexifie, elle appelle à multiplier les échelles d’observation pour trancher entre plusieurs hypothèses et à rapprocher des éléments de catégories différentes (voie ferrée et gare) dans un même ensemble. Les mots « écluse », « port », « bassin »... constituent des indicateurs exploitables suivant le même protocole.

A Angoulême, le long du boulevard Desaix, le plan de la ville publié en 1907 présente la particularité de figurer, dans un lacs de voies de circulations en biseau, des îlots linéaires occupés par des hachures plus denses que celles chargées de représenter les autres îlots (figure 17). Le sens de ces hachures s’éclaire en les confrontant avec la figuration du quartier situé immédiatement à l’est où on trouve le même symbole dont la valeur se dégrade vers l’est. C’est en mobilisant sa culture cartographique que l’opérateur identifie une représentation de la pente du terrain – plus les hachures sont courtes et serrées, plus la pente est forte. Comme en témoigne l’entrée de la ligne de chemin de fer sous un tunnel et la figuration de la trémie précédant cette entrée, le terrain s’élève vers l’ouest. La ville proprement dite, celle des îlots en hachures légères, est située sur un plateau et bordée par le boulevard Desaix tandis que le faubourg Saint-Martin est en contrebas. Les îlots hachurés en question figurent les talus qui permettent aux rampes de passer de la rue Basse au boulevard Desaix.

De nombreuses figurations cartographiques portent des petites croix qui ne sont pas toujours aisées à interpréter. Si elles sont groupées dans un périmètre donné, il s’agit sans doute d’un poncif (signe cimetière), si elles sont plutôt dispersées, elles peuvent correspondre à un signe ponctuel (calvaire, par exemple) et, dans ce cas, il est peu probable qu’un calvaire soit placé au milieu d’une prairie ou bien sur la toiture d’un immeuble. Les croix dans cette situation sont susceptibles de figurer un graticule. Il reste cependant à vérifier que leur répartition dans le champ de la représentation présente une régularité géométrique (pas nécessairement un quadrillage). Dans ce cas, on n’attend pas de la vectorisation une reproduction des croix mais une restitution du graticule et de son système de coordonnées, dont les valeurs sont le plus souvent indiquées en marge de la figuration

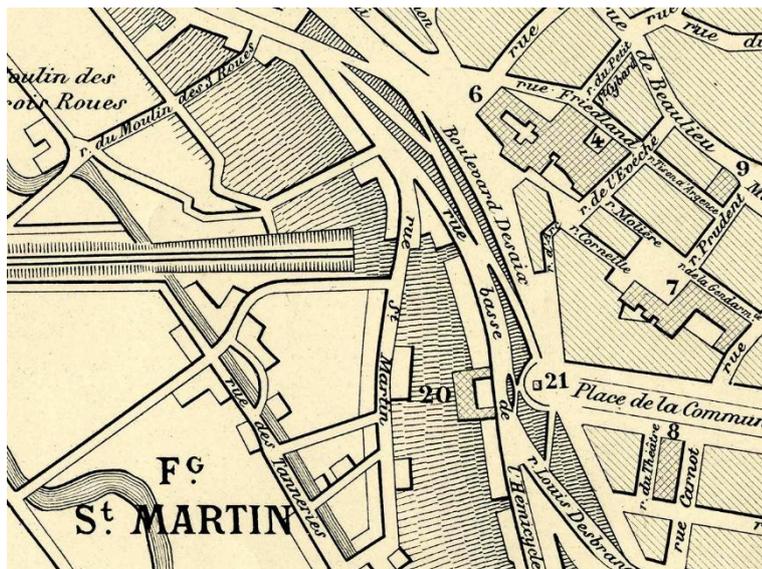


Figure 17. Hachures dans un plan d'Angoulême.

Légende. A Angoulême, le long du boulevard Desaix, le plan présente la particularité de figurer, dans un ensemble de voies de circulations en biseau, des îlots linéaires occupés par des hachures plus denses que celles chargées de représenter les autres îlots. Elles figurent la déclivité du terrain Extrait de *Plan d'Angoulême*, s.l. n.d. [1907]. Document IGN.

cartographique proprement dite. Une telle restitution peut se révéler cruciale pour procéder au géoréférencement du document.

