

Au commencement était la BDD (des *Parisii*)

discussion de projet

- Thèse : les *Parisii*, dynamiques territoriales et identité(s) culturelle(s)
- thématiques de l'atelier concernées : 1 - méthodes et outils d'enregistrement et de traitement des données de terrain ; 4 - exploration multidimensionnelle de sources croisées pour l'archéologie ;

résumé de présentation

Carole Quatrelivre

Dans le cadre d'un projet doctoral portant sur la formation de la cité des *Parisii* (entendue comme entité territoriale et culturelle indépendante ... ou non) entre le IIIe et le Ier s. av. J.-C., plusieurs problématiques ont pu émerger : l'évolution de l'occupation des sols et des pratiques économiques, la question des réseaux et de l'intégration de la région dans le contexte gaulois global, et l'épineux problème de la caractérisation culturelle. Alors, la nécessité de renouveler et synthétiser les connaissances franciliennes pour le Second âge du Fer s'est imposée. Le croisement de sources étant au cœur du projet, seront prises en compte des approches traditionnelles (toponymie, numismatique) et des méthodes propres aux analyses spatiales. Ce projet a déjà fait l'objet d'une présentation aux Journées Archéologiques d'Île-de-France en novembre 2018 (Marion et Quatrelivre à paraître)

Pour ce faire, une base de données relationnelle a été créée (la célèbre Carobase) autour de plusieurs axes permettant l'évaluation de la documentation et l'enregistrement des données (géographiques, et archéologiques) à l'échelle de chaque site, tout en respectant le découpage en phases (voir le modèle conceptuel, figure 1). Le premier objectif de cet exposé serait donc de soumettre ce travail principalement autodidacte à une critique constructive, afin d'en optimiser la structuration tout en tenant compte des particularités du logiciel utilisé (...Filemaker).

Enfin, la quantité de données traitées, issues d'un dépouillement minutieux des rapports d'opérations franciliennes, permet d'envisager le recours à certains traitements statistiques pour répondre à certaines questions (hiérarchisation de l'habitat par exemple). Un certain nombre de dispositions ont déjà été prises en ce sens dans la construction de la BDD, mais la présentation-discussion aura également pour objectif l'évaluation de leur pertinence en fonction des types d'analyses envisagées (AFC, ACM) et pourront faire l'objet d'un réajustement.

