

Remerciements :

Les auteurs remercient l'ensemble des collègues du programme de recherche MOBI'KIDS (financement ANR 16-CE22-009, 2017- 2021).

Les traces : nouvelles données pour l'étude des mobilités enfantines

Définie comme *“marque laissée par une action quelconque”* (en référence au dictionnaire Larousse) la notion de trace prend aujourd'hui une toute autre dimension avec le numérique. Dans le contexte du développement des données massives (Big data), les traces sont au cœur d'enjeux majeurs tant économiques, sociétaux que scientifiques. En qualité de marqueurs d'activités individuelles, de par leur forme et leur volume, elles constituent un type de donnée très convoité dans l'économie du numérique mais aussi de nouvelles ressources informationnelles qui questionnent largement les pratiques scientifiques (Boullier, 2015 ; Bastin et Tubaro, 2018).

Engagé.e.s dans un programme de recherche portant sur l'étude des mobilités quotidiennes enfantines, nous - auteur.e.s issu.e.s de la psychologie environnementale, sociologie, géographie, géomatique et informatique - proposons d'exposer et de discuter ici comment la notion de “trace” nous réunit pour délimiter conceptuellement et méthodologiquement l'objet de notre recherche avant de le mettre en œuvre empiriquement.

La notion de trace n'est pas nouvelle en sciences humaines et sociales (Ginzburg, 1980, Serres, 2012, Galinon-Méléneq, 2017), mais elle revêt depuis quelques années un caractère particulier dans le champ de la géographie (Gagnol, Mounet et Arpin, 2018 ; Mericskay, Noucher et Roche, 2018) et des mobilités spatiales en particulier. En effet, les objets connectés et les réseaux sociaux sont vecteurs de traces (numériques) qui sont de plus en plus explorées et mobilisées par les chercheurs dans la mesure où elles permettent de saisir et de donner à voir des pratiques spatiales au niveau des individus et à des granularités spatio-temporelles fines (Beaude, 2015). Dans le domaine des sciences de l'information, cette masse de données procure aussi de nombreux défis en termes de stockage, d'interrogations, d'analyse et d'extraction de connaissances (Tampakis et al. 2020). Aussi, ces nouvelles données et méthodes participent à un renouvellement des modalités de collecte, des traitements et analyses des mobilités individuelles et collectives (Depeau et Quesseveur, 2014, Drevon et al., 2014, Vidal, 2015). Les traces (numériques) ont donc un

fort potentiel pour produire des connaissances qui répondent aux exigences de validité écologique des enquêtes de terrain d'une part et aux besoins de reproductibilité des collectes et des analyses d'autre part (volume, continuité dans le temps et dans l'espace, collecte automatisée). Toutefois, penser les mobilités par le seul prisme des traces numériques ne suffit pas pour comprendre ce qui motive et explique les déplacements des personnes en termes de choix des itinéraires, des modes de transport, de l'organisation de la mobilité au sein du ménage et plus encore en termes de sens et de vécu associés à ces pratiques. C'est pourquoi, associer aux traces numériques des informations qui leur affectent du sens s'avère être un enjeu crucial pour produire des connaissances renouvelées sur les pratiques de mobilité.

Le programme Mobi'kids^[1] (MK) s'inscrit dans cette lignée de recherches qui s'attachent à comprendre les mobilités individuelles en mixant différentes méthodes de collecte et d'analyse pour produire des traces de natures différentes et complémentaires. Dans ce contexte, il s'agit de comprendre l'évolution des mobilités quotidiennes d'enfants, en considérant la période de transition dite « normative » (Bronfenbrenner, 1986) qui caractérise l'élargissement du territoire des enfants, lié au passage de l'école primaire au collège d'une part, et en tenant compte de la diversité du contexte socio-spatial d'autre part (Depeau et al., 2017). Plus spécifiquement, deux orientations guident notre recherche pour comprendre l'élargissement du territoire et les processus d'apprentissage des déplacements :

- Une première consiste à étendre la définition d'autonomie de déplacement des enfants en considérant les types de *“réaction face aux formes d'imprévu, de nouveauté, d'absence de familiarité”* (Depeau, 2003). Autrement dit, en plus de décrire des routines (Pradel, Chardonnel et al., 2015), il s'agit ici de rechercher et de comprendre l'impact des formes spontanées voire non programmées des pratiques des enfants en ville dans le processus d'apprentissage de l'autonomie des déplacements.
- La seconde orientation relève des conditions d'apprentissage, en particulier du rôle du contexte social-spatial de l'enfant et de sa famille qui pourrait être en partie expliqué par des *“cultures éducatives urbaines”* différenciées. Il s'agit ici de dépasser les approches « physicalistes », pour comprendre la notion de contexte de façon plus systémique où s'articulent les dimensions sociale (l'entourage, la famille), temporelle (les temps de vie, les horaires des services, des activités), spatiales (son environnement, son accessibilité), pratiques (activités, pratiques sociales) et perceptives (modes attentionnels, ambiances).

Afin de répondre à ces objectifs, la trace représente une notion partagée et structurante tant dans les échanges entre chercheurs que dans la mise en œuvre des collectes et des traitements de données. Et au-delà de matériaux empiriques visant à mieux comprendre les évolutions de mobilités enfantines, le potentiel heuristique des traces est au cœur des questionnements de MK. De simples données “brutes” d’enregistrements issus de capteurs, elles deviennent des informations structurées à partir du moment où on leur affecte du sens grâce à des enquêtes auprès des individus. Elles permettent enfin de produire des connaissances, quand elles sont agrégées, croisées et analysées.

Fil d’Ariane du programme MK, la notion de traces permet de guider l’ensemble des étapes du processus de recherche. Nous proposons dans cet article une présentation et une discussion autour des modalités de fabrication des traces comme objets épistémologiques, méthodologiques et techniques au centre d’une recherche interdisciplinaire.

La trace comme concept dans la littérature et mise en application dans MK est développée dans la première partie afin de définir de façon interdisciplinaire les mobilités et rapports à l’espace des enfants et de leur famille. La trace comme entité observable est ensuite présentée dans une deuxième partie qui explique la fabrication des traces d’un point de vue méthodologique et décrit le corpus original de données. Enfin, la trace comme concept analytique pour l’exploitation des données est discutée dans la dernière partie qui s’attache à détailler la question de la transformation des traces à travers notamment les opérations de catégorisation dans une optique interdisciplinaire.

