TOPIC: MODELS TRANSFORMATION AND DATA INTEGRATION IN OPENSTREETMAP FOR ACCESSIBILITY: FEASIBILITY STUDY FROM THE GEOSTANDARD

MSc Research Project

International of Computer Science

ISIMA

University of Clermont Auvergne

WRITTEN BY:

Eunice Guédaliah WADJOM KAMMEGNE

ADVISOR : Jean-Marie FAVREAU

TABLE OF CONTENTS

LIST OF FIGURES 4
INTRODUCTION 5
I. OVERVIEW OF THE RESEARCH TOPIC 6
I.1. Context 6
I.2. Problematic 6
I.3. Proposed solution 6
II. PRIOR RELATED WORK 7
II.1. Industrial solutions 7
II.1.1. Presentation 7
II.1.2. Limits 7
II.2. Literature review 8
II.2.1. OpenStreetMap 8
II.2.2. Geostandard 11
II.2.3. Data integration 14
II.2.4. Data matching 15
II.2.5. Data fusion 16
II.2.6. Data coverage 17
III. PROJECT PLANNING 18
III.1. Gantt diagram 18
III.2. Pert chart 19
IV. MATCHING STUDY BETWEEN OPENSTREETMAP AND CNIG 20
IV.1. Geometry overview 20
IV.1.1. Coordinate system 20
IV.1.2. Study of the geometries 21
IV.2. Study of semantics of OSM and CNIG data 24
IV.2.1. OpenStreetMap accessibility model → CNIG accessibility standard 24
IV.2.2. CNIG accessibility standard → OpenStreetMap tags 29
V. IMPLEMENTATION OF TRANSFORMATION MODELS 29
V.1. Study environment and tools 29
V.2. Creation of the CNIG database 30
V.3. Fusion of the CNIG database with OSM data 31
V.3.1. Connection to the CNIG database and loading of OSM accessibility data 31
V.3.2. Design of OpenStreetMap data transformation models 32
V.3.3. Presentation of the results 43
CONCLUSION AND PERSPECTIVES 46

REFERENCES 47
APPENDIXES 50

LIST OF FIGURES

Figure 1 : OpenStreetMap Data Model (Mericskay 2017) 9
Figure 2 : Conceptual Data Model for road accessibility (CNIG Accessibility Standard 2021) 13
Figure 3 : Conceptual Data Model of access equipment (CNIG Accessibility Standard 2021) 13
Figure 4 : Conceptual Data Model of ERP (CNIG Accessibility Standard 2021) 14
Figure 5 : Differentiation of matching 16
Figure 6 : Gantt diagram 18
Figure 7 : Pert chart 19
Figure 8: Point 21
Figure 9: Simple non-closed linestring 21
Figure 10: Polygon defined by exterior ring 21
Figure 11: Multipoint with 4 parts 22
Figure 12: Simple multilinestring defined by 4 endpoints of 2 elements 22
Figure 13: Multipolygon consisting of two elements defined by exterior rings and three interior rings 22
Figure 14: Simple schema of the mapping of crossings (Biagi, Brovelli, and Stucchi 2020)... 25
Figure 15: Example of mapping the height of a traffic sign (Biagi, Brovelli, and Stucchi 2020) 27
Figure 16: Example of mapping the parking spaces reserved for people with disability(Biagi, Brovelli, and Stucchi 2020) 28
Figure 17: CNIG database 31
Figure 18: Loading OSM data from Orange 31
Figure 19: Download of Orange's footways 32
Figure 20: Generation of pathway trunks 33
Figure 21: Downloading pedestrian paths on OSM 34
Figure 22: Vertices extracted for line and polygon layer (“Géométrie Vectorielle — Documentation QGIS Documentation” n.d.) 34
Figure 23: Model cut lines with points 35
Figure 24: Transforming points into lines 36
Figure 25: Cut with lines 36
Figure 26: Geostandardization of nodes 37
Figure 27: Selection of initial and final path nodes 38
Figure 28: Joining of pedestrian paths with initial and final nodes 38
Figure 29: Creation of the geostandard path trunks 39
Figure 30: Generate obstacles of pathways in the CNIG format 39
Figure 31: Downloading of obstacles in OSM 40
Figure 32: Geostandardization of obstacles 41
Figure 33: Model for integrating OSM data into the CNIG database 41
Figure 34: Exportation of nodes into the node_tmp table 42
Figure 35: Insertion of nodes data into the "Noeud" table of the CNIG database 42
Figure 36: Pathway nodes represented according to the CNIG accessibility standard 43
Figure 37: Attribute values of pathway nodes according to the CNIG accessibility standard.. 44
Figure 38: Obstacles represented according to the CNIG accessibility standard 44
Figure 39: Attribute values of obstacles according to the CNIG accessibility standard 45
Figure 40: Pathway sections represented according to the CNIG accessibility standard........ 45
Figure 41: Attribute values of pathway sections according to the CNIG accessibility standard 45