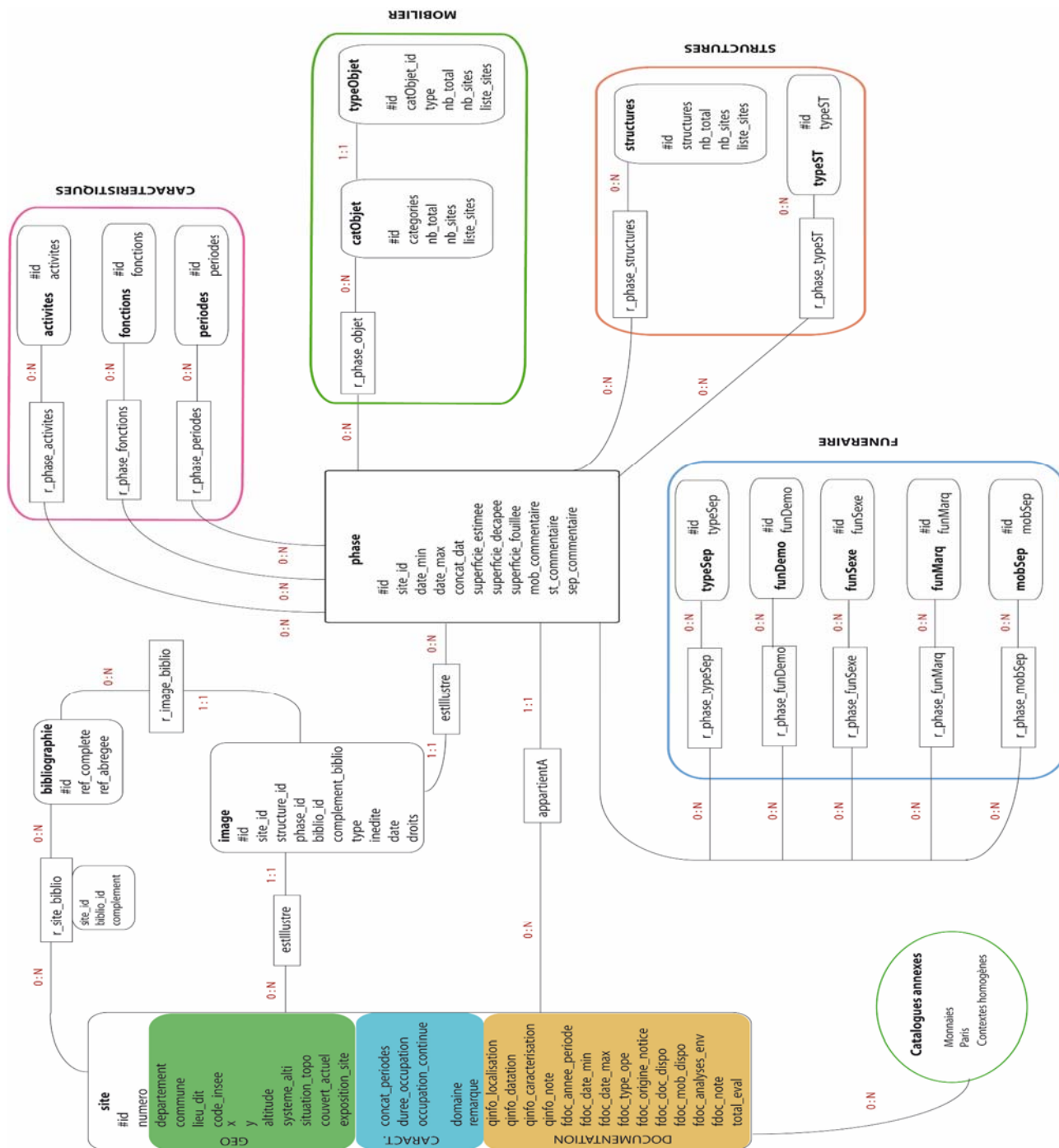


figure 1 : le modèle conceptuel de la base de données actuelle

Discussion

Gaëlle Bruley-Chabot (GBC) ; François Capron (FC) ; Vincent Dargery (VD) ; Robert Demaille (RD) ; Bruno Desachy (BD) ; Rachid El-Hajaoui (REH) ; Julie Gravier (JG) ; Léa Hermenault (LH) ; Carole Quatrelivre (CQ) ; Emilie Roux-Capron (ERC) ; Christophe Tufféry (CT).

A propos de la fiabilité de la documentation et la qualité de l'information.

- RD : Pourquoi ne pas parler de « précision de l'info » à la place de « qualité de l'information » ?
- CQ : Oui, je pourrais.