Le concept de trace dans la littérature et dans la recherche Mobi’Kids

La trace en situation de mobilité en SHS et en Sciences de l’information

La trace demeure une notion récente et de plus en plus mobilisée en SHS. Dans le domaine des sciences du territoire et des sciences des comportements où les méthodes s’articulent à celles des sciences de l’information, la notion de trace ouvre un champ d’investigation important pour l’étude des déplacements (Chen et al., 2016). Qu’il s’agisse des traces relatives aux activités de téléphonie mobile permettant d’étudier de grands échantillons d’individus sur des temporalités longues (Smoreda et al., 2013), des traces de billettique des

réseaux de transports en commun mettant en lumière des évolutions de pratiques dans le temps ou encore des logiques d'intermodalité (Zong et al., 2015; Richer et al., 2018) ou enfin des traces issues des réseaux sociaux (Twitter ou de Foursquare), de plus en plus utilisées pour appréhender sous des formes renouvelées, des dynamiques de déplacement ou de fréquentation, notamment urbaines (Cébeillac, 2018 ; Calabrese, 2013).

Dit autrement, la trace est particulièrement pertinente pour qui s'intéresse aux manières de se déplacer, de pratiquer, de percevoir les espaces de vie et donc d'habiter les espaces et le territoire. Partant, Walter Benjamin remarquait très justement ([1935] 1971) : "*Habiter signifie laisser des traces*" (p.132).

En outre, la double acception de "trace", extraite de son étymologie allemande "Spur", mobilisée par Krämer (2012), faisant référence à ce qui est "nettement perceptible" (en référence au terme *piste* et tout le vocabulaire associé aux normes des voies de déplacement) et à ce qui "est à peine perceptible" permet également d'ouvrir la dimension polysémique des traces, et dans son corollaire, un ensemble d'opérations produites dans l'interdisciplinarité. En effet, dans le premier cas (ce qui est "nettement perceptible"), la notion de trace renvoie à la donnée géonumérique où la dimension d'ordre ou d'enchaînement produisant la ligne (la piste et donc la mobilité) rappelle aussi les processus de mise en sens qu'engage la notion; les plus communs à l'ensemble des données de type numérique étant le "séquençage" au sens de "*déterminer l'ordre linéaire de composants (...)*" (Ertzscheid, 2017, p. 150). La trace, abordée comme une donnée complexe qui mélange les dimensions spatiale et temporelle, est ainsi le résultat d'un processus transformant des données brutes en extraction de connaissances (Bogorny et al. 2014). Au-delà de l'évaluation automatique de la qualité (Ivanovic 2018) et de la prise en charge de capteurs complémentaires pour suivre des objets en indoor et en outdoor (Momose et al. 2017), les traces s'inscrivent dans une segmentation en STOP et MOVE (Yan et al. 2011) et un enrichissement sémantique spécifique à chaque domaine d'application (par exemple, association de points d'intérêt à des STOP, de mode de déplacement à des MOVE ou d'ajout d'activités à ces segments (Bisone et al. 2018). Ces opérations conduisant à des modes de re-présentation scientifique constitutive de la production de connaissances peuvent aller jusqu'à la recherche de structures ou de patterns (Moreau et al. 2018). Dans le second cas (ce qui "est à peine perceptible"), la notion de trace renvoie à la dimension d'empreinte qu'on assignerait au point (spatial et temporel), à la marque (par exemple corporelle) à l'impression (par exemple mémorielle) ou à toute autre forme de pratiques non habituelles (outliers). Ces deux types d'acceptions rapportés au terme allemand rappellent la distinction faite par Steck (2011) entre « *trace flux* » et « *trace-marque* » qui permet de différencier les déplacements et mouvements (ou MOVE) relatifs à des pratiques de mobilité, des arrêts ou

positions (ou STOP) liés à des lieux et des activités.

En outre, dans le champ des travaux sur les mobilités spatiales, le processus de “tracking” ne peut se limiter aux seules traces collectées (position géographique et estampillage temporel). Celles-ci ne peuvent être analysées et comprises qu’à partir des informations sémantiques complémentaires et/ou à partir de données de contexte (Raza et al., 2015). Ces différents types d’informations permettent d’associer à la trace une identité individuelle d’une part, et les caractéristiques des comportements engagés dans l’espace et dans le temps (comme les modes de déplacements, modes d’accompagnement, motifs, etc.) d’autre part. L’ensemble de ces caractéristiques dépend des « données sémantiques » qui enrichissent le sens des traces. Et leur interprétation dépend dès lors d’un ensemble d’opérations scientifiques (extraction, nettoyage, séquençage, enrichissement, agencement, contextualisation) qui convoque des champs d’expertise disciplinaires spécifiques amenés à s’accorder sur les objectifs visés et à composer l’objet trace à partir de multiples dimensions. Ainsi, pour faciliter la manipulation de ces données sémantiques et les croiser avec des sources de données extérieures, il est fréquent d’avoir recours à des ontologies. Dans un premier temps, l’ontologie était uniquement une notion philosophique qui s’intéressait à l’étude de ce qui est, contrairement à la métaphysique. En sciences de l’information, cette notion a été extrapolée. Elle désigne l’organisation des connaissances par des experts pour un domaine d’étude. Dans le cadre de MK, l’utilisation d’une ontologie commune a permis aux membres du projet de s’accorder sur le sens des concepts et leurs liens. Parallèlement, l’extraction de sens, voire de connaissances, peut être définie en sciences de l’information selon diverses méthodes (Renso et al. 2013) : l’extraction de motifs fréquents de trajectoires (pattern mining), la constitution de clusters d’objets mobiles aux comportements identiques ou encore l’identification d’objets aux comportements inhabituels (outlier detection).