Le contexte peut être plus générique dans la mesure où des règles qui orientent l'identification des signes sont applicables à la plus grande part des documents cartographiques. Ainsi les courbes de niveau ne se croisent jamais. Les voies ferrées suivent des courbes très larges et leurs embranchements sont toujours en biseaux. Les voies d'eau suivent toujours le même sens par rapport à la pente du terrain. La confrontation de chaque tracé avec les informations relatives au relief permet d'exclure des hypothèses d'affectation à certaines catégories. Les unités de propriété sont figurées par des périmètres fermés. Hors des unités de propriété, l'espace public est au contraire ouvert et continu. Dans le plan d'une ville, cette propriété permet de distinguer une section de rue d'une cour longue et étroite. Plus généralement, une ligne qui comporterait l'aspect d'une voie de circulation ne peut pas être considérée comme telle si elle n'est pas en relation avec d'autres voies de circulation. De telles indications, géométriques et topologiques, facilitent l'identification des catégories d'objets figurés et s'avèrent souvent très opératoires pour les documents imprimés en une seule couleur. Elles sont cependant délicates à formuler. Ainsi, les voies ferrées destinées aux tramways (voies normales mais à gabarit réduit) ont des rayons de courbure qui peuvent descendre à 50 mètres et, dans les installations industrielles les tables tournantes donnent lieu à des carrefours de voies ferrées à angle droit ; pour sa part, le continuum qui caractérise l'espace public peut être interrompu dans sa figuration par des lignes qui représentent des passages sous voûtes, des emmarchements ou encore des murs de soutènement. Enfin, le contexte est aussi constitué par des informations consignées en dehors de la représentation cartographique, voire hors du document lui-même¹⁰. Elles sont multiples mais deux variables se révèlent fondamentales : la date de production ou de révision du document et son échelle de réduction.

¹⁰ . Des notices d'accompagnement, comptes rendus de réunions, rapports d'activités...

A chaque échelle correspond une liste des catégories qu'il est envisageable de figurer. De manière générale, plus l'échelle est petite, plus les informations sont généralisées, simplifiées et symbolisées ; de fait, plus le spectre des catégories représentées est étroit. Ce n'est pas un hasard si la plus grande part des planisphères ne porte pas de légende. La culture cartographique la plus rudimentaire est suffisante pour interpréter la répartition des lignes et des couleurs – le bleu étant réservé aux espaces maritimes, bien entendu. Par la réduction de l'échelle, on passe de plans topographiques à des cartes géographiques dans lesquelles l'emprise de chaque signe est sans commune mesure avec la zone géographique effectivement occupée par l'objet correspondant. Les cartes thématiques ou bien celles dont l'utilisation est très circonstanciée – les plans cadastraux par exemple – brouillent cette progression ; elle est cependant très opératoire en ce sens qu'elle permet de restreindre les hypothèses d'interprétation. Ainsi, par exemple, les arbres d'alignement, très fréquents aux grandes échelles, disparaissent au-delà du 1:80 000.

La période de production de chaque document constitue aussi une variable déterminante. Les voies ferrées n'existent pas avant 1830 tandis que la catégorie éolienne, apparue sur les cartes de France au début du XXe siècle, tombe en désuétude à la fin des années cinquante et reprend du service, avec un nouveau signe, au début des années deux mil. Ces quelques exemples montrent que l'examen du contexte – suivant un spectre très large – constitue une étape fondamentale dans la préparation du processus de vectorisation. *De facto*, une part importante du contexte n'est pas un *donné* mais le résultat d'un processus hypothético-déductif qui se construit au fur et à mesure de l'investigation. Ainsi, en première approche on pourrait imaginer de caractériser chaque document à traiter à travers un processus d'apprentissage supervisé. Pour les documents en noir seul, plus complexes à interpréter, l'avancement de leur compréhension procède de proche en proche. Sachant que chaque signe correspond à une seule catégorie et que chaque catégorie est figurée par un seul signe, chaque identification constitue une étape qui facilite les suivantes.

Vers une multiplication des échelles d'observation : déconstruire pour mieux catégoriser

Modéliser les connaissances, c'est tout d'abord saisir la manière dont elles ont été construites. Les cartes ne sont jamais purement topographiques ou bien géographiques, elles comportent toutes des aspects thématiques dont la figuration résulte du code graphique adopté par le rédacteur en fonction de l'objectif assigné au document par son commanditaire. Ces objectifs déterminent une part importante des spécifications telles que le mode de projection, l'échelle de réduction, la catégorisation des données, la distribution des couleurs, la légende, la liste des mentions écrites... En ce sens, la modélisation d'une carte – à travers sa vectorisation – suppose, en amont de disposer d'une bonne connaissance de son contexte et de son processus de production. Les historiens nomment cette opération : critique des sources. Dans la mesure où les représentations cartographiques sont toutes à la fois partielles et partiales, cette analyse se révèle cruciale quelle que soit la période de publication considérée.

Chaque carte résulte de l'association sur un même support de multiples informations de diverses natures. C'est justement cette association qui confère tout son intérêt au document cartographique. Comme nous l'avons examiné, vectoriser une carte n'est pas la recopier, pour donner un sens à chaque objet – lui affecter une catégorie – l'approche doit être multi-scalaire ; c'est en faisant varier l'échelle d'observation que les

points, les lignes et les surfaces imprimées sur la carte deviennent intelligibles. Non seulement, chaque échelle successive d'examen apporte son lot d'informations mais encore, le changement d'échelle lui-même constitue un des meilleurs moyens pour lever une grande part des ambiguïtés. Ainsi, le protocole d'examen le plus efficace ne consiste pas à multiplier les échelles d'observation mais à les faire varier – vers le haut et vers le bas – suivant un processus d'aller-retour systématique. Chaque hypothèse de catégorisation formulée à une échelle donnée gagne à être vérifiée aux échelles supérieures et inférieures.