- JG : une référence utile pour discuter le vocabulaire de l'imperfection des données archéologiques pourrait être celle de la thèse de Cyril de Runz (De Runz 2008) (il propose notamment une typologie de l'imperfection, comprenant les notions d'incertitude, d'imprécision et d'incomplétude). L'intérêt est de partir d'une sorte de référentiel – sa typologie a par exemple été reprise par Johanna Fusco dans sa propre thèse (Fusco 2016) – et de voir comment est-ce que tu peux te l'approprier et, possiblement, la faire évoluer par rapport à tes questions et données.
- CT : il faut se fabriquer son propre dictionnaire pour ce type de question. Il faut établir ses propres critères d'évaluation, et ne pas se baser sur ceux utilisés par les différents auteurs des rapports par exemple, et aussi sortir de la logique booléenne pour ne pas raisonner en binaire (exemple : « fiable », « pas fiable »)
- BD : dans cette veine de l'évaluation de la qualité de la documentation, autre référence par rapport à laquelle tu peux te situer : les travaux de Martin Carver, qui prône un enregistrement « réflexif » dès le terrain, c'est-à-dire incluant des éléments permettant d'apprécier la qualité des données (par exemple, noter les conditions d'observation, noter que plusieurs interprétations sont possibles quand c'est le cas...) (Carver 2009, 2005). Concernant la gestion de l'incertitude et la sortie de la logique binaire dont parle Christophe, il y a les approches qui se rapportent de près ou de loin à la logique des sous-ensembles flous (*fuzzy logic*) qu'utilise C. de Runz notamment, c'est-à-dire basées sur des valeurs de vérité attribuées par le chercheur et variables entre 0 et 1 ; les premiers exemples pour l'archéologie sont les essais, toujours intéressants bien qu'un peu tombés dans l'oubli, d'application informatique de systèmes expert, dans les années 1980, par Jean-Claude Gardin et son équipe (Gardin et al. 1987). Clara (Filet) aussi, par exemple, procède ainsi pour sa thèse.
- CQ : il y a aussi la question de la fiabilité spatiale à prendre en considération, comme le fait par exemple Célia Basset dans sa thèse sur l'occupation du bassin versant de la Seine en aval de Paris.
- JG (post-réunion) : pour ces questions, Célia s'est d'ailleurs inspirée des travaux développés dans le projet *Archaedyn* et de la méthode de la « carte de confiance » ; voir en particulier (Nuninger et al. 2012)

A propos des phases d'occupation dans chacun des sites.

- RD : comment définir une phase ?
- CQ : c'est un ensemble homogène de structures.
- FC : les changements de fonctions ont-ils une influence sur le changement de phase ?
- GBC : est-ce que tu reprends ce qui a été écrit par le RO ou bien fais-tu ton propre découpage en phase ?
- CQ : je reprends les analyses du RO. Je ne réinterprète pas les données.
- RD : mais comment comparer des phases entre plusieurs sites si les RO n'utilisent pas les mêmes définitions de structures par exemple ?
- CQ : il y a quand même un travail de mise en cohérence, surtout pour les gros sites.
- ERC : des études ont-elles déjà été faites sur la typo-chrono dans la région ?
- JG : tu pourrais penser la saisie et l'interprétation des données en deux temps : dans un premier temps tu saisis les données de manière « brute » (tu reprends les phases proposées dans les rapports par les différents RO) et dans un second temps tu interprètes les données (en ayant les phases que tu as toi-même construites). L'intérêt est à mon avis double : 1) tu pourras comparer les différences entre « phases des RO » et « phases revues par toi » (traitements en aval) ; surtout 2) le fait d'avoir les phases construites par toi-même te permet d'harmoniser les données de ta BDD.
- BD (avec compléments post réunion) : cette question de la distinction des « phases des RO » et des « phases construites par le chercheur » avait déjà été discutée dans une réunion précédente lorsque Camille Roulot avait présenté sa BDD sur l'étude des phases d'occupation des sites du haut Moyen Age en Pays de France. Voir notamment le CR 2016-2017, pp. 29 à 32 : http://sitrada.hypotheses.org/files/2018/04/2017_CR_SITRADA.pdf. Pour la datation de ces phases, on retrouve la question de l'imprécision, ici chronologique, dont il faut tenir compte si l'on veut faire du traitement de données basé sur la chronologie, mais qui reste souvent implicite et non traitée dans les publications archéologiques. Il faut en particulier éviter la confusion entre durée d'un événement ou d'une occupation et intervalle d'imprécision chronologique dans lequel on peut situer cet événement ou cette occupation. Ce sont des questions qui nourrissent des échanges dans Sitrada depuis quelques temps ; voir les comptes rendus des précédentes « saisons » (<https://sitrada.hypotheses.org/comptes-rendus>) qui renvoient à des travaux récents (Desachy 2016, 2018b; Gravier 2015, 2018; Hermenault 2017b, 2017a) ou en cours (la thèse de Clara Filet, ici présente, notamment) ; en pratique, il existe un petit outil expérimental (« Chronophage » sous LibreOffice) qui prend en compte la position dans le temps et la répartition par pas de temps choisi par le fouilleur de données de précision inégale (Desachy 2018a)

A propos de la description des structures mises au jour.