La trace est ainsi, comme le précise Jeanneret (2011), une donnée que l’on dote d’un sens particulier (la trace étant par nature à la fois labile et muette). Mais la trace est également comprise comme ce qui, dans ses pratiques quotidiennes, imprègne ou marque l’individu d’un point de vue cognitif, affectif et sensoriel et peut être transcrit dans une forme verbale et comportementale apte à enrichir certaines formes d’expériences en lien avec le contexte et plus spécifiquement les ambiances. Autrement dit, la trace fait référence à l’empreinte laissée en mémoire qui peut être transmise par l’individu sous des mises en forme variées tant narratives, visuelles qu’auditives et qui peuvent alors être retranscrites, assemblées voire recomposées pour une interprétation relationnelle du système des traces. Ainsi, le sens particulier est produit et (re)construit sous le prisme de données complémentaires aptes à documenter les mobilités (Mericskay, 2019) et à recomposer, en contexte

d'écologisation des pratiques individuelles, l'univers et les conditions de production des traces numériques captées.

En ce sens, malgré un réel potentiel analytique, la trace collectée de façon numérique ne peut se substituer à d'autres données collectées plus classiquement (via l'enquête). C'est une donnée qui initie des enrichissements complémentaires voire, en empruntant la formule relative au tracking corporel, « *un croisement de données qui, ensemble, forment un agrégat d'authentification en traces* » (Leleu-Merviel, 2017, p. 79). Dès lors, la mobilité comme fait social (Bassand et Brulhardt, 1980) ne peut uniquement être appréhendée par le traitement de gros volumes de données standardisées indiquant des positions spatiales et temporelles. Ce que montrent Loebach & Gilliland (2016) à partir des données de mobilité enfantines. Les données plus fines ou small data (données qualitatives) doivent permettre de dépasser les limites relevées par Quesnot (2016) en ce qui concerne l'usage des Big Data dans le champ de la géographie et plus spécifiquement de la définition de l'espace social.

La trace dans le projet Mobi'kids

Au sein du projet MK la trace est mobilisée comme cadre opératoire qui permet à la fois de : (1) rassembler et faire dialoguer des chercheurs issus de domaines différents (psychologie, sociologie, géographie, géomatique, informatique) afin de définir de manière interdisciplinaire l'univers du phénomène que l'on étudie ; (2) structurer et articuler un protocole méthodologique d'enquête et d'analyse mixte autour de la collecte et de l'enrichissement de données spatio-temporelles précises, riches et inédites. Au centre des questionnements sur la mobilité des enfants et la recherche de régularités/irrégularités dans les pratiques et les comportements, la trace est mobilisée au sein de toutes les étapes du projet.

En premier lieu, en qualité de concept, la trace ne peut être définie de manière unique tant les approches et les visions des chercheur.e.s impliqué.e.s dans le programme sont diversifiées (cf. vignette 1).

Vignette 1 : Univers sémantique de la trace dans MK

La notion de trace est ici définie par les auteurs membres du collectif des chercheur.e.s de MK. L'univers sémantique de la trace pour ces chercheur.e.s relève de l'imbrication de 6 dimensions principales : l'espace (spatial, espace,

forme d'interdisciplinarité conceptuelle et de création méthodologique partagée qu'il s'agit. La trace joue ici le rôle de traducteur entre des disciplines qui reposent chacune sur un langage et un univers conceptuel différents. Elle contribue à développer un vocabulaire commun, induit un effort de clarification conceptuelle au sein de chaque discipline concernée et rend possible les échanges et les apports réciproques entre les diverses perspectives mobilisées (cf. vignette 2). De plus, ce travail de composition interdisciplinaire permet de mettre à l'épreuve le sens des traces des mobilités enfantines au moyen d'un dispositif multidimensionnel.

Vignette 2 : Production d'un langage commun - Glossaire

Afin de faciliter l'interdisciplinarité et favoriser l'échange entre des approches scientifiques ayant chacune leur propre lexique et cadre conceptuel, un glossaire partagé a été élaboré au cours de la recherche. Composé de plus d'une soixantaine de termes, ce glossaire a permis à la fois de construire ensemble un langage commun minimal permettant l'articulation entre les domaines concernés, de mettre en perspective les diverses cultures scientifiques les unes par rapport aux autres, et de stabiliser l'acception partagée de certaines notions. Par exemple, la mise en catégories des traces enrichies liées aux activités a conduit à clarifier et distinguer les notions de pratiques et comportements.

En second lieu, en tant qu'outil méthodologique la trace est appréhendée comme un observable qui doit aider à comprendre le rôle des formes imprévues des activités spatiales dans les apprentissages et l'élargissement du territoire de vie des enfants d'une part et l'impact des contextes socio-spatiaux d'autre part. Au même titre qu'au niveau conceptuel, la trace comme matériel empirique n'est pas uniforme au sein du projet MK. Ici, les traces comme marqueurs spatio-temporels des pratiques de mobilité infantine renvoient à des formes hétérogènes et des modalités de collecte et d'enrichissement diversifiées selon plusieurs temporalités et différents "points de vue".

La production des traces sur le terrain de Mobi'kids

La production des traces dans le programme MK s'appuie sur une enquête menée entre

2018 et 2019 auprès d'une cohorte de 88 familles, soit 182 personnes à raison d'un enfant et un parent par famille consentants pour participer à l'ensemble du protocole instrumenté avec des capteurs GPS au cours de deux périodes (enfant en école primaire et enfant au collège). Le recrutement s'est fait par l'intermédiaire des écoles primaires (classes de CM1-CM2), dans le but de suivre tous les enfants d'une même classe ayant potentiellement des interactions sociales y compris en dehors de l'école. Le terrain se compose de deux sites d'étude localisés dans la métropole de Rennes : le premier, situé dans le centre de Rennes concerne 3 écoles et 5 collèges, le second localisé dans une zone périurbaine de Rennes couvre Orgères et Saint-Erblon et comprend 2 écoles et 4 collèges.

Un protocole de collecte hybride et itératif

Afin de répondre aux objectifs de recherche, plusieurs dispositifs originaux de collecte et de traitements de traces, combinant des approches et des méthodes de prime abord différentes, ont été développés. Le protocole d'enquête est conçu sous la forme d'une chaîne de collecte qui s'appuie sur trois dispositifs d'enquêtes complémentaires. Un premier consiste à équiper une cohorte de familles (binôme enfant-parent) d'un datalogger (munis de capteurs GPS et accéléromètres) enregistrant en continu pendant cinq jours au minimum l'ensemble de leurs positions géographiques et temporelles (traces brutes). Le deuxième dispositif consiste en un examen exhaustif du séquençage des traces collectées des personnes enquêtées via une tablette (Depeau et al., 2019) qui vient enrichir, *a posteriori*, la première collecte en questionnant les personnes sur les motivations et les conditions de réalisation de leurs déplacements et activités enregistrés ainsi que les rapports aux espaces, à l'autonomie et divers éléments relatifs aux attitudes éducatives. Enfin, une ultime phase de collecte (réalisée avec un sous-échantillon d'enfants ayant réalisé les deux phases précédentes d'enquête) s'attache à observer le vécu des enfants en situation de mobilité, tout en leur demandant de commenter leur expérience.