Ces données et leurs relations inter-scalaires pourraient être acquises à travers des traitements supervisés par des cartographes et enregistrées dans une ontologie, enrichie à la faveur de la multiplication des documents traités. Elle serait d'autant plus efficace qu'elle rendrait compte des relations entre les différentes échelles d'observation. Une telle ontologie constituerait une alternative à l'utilisation de *vérité-terrains*. Cette ontologie établirait également des relations de type *traversant / traversé* entre les catégories pour faciliter la compréhension des discontinuités graphiques (et donc leur restitution). Suivant ces relations, elle offrirait la possibilité de regrouper différentes sections – d'une ligne ou d'une surface de même catégorie – dans des unités cohérentes. Lorsque la séparation entre plusieurs éléments de la même catégorie correspond à l'emprise d'un autre élément (voie de communication le plus souvent), on peut raisonnablement les rassembler dans un même périmètre, ou, pour le moins, dans un même ensemble. On l'a examiné, le regroupement des sections d'un élément dans une unité spatiale unique n'est pas toujours souhaitable. Par contre, leur identification comme des composantes d'un même ensemble, voire d'un même sous-ensemble, faciliterait le développement d'un processus d'investigation de proche en proche. Sachant que les premières affectations catégorielles ont pour effet de réduire le spectre des hypothèses envisageables pour les éléments non affectés, chaque étape de traitement supprime du « bruit » et d'éventuelles ambiguïtés. Par exemple, le fait d'identifier et d'isoler les voies de chemin de fer, facilite le repérage des discontinuités graphiques entre les éléments qu'elles traversent.

Cette approche multi-scalaire constitue une piste prometteuse. Elle suggère qu'il est envisageable de remonter dans le temps du processus de production de chaque document cartographique et d'en restituer la catégorisation. Une large part des outils nécessaires pour obtenir un tel résultat reste à élaborer. Il semble cependant que la modélisation des connaissances rassemblées par chaque document cartographique ne puisse pas faire l'économie de cette forme de déconstruction.

Références

ARNAUD, Jean-Luc, 2022. *La carte de France – Histoire et techniques*, Marseille, Parenthèses.

Automatic Vectorisation of Historical Maps. International workshop organized by the ICA Commission on Cartographic Heritage into the Digital, Budapest, 13 March, 2020.

CHAPPART Gérard et REYNARD, Nicolas, 2007. « La carte topographique française de 1887 à nos jours », *Le monde des cartes*, n° 191, mars 2007, p. 53-67.

- CHEN, Yizi., *et al.*, 2021. « Vectorization of historical maps using deep edge filtering and closed shape extraction » dans *Document Analysis and Recognition – ICDAR 2021*, Lausanne, p. 510-525. hal-03256073.
- CHIANG Yao-Yi., LEIK, Stephan, et KNOBLOCK, Craig A., 2014. « A Survey of Digital Map Processing Techniques », *ACM Computing Surveys*, vol. 47- 1.
- COUZINET, Monique, 1947. *Etude comparative des signes conventionnels de la carte française au 20 000^e et de la carte allemande au 25.000^e*, Paris, Ministère des Travaux publics et des Transports.
- COUZINET, Monique, 1948. *Etude comparative des règles et des signes conventionnels de la carte de France au 20.000^e et de la carte d'Italie au 25.000^e*, Paris, Ministère des Travaux publics et des Transports.
- DAINVILLE, François De, 1964. *Le langage des géographes*, Paris, Picard.
- LEPETIT, Bernard, 1993. « Passé, présent et avenir des modèles urbains d'auto-organisation », dans Bernard Lepetit et Denise Pumain, *Temporalités urbaines*, Paris, Anthropos, p. 113-134.
- Mémorial topographique et militaire rédigé au dépôt général de la Guerre*, Paris, Imprimerie de la République, an XI.
- Rapport du service géographique de l'armée pour 1914-1919*, Paris, Service géographique de l'armée, 1924.
- Tableau commenté des signes conventionnels utilisés pour les levés au 1/10 000 et au 1/20 000*, Paris, Institut géographique national, 1960,
- VALLONGUE, général, 1826, « Langage topographique », dans *Mémorial du Dépôt général de la Guerre*, tome III, Paris, C. Picquet, p. 142-165.

Jean-Luc Arnaud, laboratoire Telemme, UMR 7303 CNRS, université d'Aix-Marseille, Aix-en-Provence, France

Jean-Luc Arnaud, directeur de recherches au CNRS, est historien contemporanéiste, spécialiste d'histoire de la cartographie.

jean-luc.arnaud@univ-amu.fr – Orcid n° 0000-0002-5040-2169

Romain Suarez, Maison méditerranéenne des sciences de l'homme, USR 3125, université d'Aix-Marseille, Aix-en-Provence, France

Romain Suarez, ingénieur d'étude à l'université d'Aix-Marseille, est géomaticien.

romain.suarez@univ-amu.fr

Mots-clés

cartographie, numérisation, catégorisation, histoire, échelle, map, digitisation, categorization, history, scale