- ERC : attention « oppidum de type Zemplin » est une description très critiquable. Il faut faire attention à ne pas confondre les termes qui décrivent et les termes qui interprètent. Par exemple dans l'expression « habitat fortifié

aristocratique », il y a deux choses : le « fortifié » est là pour décrire la structure, mais le « aristocratique » est une interprétation.

- LH : de la même manière, « parcellaire » n'est pas la description d'une structure excavée mais son interprétation.
- JG : c'est important de distinguer les différentes dimensions sémantiques des choses. Pour cela, il faut multiplier les listes : une liste permet de décrire la fonction d'une structure, une autre permet de décrire sa matérialité et une troisième permet de caractériser les groupes sociaux en jeu par exemple. Cela te fera au final trois variables. Tu pourras ensuite faire de belles analyses de données (type AFC), où tu pourras étudier les associations et les oppositions préférentielles entre les structures.
- ERC : peut-on imaginer utiliser les fonctions du thesaurus du CNAU pour une liste de ce type?
- JG : oui pourquoi pas
- ERC : par ailleurs pour la caractérisation des structures, les travaux de Buchsenschutz commencent à dater un petit peu. Il vaut peut-être mieux se baser sur ce que fait François Malrain
- ERC : il faut sortir de la typologie traditionnelle de l'âge du Fer.
- REH : il faut se méfier de l'opposition entre « excavé » et « construit » également : l'un n'empêche pas l'autre dans certain cas. Par exemple, une cave est une structure excavée mais elle peut être construite.

- CT : est-ce que tu t'es intéressée au modèle OH_FET ? (Objet Historique - Fonction Espace Temps)
- CQ : non
- JG : parmi les références du modèle OH_FET, je conseille de lire la référence liée au colloque de géomatique SAGEO de 2007 qui est (à mon avis) très claire (Rodier et Saligny 2007). Il y a aussi la publication de 2010 dans *Cybergeo* (Rodier et Saligny 2010) (en open access)
- BD le modèle OH_FET repose sur un triplet relationnel définissant une unité formée d'une matérialité rattachée à une fonction et à une plage de temps, obtenue par découpage en sous-ensembles élémentaires du volume et l'espace d'ensembles archéologiques qui ont connu plusieurs fonctions successives et plusieurs remaniements architecturaux (comme un château transformé en prison). C'est une approche quasiment stratigraphique ! Sans aller jusqu'à l'échelle de l'unité stratigraphique, il est plutôt prévu pour de grandes échelles (intra-site, voire édifice en archéologie du bâti) ; je ne suis pas sûr qu'il soit adapté à l'échelle territoriale plus globale.

A propos de la description du mobilier mis au jour site par site.

- GBC : faut-il repartir de la donnée brute pour la céramique ? Est-ce que par exemple, ce n'est pas tout aussi bien de travailler à partir des pourcentages de présence de mobilier pour faire une synthèse (ex : pourcentage de tel ou tel type céramique dans l'ensemble) ? Cela car repartir des données initiales brutes demande un boulot immense.
- JG : peut-être que dans l'optique d'une synthèse régionale, et pour aller dans le sens de GBC (c'est un travail immense !), tu pourrais distinguer l'étude du mobilier de l'étude de l'architecture. Pour le mobilier, qui est relatif à une résolution d'analyse particulièrement fine des sociétés, tu fais un « pré-filtre » dans l'enregistrement des données (ex : comme le dit GBC, tu prends un pourcentage de présence de mobilier). Pour l'architecture, la résolution d'analyse est moins fine, donc là tu pourrais directement enregistrer ce qui est donné dans les rapports. Il me semble que cette procédure d'enregistrement est tout à fait justifiable pour deux raisons (combinées) : 1) les données ne sont pas relatives aux mêmes échelles et il n'est donc pas étonnant d'appliquer un processus de travail différencié selon les échelles d'analyses ; 2) le temps de travail que cela demande de reprendre toutes les données.
- CQ : je ne repars aux données brutes que dans le cadre d'étude de cas spécifiques
- ERC : il faudrait quantifier la présence/absence d'importation et leur proportion dans le corpus global
- GBC : pour ce qui est de la faune il y a des synthèses probablement de déjà faites sur le mobilier faune.
- ERC : je te conseillerai d'ajouter dans la catégorie du matériel un item « clous », car c'est en effet un marqueur architectural intéressant. Cf. les travaux de Pierre Péfau qui travaille sur l'architecture en pan de bois gauloise.
- BD : l'utilisation du mobilier pour caractériser un site dans ton cas est liée aux 4 éléments de ta problématique, que tu as exposée : les pratiques économiques, sociales, funéraires et de production. La question est donc, en quoi les données de mobilier disponibles renseignent ces aspects ? C'est peut-être cela le filtre à appliquer, plutôt que de considérer le mobilier en lui-même ou de chercher à être exhaustif ; concernant les aspects sociaux, en particulier la distinction d'occupants « d'élite » ou privilégiés, je pense à la publication du château de Tours qui est un exemple intéressant d'interaction (et pas de simple juxtaposition) entre études spécialisées de mobilier et discours de l'archéologue « généraliste » sur l'évolution de l'occupation du site ; en effet, on voit dans la publication qu'une commande a été explicitement passée aux spécialistes, pour identifier (si possible) les éléments de mobilier constituant des marqueurs sociaux (Galinié, Husi, et Motteau 2014)