Cette imbrication de méthodes et de dispositifs permet de constituer un corpus hybride de données quantitatives et qualitatives (cf. Figure 1) ayant trait aux formes et aux conditions de mobilités quotidiennes des enfants.

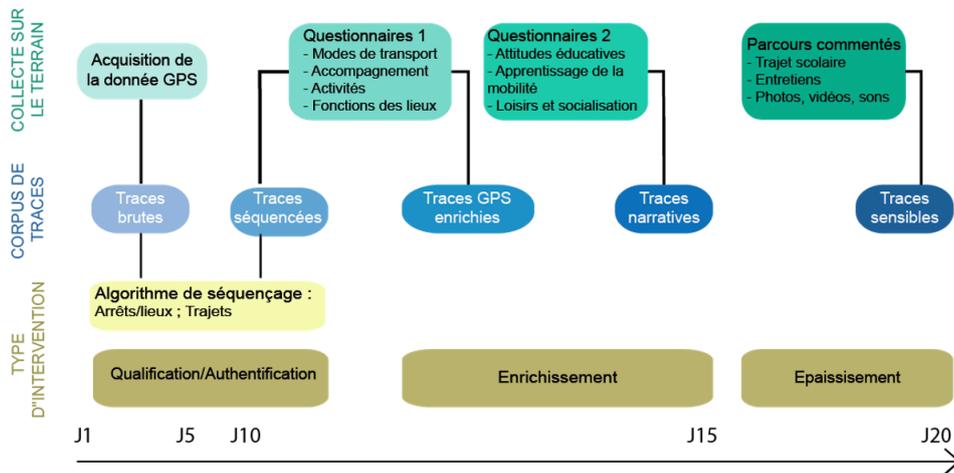


Figure 1: Chaîne de collecte et de traitement des traces

Quatre corpus de traces ont ainsi été collectés :

- un corpus de *traces brutes* qui se matérialise par une suite de positions spatio-temporelles (7,6 millions d'enregistrements sur l'ensemble du corpus).
- un corpus de *traces géolocalisées séquencées*, issues des dataloggers portés par les individus, qui sont ensuite "nettoyées" puis segmentées par un algorithme dédié afin de les transformer en trajectoires spatio-temporelles décrivant une suite de déplacements et de lieux d'activités (où l'individu est "immobile"). La trace est alors formalisée dans une base de données spatio-temporelle, sous la forme de trajets (lignes) et de lieux (points).
- un corpus de *traces enrichies sémantiquement et narratives*, collectées lors des entretiens, qui caractérisent les modes de déplacement et les accompagnements liés aux trajets, les types d'activités attachées aux lieux, et les formes d'apprentissage de mobilité des enfants et les attitudes éducatives des parents. Ces traces sont également accompagnées de notes qui permettent de corriger certaines erreurs et d'apporter plus d'informations sur les pratiques.
- un corpus de *traces sensibles* (traces géolocalisées, photographiques, vidéo et sonores) produites durant des trajets de parcours commentés réalisés *in situ* entre l'école et le domicile,



Figure 2. Représentations spatiales des 4 corpus de traces

Diversité des formes de “captures” et production des traces

Sur le plan des méthodes de recueil de donnée, la boîte à outils du terrain MK (cf. vignette 3) est riche et composite. On peut toutefois distinguer deux principales approches dans le processus de collecte des traces en fonction du degré d’implication des enquêtés et des formes de mobilisation des enquêteurs.

En premier lieu, la collecte “passive” renvoie à l’utilisation de dispositifs techniques dédiés, tels que des dataloggers ou des caméras qui permettent d’enregistrer des positions et des images de manière automatique. Au niveau de l’enquête, ces types de dispositifs sont intéressants pour leur capacité à enregistrer des grands volumes de traces homogènes. Ces formes de collecte, “simples” de prime abord du fait de la faible mobilisation d’enquêteurs, renvoient toutefois à des exigences spécifiques d’ordre technique auxquelles il a fallu répondre. Le recours aux dataloggers supposant un dispositif adapté aux besoins de l’enquête, a par exemple nécessité en amont, en partenariat avec des entreprises spécialisées, un long travail de développements spécifiques et de tests sur la base du dispositif existant (Depeau et al. 2019). Ensuite, sur le plan des données collectées, les questions de stockage (conformes aux règlements de protection des données personnelles) comme d’analyse impliquent un ensemble d’opérations systématiques laborieuses (nettoyage et séquençage des positions GPS ou montage vidéo) pour transformer les données “brutes” en connaissances (Duroudier et al., 2020).

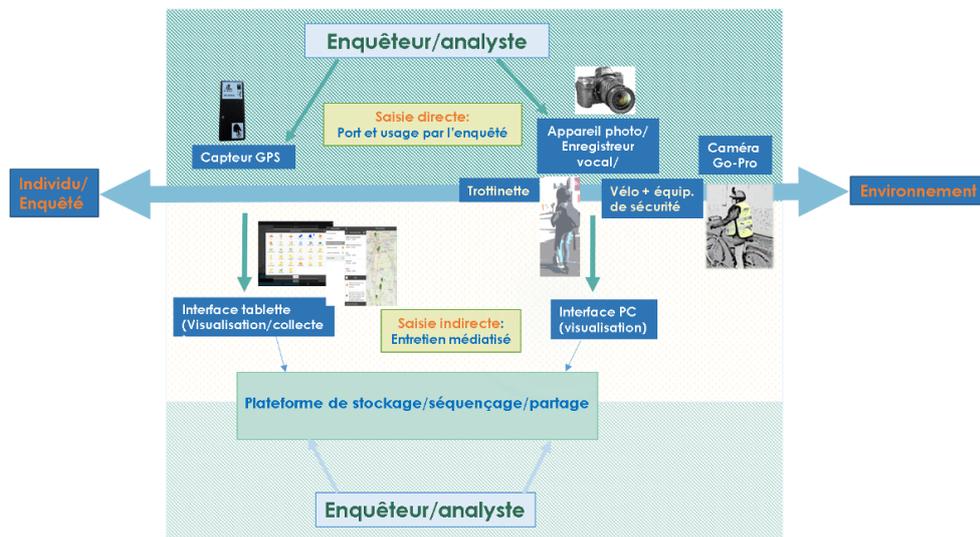
En complément de l’approche “passive”, se situe la collecte dite “active” de données, qui nécessite l’intervention d’enquêteurs pour enrichir les traces d’informations précises. Ce type d’approche, plus classique, s’appuie, dans le cadre de MK, sur des méthodes qualitatives (entretiens, parcours commentés) plus chronophages (tant dans la préparation