INTRODUCTION

OpenStreetMap (or "OSM") is an international project founded in 2004 with the goal of creating a free map of the world. This map of the world is freely editable and made by Internet users. OSM data can be used in a variety of ways, including producing paper and electronic maps, geocoding addresses and place names, and route planning. However, there is often a lack of accuracy when it comes to road accessibility for people with reduced mobility. Hence, The French Ministry of Environment has been working since 2020 on the design of a data standard, led by the CNIG and in collaboration with the CEREMA, professionals in the field and representatives of the people concerned for collecting data on road accessibility, for people with disabilities or reduced mobility. It is in this perspective that we will work on the project: "Models transformation and data integration in OpenStreetMap for accessibility: feasibility study from the geostandard". It will be a question of studying the possibilities and measures of merging CNIG data with OpenStreetMap data in order to improve the accessibility of pathways for people with reduced mobility.

I. OVERVIEW OF THE RESEARCH TOPIC

I.1. Context

The setting up of the project of model transformation and data integration into OpenStreetMap for accessibility is justified by the recent initiative of the Ministry of Environment, in application of the law of orientation of mobilities (LOM) to facilitate the constitution of a collaborative database, resulting from a regular and coordinated collection by local communities. This initiative will lead to the establishment by each community of a high precision geostandard database to facilitate road accessibility. Therefore, it will be essential and necessary to be able to transform and integrate these data into OpenStreetMap, which is where our project comes from. It will contribute in an undeniable way to the resolution of problems such as:

- Road accessibility of people with reduced mobility;
- Lack of geographical accuracy of data in OpenStreetMap;
- Heterogeneity of data quality;
- Evolution of OpenStreetMap for accessibility.

I.2. Problematic

Moving for people with disabilities is often difficult with a high level of uncertainty due to the lack of geographic accuracy of data in OpenStreetMap. Accessibility is not yet guaranteed on all travel routes. Until now, the only sources that could really be used with a sufficiently high level of accuracy were proprietary databases, hence the Mobility Orientation Law (published in the Official Journal on December 24, 2019) stipulated the obligation for road management authorities to create databases describing the accessibility of the route in order to feed the route planners and allow disabled people to build accessible routes. So, a way must be found to help communities to integrate these data into OpenStreetMap. Designing a tool to transform models and integrate data into OpenStreetMap for accessibility is therefore necessary and imperative.

I.3. Proposed solution

There is a need to improve the autonomy, safety and fluidity of travel for people with reduced mobility. To achieve this, the need for an accurate, comprehensive and universally accessible route planner is clear.

On one hand, it is urgent to collect data from local authorities in the CNIG accessibility standard.

On the other hand, it is necessary to integrate these data into OpenStreetMap in order to:

- Create a safer route calculation for people with disabilities or reduced mobility;
- Reduce the uncertainty and lack of precision of geographical data in OpenStreetMap;

- Ensure the accessibility of routes to anyone in need.

We therefore thought of the project: Models Transformation and Data Integration in OpenStreetMap for Accessibily, which fits perfectly with the impetus given by the Ministry of the Environment to design a collaborative database that meets the needs of people with disabilities or reduced mobility. This project will help communities to integrate their data directly into OpenStreetMap in order to make them accessible and usable by all.

II. PRIOR RELATED WORK

II.1. Industrial solutions

II.1.1. Presentation

Until now, the only sources that can really be used to calculate routes for people with reduced mobility with sufficient accuracy are proprietary databases.