- REH, ERC : la présence localisée de pierres sur des agglomérations gauloises, qui signifie un mode d'aménagement particulier, est par exemple un marqueur social et fonctionnel de certains espaces spécifiques

A propos de la base de données et de la structuration des données.

- CQ : on a parlé des notions de fond pour décrire ou interpréter les données, mais sur la base de données elle-même et son organisation, avez-vous des critiques ?

- BD : le document présentant l'organisation de la base que tu nous a montré (cf. résumé de présentation) est visiblement un vrai modèle conceptuel de données, non dépendant du logiciel utilisé (en l'occurrence Filemaker), et qui précise bien la structure relationnelle de la base ; et ça c'est très bien car cela permettra plus facilement d'opérer des requêtes pour des traitements de données (ou de sauvegarder ou transférer les données dans d'autres systèmes).

- CQ : la discussion que l'on vient d'avoir amène éventuellement à des compléments en termes de champs et de listes de valeurs, mais elle ne remet pas en cause la structure relationnelle ?

- JG, BD, RD : a priori, non. Bien sûr il faudra voir à l'usage.

références citées

https://www.zotero.org/groups/2280984/atelier_sitrada

- Carver, Martin O. H. 2005. « Key ideas in excavation ». In *archaeology: the key concepts*, 106-10.
- . 2009. *Archaeological investigation*. London ; New York: Routledge.
- De Runz, Cyril. 2008. « Imperfection, temps et espace : modélisation, analyse et visualisation dans un SIG archéologique ». Phdthesis, Université de Reims - Champagne Ardenne. <https://tel.archives-ouvertes.fr/tel-00560668/document> .
- Desachy, Bruno. 2016. « From Observed Successions to Quantified Time: Formalizing the Basic Steps of Chronological Reasoning ». *ACTA IMEKO* 5 (2): 4-13. https://doi.org/10.21014/acta_imeko.v5i2.353 .
- . 2018a. « Chronophage – téléchargement et documentation ». Billet. *ArchéoFab - Archéologies du Bassin Parisien* (blog). 2018. <https://abp.hypotheses.org/4284> .
- . 2018b. « Recoudre des lambeaux de temps - notes sur le raisonnement chronologique en archéologie de terrain ». In *Pour une archéologie Indisciplinée : Réflexions croisées autour de Joëlle Burnouf*, édité par Florence Journot, 79-88. Editions Mergoïl. <https://sitrada.hypotheses.org/files/2019/02/Desachy-2018-Recoudre-des-lambeaux-de-temps.pdf> .
- Fusco, Johanna. 2016. « Analyse des dynamiques spatio-temporelles des systèmes de peuplement dans un contexte d'incertitude. Application à l'archéologie spatiale. » Phdthesis, UMR 7300 ESPACE (Université Nice Sophia-Antipolis) ; UMR 7264 CEPAM (Université Nice Sophia-Antipolis). <https://hal.archives-ouvertes.fr/tel-01317630/document> .
- Galinié, Henri, Philippe Husi, et James Motteau. 2014. *Des thermes de l'Est de Caesarodunum au Château de Tours : le site 3*. Tours: FERACF.
- Gardin, Jean-Claude, Olivier Guillaume, Peter Q. Herman, Antoinette Hesnard, Marie-Salomé Lagrange, Monique Renaud, et Elisabeth Zadora-Rio. 1987. *Systèmes experts et sciences humaines: le cas de l'archéologie*. Paris, France: Eyrolles.