que dans la passation) et dont la garantie de qualité de la collecte dépend à la fois du nombre d'enquêteurs mobilisés (souvent important pour ce type d'enquête) et du type de relations enquêteur-enquêté qui se joue selon les formats d'enquête requis.

Cette posture méthodologique mixte permet la collecte et la production itératives d'un ensemble de traces dont la qualité d'information déroge aux principes classiquement adoptés dans le champ des Big Data, à savoir l'obtention (quasi)-automatique (via des plateformes par exemple) de données personnelles ou complémentaires documentant les traces. En effet, dans le cadre de MK, les opérateurs humains (chercheurs, enquêteurs) jouent un rôle important dans les processus de collecte et de transformation des traces brutes en traces sensibles, que ce soit au niveau du paramétrage des pré-traitements (notamment l'algorithme de séquençage) ou dans les phases d'enrichissement discursif (questionnaire tablette et parcours commentés).

Vignette 3 : Le rôle des dispositifs "matériels" (les boîtes à outils du terrain MK et l'organisation du terrain)

Faire porter des capteurs (GPS, Caméra, Appareil Photo, Enregistreur) aux enfants et à leur parent donne une tonalité particulière à la relation enquêté/enquêteur et plus généralement à l'engagement des personnes dans cette étude. D'un côté cela permet de s'affranchir de l'enquêteur lors du suivi sur 5 jours, d'un autre côté cela exige des personnes d'être dans une posture d'expérience instrumentée avec des équipements matériels qui peuvent transformer les gestes quotidiens. La réception du dispositif a été globalement très positive, notamment parce que les personnes ont pu se représenter concrètement leurs données personnelles par l'intermédiaire du dispositif d'enquête sur tablette qui visualise les traces segmentées sur une carte. Les parcours commentés ont aussi fait l'objet de restitutions sous forme de posters personnalisés à destination des enfants. En outre, nous avons pris soin de bien communiquer sur les encadrements légaux qui entourent la collecte et le traitement des données, à savoir le suivi des règles du RGPD et la mise en place d'un Plan de Gestion des Données.



La donnée « trace » associée à « la donnée sémantique » nécessaire pour enrichir le sens et son interprétation dépend dès lors d'un ensemble d'opérations scientifiques qui convoque des champs d'expertise disciplinaires spécifiques amenés à s'accorder sur les objectifs visés pour enrichir les niveaux de pratiques et de vécus ainsi que le sens donné aux différentes pratiques des espaces. Dans cet objectif, trois principes préfigurent à la production des traces (cf. figure 1) : un premier principe de **qualification/authentification** de la trace qui relève de la tâche de séquençage d'une part puis de la tâche d'enrichissement qualitatif direct. Cette tâche accomplie dans l'interaction enquêteur-enquêté permet en outre de qualifier les déplacements et les arrêts sur des dimensions sociales et comportementales, d'authentifier les dimensions spatiales et temporelles des traces enrichies.

Un deuxième principe **d'enrichissement indirect** consiste à collecter un ensemble d'informations sociales, psychologiques (via des traces narratives collectées en entretien), et environnementales (par le biais de données externes caractérisant les environnements de vie).

Enfin, un troisième principe **d'épaississement** des traces s'attache à **l'agrandissement** d'une trace type (un trajet scolaire) via la collecte et la composition de traces sensibles. En portant la focale sur une situation extraite de la réalité quotidienne des enfants, il s'agit ici d'épaissir certaines traces en particulier à travers un faisceau de dimensions qualitatives plus diversifiées (traces sensibles, traces iconographiques, traces auditives, ...). Dès lors, une trace enrichie devient une trace augmentée qui contribue à décrire et expliquer

certains processus associés aux déplacements des enfants, et notamment ici l'autonomie de déplacement.

Ainsi, et par analogie aux 3 V (vitesse, variété et volume) caractéristiques des Big Data, ces principes sont guidés par un triptyque d'exigences liées aux traces, que l'on pourrait résumer aux 3C, à savoir, *continues, complètes et cohérentes* et qui pourraient être définies comme suit.

- *Continues* dans le temps et dans l'espace au niveau de l'enregistrement des données ;
- *Complètes* au niveau des données nécessaires relatives aux caractéristiques et activités de l'individu comme de la cohorte ;
- *Cohérentes* entre les informations spatiales (distance, localisation), temporelles (heure, durée) et sémantiques (modes d'accompagnement ou de déplacement, activités...) associées à chaque trajet et lieu.

La transformation et l'analyse des systèmes de traces dans les laboratoires de Mobi'kids

Cette exigence de complémentarité de données qui conduit à l'enrichissement et l'épaississement des traces dans la phase de collecte et de production des traces de MK est aussi doublée d'un besoin de comprendre les systèmes des mobilités enfantines en contexte d'écologisation des pratiques, c'est-à-dire en considérant les contextes au sens large et de façon dynamique. Partant, elle oriente également une grande part du processus de transformation des traces collectées, étape préalable aux analyses interdisciplinaires qui nécessiteront de faire varier plusieurs niveaux d'échelles des traces et de les comparer selon plusieurs formes de catégorisation.

Niveaux d'échelle et d'adhérence de l'analyse des traces

Afin de mettre en dialogue les divers niveaux de traces avec les espaces de vie, les données de traces sont organisées à différents niveaux "d'échelle" spatiale et temporelle et "d'adhérence" au terrain.

- La dimension spatiale du système de mobilité des enfants doit être appréhendée tout à la fois à l'échelle de l'espace de vie lié aux activités quotidiennes, à l'échelle des déplacements entre lieux, voire à l'échelle des lieux eux-mêmes où peuvent être observées les différentes formes de coprésences et de sociabilités. Les traces inspectées à ces différents niveaux d'échelle permettent ainsi de définir et comprendre différents types d'espaces : espace de vie, espace d'action, espace des mobilités.

- La dimension temporelle des pratiques spatiales et sociales est abordée à deux niveaux.

Le premier niveau permet d'appréhender les processus de changement dans les routines et les territoires de vie des enfants sur la période de transition scolaire (école/collège).

L'approche longitudinale du protocole permet ainsi, de façon répétée, d'observer et mesurer à l'identique les comportements des enfants et de leurs parents aux deux périodes afin d'appréhender la dynamique des processus psycho-sociologiques engagés dans cette transition.

Le second niveau porte sur les dynamiques quotidiennes. Le temps est alors considéré dans des granularités fines (de l'ordre de la minute) et continues (durant plusieurs journées consécutives) pour mesurer des régularités (récurrence de lieux, de déplacements, d'activités) à l'échelle de la semaine. Cela permet aussi de détecter des niveaux plus fins de traces "encapsulées" dans les activités et déplacements quotidiens, de nature plus imprévues et spontanées mais souvent négligées car insignifiantes ou encore très peu perceptibles. Ces formes plus spontanées de traces de déplacements et d'activités peuvent alors être (comme nous le verrons plus loin) le résultat d'analyses de patterns ou d'outliers.

- Enfin, traversant ces deux premières dimensions, différents degrés d'adhérence (Amar, 1993) aux lieux habités doivent pouvoir être documentés : celui de l'espace de vie global décrit par les morphologies urbaines où s'inscrivent les flux de déplacements, celui des lieux et des moments décrits par leurs fonctions et leurs usages et enfin celui des interactions entre les individus en mouvement et leur environnement le plus direct et le plus immédiat. Ainsi, rapportées aux données MK, divers niveaux d'adhérence peuvent être traduits: les traces géolocalisées correspondent plutôt à l'échelle du territoire des enfants, les traces narratives à celle des activités sociales et de leurs enchaînements, et les traces sensibles à celle des cours d'action et des pratiques en situation.



Figure 3 : Niveaux d'échelles et d'adhérence (1- haut gauche) Quantification des déplacements dans l'espace public : on visualise ici les trajets des enfants se déplaçant seuls à pied et on identifie les rues les plus utilisées. (2 - bas gauche) Quantification du temps passé dans les lieux : on observe les différences des lieux les plus fréquentés par les enfants en dehors de l'école en fonction du temps d'enquête (primaire et collège). (3- droite) Qualification de la fonction des lieux fréquentés : on peut faire ressortir des logiques spatiales de spécialisation dans les pratiques des enquêtés.

Préparation et mise en catégorie des traces enrichies et narratives entre les disciplines de Mobi'kids

Une des étapes fondamentales du processus d'analyse des traces réside dans la constitution d'une base de données qui réponde aux objectifs interdisciplinaires de MK. Celle-ci doit permettre de composer un univers de traces qui soient utilisables par divers disciplines et surtout complémentaires et articulables entre elles. Les traces transformées et enrichies sont ainsi mobilisables à différents niveaux de granularité (individu, classe, famille) et à différentes échelles spatiales (école, quartier, commune) et temporelles (quotidienne, longitudinale). Dans cette perspective, les différents types de traces nécessitent un processus de réduction qui réponde à divers niveaux d'attente (cf. figure 4) et plusieurs niveaux de formalisation, selon que les analyses s'inscrivent dans une démarche descriptive,

illustrative ou explicative.

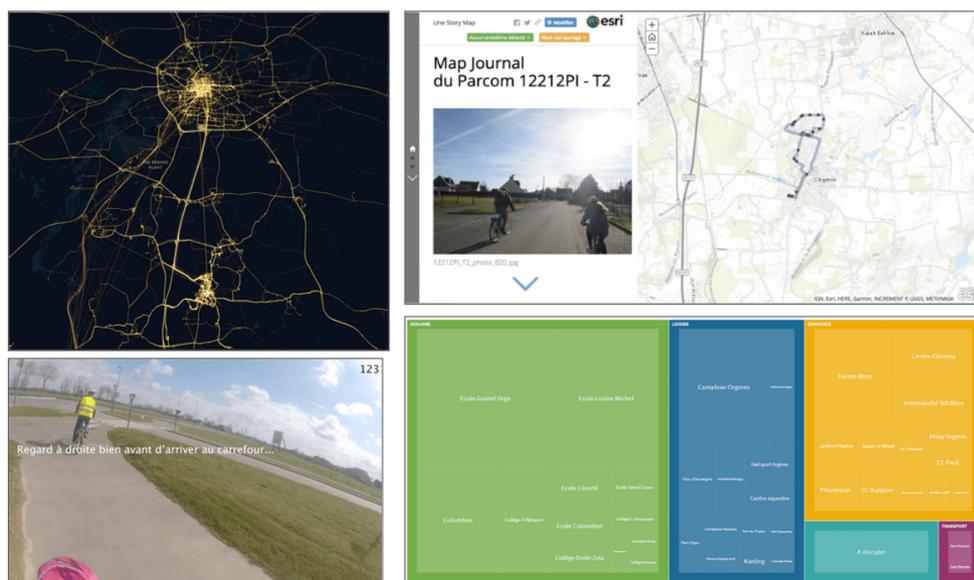
	Nature des données	Attentes	Procédés	Données produites
Traces enrichies	Traces GPS enrichies (déplacements et arrêts qualifiés)	Précision Cohérence Complétude	Corrections automatiques et manuelles	Trajectoires spatiales et temporelles définies comme une suite de lieux, de déplacements (et d'arrêts transitoires en cours de déplacements) qualifiés par les activités et les modes de transport.
	Traces narratives	Contextualisation et relativisation Confirmation	Extraction production de profils et indicateurs	Indicateurs socio-démo Indicateurs psycho-socio
Traces sensibles	Audiovisuel (son, photo, vidéo)	Identification de phénomènes sensibles	Montage audiovisuel	Descripteurs : attention, ambiance, mémoire
	Parole enregistrée	Compte rendu d'expérience	Transcription de la parole et écoute des tonalités	Descripteurs : perception située et vécu sensible