- Gravier, Julie. 2015. « Recognize temporalities from formalization of dating urban units: three case studies ». In *CAA2014: 21st century archaeology: concepts, methods and tools: proceedings of the 42nd annual conference on Computer Applications and Quantitative Methods in Archaeology*, p.371-379. Oxford [UK]: Archaeopress. <http://www.archaeopress.com/Public/displayProductDetail.asp?id={E35F9954-5653-493D-884B-4A7D2DE66610}> .
- . 2018. « Deux mille ans d'une ville en système. Proposition d'une démarche appliquée au cas de Noyon ». École doctorale de Géographie de Paris (ED 434) Spécialité: Géographie, Paris: Paris I – Panthéon-Sorbonne.
- Hermenault, Léa. 2017a. « La ville en mouvements. Circulations, échanges commerciaux et matérialité de la ville ». Phdthesis, Université Panthéon-Sorbonne - Paris I. <https://halshs.archives-ouvertes.fr/tel-01551802/document> .
- . 2017b. « Percevoir les rythmes d'évolution du tissu urbain parisien dans la seconde moitié du XVème siècle grâce à l'étude des transactions foncières : le traitement des intervalles d'imprécision des datations. » *Archéologies numériques* 1 (1). <https://www.openscience.fr/Percevoir-les-rythmes-d-evolution-du-tissu-urbain-parisien-dans-la-seconde> .
- Marion, S., Recherches sur l'âge du fer en Île-de-France : entre Hallstatt final et La Tène finale : analyse des sites fouillés, chronologie et société. Oxford, J. Hedges, 2004.

- Marion, Stéphane, et Carole Quatrelivre. à paraître. « De la guerre à la terre : une archéologie du territoire des Parisii ». In *Territoires, sociétés et conflits en Île-de-France, Actes des Journées Archéologiques d'Île-de-France, 23-24 novembre 2018*.
- Nuninger, Laure, Laure Saligny, Krištof Oštir, Nicolas Poirier, Élise Fovet, Cristina Gandini, Estelle Gauthier, Ziga Kokalj, Florian Tolle, et Klemen Zakšek. 2012. « Models and tools for territorial dynamic studies ». In *ArchaeDyn seven millennia of territorial dynamics: Settlement Patterns, Production and Trades from the Neolithic to the Middle Ages*, édité par Gandini C, Favory F, et Nuninger L, 23-38. BAR International Series. Archaeopress.
https://www.researchgate.net/publication/274385789_Models_and_tools_for_territorial_dynamics_studies .
- Rodier, Xavier, et Laure Saligny. 2007. « Modélisation Des Objets Urbains Pour l'étude Des Dynamiques Urbaines Dans La Longue Durée ». In *SAGEO 2007, Rencontres Internationales Géomatique et Territoire, Jun 2007, Clermont-Ferrand, France*, 1-24. Aménagement Du Territoire. AgroParisTech-ENGREF.
https://www.researchgate.net/publication/228621488_Modelisation_des_objets_urbains_pour_l'etude_des_dynamiques_urbaines_dans_la_longue_duree .
- . 2010. « Modélisation des objets historiques selon la fonction, l'espace et le temps pour l'étude des dynamiques urbaines dans la longue durée ». *Cybergeo : European Journal of Geography*, juin.
<https://doi.org/10.4000/cybergeo.23175> .