Figure 4 : Réduction/transformation des données collectées en données d'analyse

Vignette 4 : Donner à voir ces traces sous des formes diversifiées

La diversité des traces collectées et transformées au sein du projet MK peut être rendue "visible" de multiples manières. En premier lieu, la représentation cartographique sous des formes diversifiées constitue une première méthode pour donner à voir les trajectoires dans le temps et l'espace issues des traces géolocalisées. En complément l'utilisation de la visualisation de données via des représentations graphiques offre un autre angle de lecture des pratiques comme des dynamiques. Le montage vidéo et photo, accompagné d'une bande son ou de commentaires représente un autre médium pour valoriser et structurer les traces

sensibles. Enfin, le recours aux cartes narratives (story map) permet de combiner à la fois traces géolocalisées et sensibles au sein d'une même interface pour donner à voir de manière originale et dynamique les pratiques de mobilité.



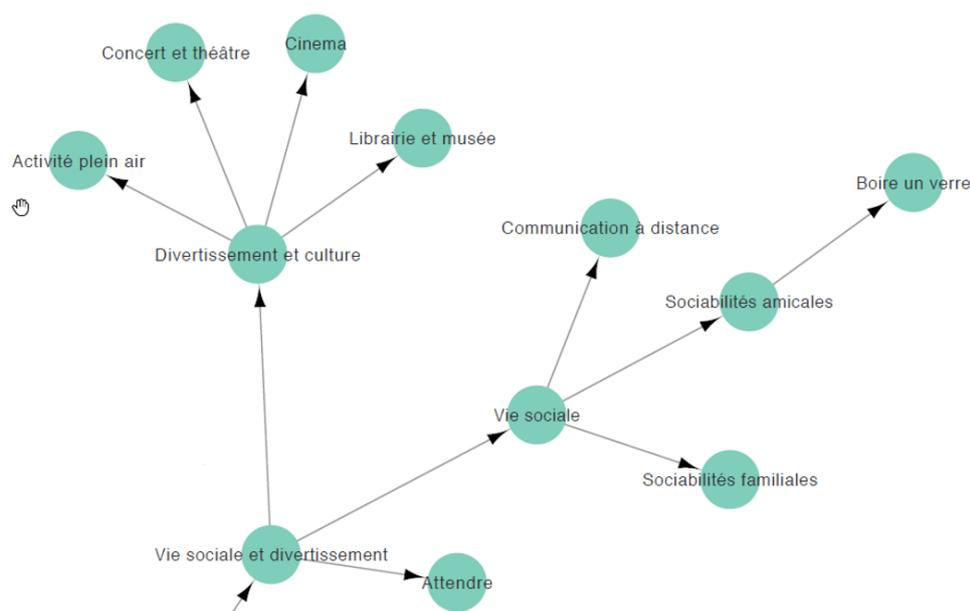
Une autre étape représentative de discussions et d'opérations interdisciplinaires consiste en la mise en catégories des traces enrichies, narratives ou sensibles. Cette opération doit permettre de rendre un certain nombre de traces et données interopérables d'une part et d'épaissir et consolider certains niveaux d'analyses d'autre part.

Dans le champ du traitement des traces, et en particulier en sciences de l'information, la catégorisation a plusieurs significations et différentes fonctions. Par exemple, la catégorisation de traces enrichies liées aux activités associées aux lieux fréquentés correspond à la construction d'ontologies (cf. vignette 5).

Vignette 5 : L'ontologie de MK

L'ontologie définie pour ce projet reprend l'ensemble des concepts liés aux pratiques observées de MK et s'appuie sur le thésaurus européens des activités HETUS : Harmonised European Time Use Surveys (Eurostat 2019). Elle regroupe

des concepts liés aux activités, aux modes de déplacement et aux types d'accompagnement. Seules les relations de subsumption sont actuellement employées. Cette ontologie définit donc des catégories et elle les classe en groupes et sous-groupes en fonction de leur similarité. La figure suivante représente une partie de l'ontologie portant sur la vie sociale et les divertissements. Dès lors, nous pouvons constater que l'activité "cinéma" est plus proche de l'activité "concert et théâtre" que de l'activité "communication à distance". Néanmoins, ces trois activités sont des activités du groupe "vie sociale et divertissement". La proximité entre concepts se traduit par une mesure de similarité. Elle est ainsi très forte entre "cinéma" et "concert et théâtre" et faible entre "cinéma" et "communication à distance".



La construction de cette ontologie permet également des analyses de patterns de trajectoires (comportements similaires). Cette opération d'analyse des traces enrichies liées aux activités a aussi conduit à ré-interroger les notions de pratiques versus comportements dans leurs nuances. Ainsi un comportement défini comme un acte observable et objectivable serait différent d'une pratique qui définit un ensemble de comportements soumis à diverses règles ou structures d'organisation (temporelles, spatiales, sociales). Par ailleurs, l'ontologie associée aux traces d'activités peut, dans certaines situations observées dans MK, être

affinée et ainsi définir une nouvelle catégorie d'activité qu'il faudra relier à l'ontologie initiale pour permettre un traitement automatique et une recherche de clusters (au sens de classes d'une typologie statistique). C'est le cas, par exemple, des traces enrichies d'activités associées aux arrêts transitoires, comme celle de cueillir un végétal, de contempler le paysage ou encore de discuter de façon informelle et courte avec une connaissance durant un trajet.

En outre, une autre forme de mise en catégories, cette fois plus analytique, renvoie à l'opération de clustering automatique. Par exemple, au niveau des traces enrichies, des catégories de traces proches (au niveau de certains "concepts") peuvent être obtenues à l'aide de cette opération et ainsi, être analysées, caractérisées et contribuer à l'extraction de connaissances qui pourront ensuite être soumises à d'autres opérations d'analyses confirmatoires ou explicatives. Ainsi, des résultats obtenus à un niveau peuvent être confirmés à d'autres niveaux.

En effet, ces clusters de traces enrichies qui regroupent des grandes quantités de traces sémantiquement ou spatialement similaires peuvent composer une première étape d'exploration des traces qui pourront ensuite être analysées sous d'autres approches et à d'autres échelles spatiales et temporelles, par exemple.

Ainsi, l'analyse par clustering de traces liées à des déplacements d'enfants peut conduire à montrer une optimisation des distances lors de trajets routiniers à l'opposé de déplacements réalisés dans le cadre d'activités occasionnelles décrits comme moins directs. Cette connaissance par clustering peut alors être approfondie et consolidée lors de l'analyse des traces sensibles extraites des parcours commentés.

Les corpus de traces sensibles issues des parcours commentés sont multiples et hétérogènes, relevant aussi bien d'enregistrements de paroles en marche et de leurs transcriptions (traces narratives augmentées), de prises de vue photographiques ou d'enregistrements vidéo de parcours (traces iconographiques). Aussi, l'analyse de ces corpus procède à partir de catégories qui permettent de décomposer une part non détectable (à partir des autres traces) des rapports à l'espace des enfants. Plus précisément, celles-ci sont documentées à partir d'un ensemble de descripteurs qui permettent de rendre compte de l'expérience in situ des enfants, de rendre manifestes leurs manières d'être en mouvement, et de rendre observables leurs formes de conduite durant les parcours (Thibaud, 2001). Ces descripteurs s'appliquent aussi bien à ce que disent les enfants (leurs paroles) qu'à ce qu'ils font (leurs actions), aussi bien aux traces narratives qu'aux traces

audio-visuelles produites au cours des parcours. Quatre descripteurs principaux donnent sens au corpus de traces : les formes d'habitation permettent d'identifier les répétitions, routines, variations et événements au cours des parcours ; les conduites attentionnelles permettent de comprendre la manière dont les enfants se rendent plus ou moins attentifs à l'environnement de leurs parcours ; les traces mnésiques permettent de montrer l'incidence de la mémoire et des expériences passées dans l'expérience des parcours ; et les tonalités affectives permettent de relever la part des émotions et des dispositions affectives dans le vécu des parcours des enfants. Habitude, attention, mémoire et affect sont donc les quatre notions qui permettent de décrire les catégories d'analyse des traces sensibles à partir d'une grille élaborée dans un tableur qui organise les diverses traces issues des parcours commentés. Ainsi, pour chaque descripteur, aux traces narratives augmentées (verbatim extraits des parcours et entretiens) et traces iconographiques répertoriées sont associés des tags temporels valant opérateur d'articulation des différents niveaux de traces sensibles.

La trace : notion opératoire interdisciplinaire

A travers la préparation et la mise en œuvre d'une enquête sur la mobilité des enfants, nous avons montré comment la notion de "trace" agit comme outil opérationnel tant sur le plan de la structuration de la chaîne de collecte et de traitement des données que sur le plan de la discussion scientifique interdisciplinaire. La trace est en effet une notion partagée, un objet intermédiaire évolutif dans les échanges et les interactions entre chercheurs. Partant de nos différents champs disciplinaires, l'objectif commun est de produire des connaissances sur l'expérience des déplacements des enfants, documentée et interprétée aux échelles spatiales macro, méso et micro et selon des temporalités de l'ordre de quelques minutes (le temps d'un parcours) à plusieurs mois (le temps d'une transition entre école primaire et collège). En croisant dans un premier temps les concepts disciplinaires habituellement utilisés pour décrire la mobilité spatiale, la notion fédératrice de "trace" a permis de nous entendre sur un univers conceptuel partagé et de décliner des définitions d'objets et de notions théoriques et empiriques structurants pour la recherche. Entendue d'un point de vue conceptuel comme les marques matérielle, mémorielle et sensible laissées par une pratique passée (ici les parcours réalisés quotidiennement par les enfants dans l'espace urbain), la trace est déclinée en plusieurs entités observables mobilisant les méthodes et outils des différentes disciplines impliquées. Afin de décrire et d'objectiver l'enchaînement des déplacements et des activités sur plusieurs jours, la trace prend la forme d'un ensemble de données spatiales et temporelles issues de capteurs ; des méthodes informatiques et géomatiques permettent de les rendre intelligibles et comparables dans le but de discerner des pratiques récurrentes mais aussi imprévues. Afin de comprendre ces pratiques qui structurent le processus d'apprentissage des déplacements en autonomie, la

trace est enrichie par les discours des enquêtés ; des méthodes psychologiques permettent alors de dégager des variables (psychologiques, sociales, environnementales) explicatives et contextuelles. Enfin, pour saisir les dimensions sensibles de l'expérience individuelle de la mobilité, la trace rassemble un faisceau d'indices de l'ordre des ambiances (sonores, visuelles, olfactives...) ou des conduites attentionnelles extraites de parcours commentés issus de méthodes d'observation participante d'une sociologie pragmatique et interactionniste ; les modalités d'observation mises en œuvre visent ici une interprétation de type modal des pratiques de mobilité. De surcroît, le grand intérêt de cette approche interdisciplinaire est aussi de pouvoir croiser, superposer et mettre en perspective ces différents types de traces (géolocalisées, enrichies, narratives, sensibles) grâce à la formalisation et l'opérationnalisation *co-construite* de la chaîne de collecte et d'analyse. Aussi, les traces sont en quelque sorte "interopérables" et fonctionnent à la fois en termes de superposition-complémentarité (à diverses échelles et degrés d'adhérence) et en termes de cadrage-décadrage (conduisant à zoomer ou dézoomer sur des portions particulières d'espaces de déplacement). Par exemple, pour interpréter la trace sensible d'un parcours, nous pourrions la contextualiser avec la cartographie de la trace géolocalisée enrichie indiquant les habitudes quotidiennes. En retour, pour donner du sens aux formes imprévues de pratiques détectées automatiquement dans les traces géolocalisées, nous pourrions d'une part les rendre intelligibles à partir des traces narratives extraites des attitudes et comportements déclarés des enfants ou encore des attitudes éducatives des parents et d'autre part étayer leur compréhension en puisant dans les modalités expérientielles extraites des traces sensibles. Ces perspectives d'analyses sont conçues comme une mise en discussion des points de vue et des approches disciplinaires qui devrait contribuer à une connaissance renouvelée des mobilités enfantines mettant l'accent sur les relations qui se tissent dans le temps entre les personnes, leurs pratiques et leurs espaces de vie.