Wegoto develops solutions for capturing geographical and topographical data in the field. It assists local authorities in modelling their accessibility databases and creates mobility databases for pedestrian, scooter and bicycle, for private and shared use. Wegoto has several applications to its credit, including:

- Wegoto: a mobile application to help with soft navigation that allows each user profile to calculate the most suitable multimodal itinerary and to be guided according to its profile. It is an itinerary calculator adapted for pedestrians, cyclists, PRMs, the disabled and transport.
- WegoTrack: the software WegoTrack collects all the cartographic data of the outdoor pedestrian travel chain in the field. It has an input interface designed as a metaphor of reality allowing a clear and fast input of the collected data.
- OSM converter: the program converts the data from the WegoTrack collection software to OSM format.

We also have the application "Acceslibre": a collaborative platform for accessibility. It allows everyone to inform and consult the accessibility data of ERP and thus facilitate the movements of people with disabilities.

II.1.2. Limits

The solutions developed so far are proprietary and therefore still not accessible to all.

The applications developed by Wegoto undeniably contribute to the improvement of soft mobility, but are nevertheless limited in certain aspects:

- Wegoto is not complete, the territories where they have actually made the acquisition are Grenoble and Isère.
- WegoTrack is a proprietary application, so it is not accessible to all communities and it is the same for their OSM converter.
- Acceslibre is limited and then not sufficient to guarantee the safe travel of persons with reduced mobility because it is just concentrate on the accessibility of ERP and it don't describe the pathway inside those ERP.

II.2. Literature review

II.2.1. OpenStreetMap

II.2.1.1. Presentation

OpenStreetMap is a free, editable map of the whole world that is being built by volunteers largely from scratch and released with an open-content license. OpenStreetMap is built by contributors, usually called mappers within OpenStreetMap, who gather information by driving, cycling, or walking along streets and paths, and around areas recording their every move using Global Positioning System (GPS) receivers (Brovelli, Codrina, and Coetzee 2019). OpenStreetMap contains details of much more than just roads. All mappers are free to add any geographic object they find, so people have added phone boxes, bus stop, parks, public toilets, places of worship, and much more. The data and the project are both intended to give mappers as much flexibility as possible, so they can map features as accurately as they can.

The OpenStreetMap (OSM) geographic data model has three principal object types: nodes (points), ways (polygons and polylines), and relations (logical grouping of all three object types to express real-world geographical relationships) completed by tags which describes every object (Pruvost and Mooney 2017).

- Tags: They are "key=value" pairs that can be linked to each geographic feature in OpenStreetMap. These attributes describe the geometric object. Both the key and the value are strings, with a maximum length of 255 characters (including spaces), with one constraint: each key must be unique for the same element. If a geographic object has no attributes, most rendering engines will not represent it at all.
- Nodes: a geographic point is called a node, and it is represented by its latitude, longitude, and as many attributes as necessary. For example, nodes are used to represent stores, bus stops, benches, vertices and mailboxes. A node without attributes will always be a sub-element of another element.

➢ Way: an ordered list of nodes is called a way. A way has a maximum of 2000 nodes, to ensure that tools and users are not overwhelmed by huge objects that are too difficult to manipulate. Ways are used to represent linear objects such as paths, roads, railroad lines, coastlines, and boundaries.

  Polygons are not a special type of data at this time: they are simply closed ways, i.e., the last point is merged with the first. They are used to represent buildings, parks, and areas of a land use plan.

➢ Relations: a relation is an ordered list of nodes, ways, and relations. Each member of a relation has an optional role that adds information to that sub-element. Like tags, these roles are stored in strings (limited to 255). Relations can represent road or bike routes, administrative boundaries or rivers, again depending on the tags attached to them.

➢ Identifiers: an element in the OpenStreetMap database, whether it is a node, a way, or a relation, is identified by a unique numeric identifier. This identifier has no particular meaning; it is only used to uniquely identify an element. A relation and a way use these identifiers to reference each of its sub-elements. Two ways meet only if they reference at least one node identifier (not two nodes with identical coordinates). Similarly, a closed way representing a surface must reference the same node identifier twice.

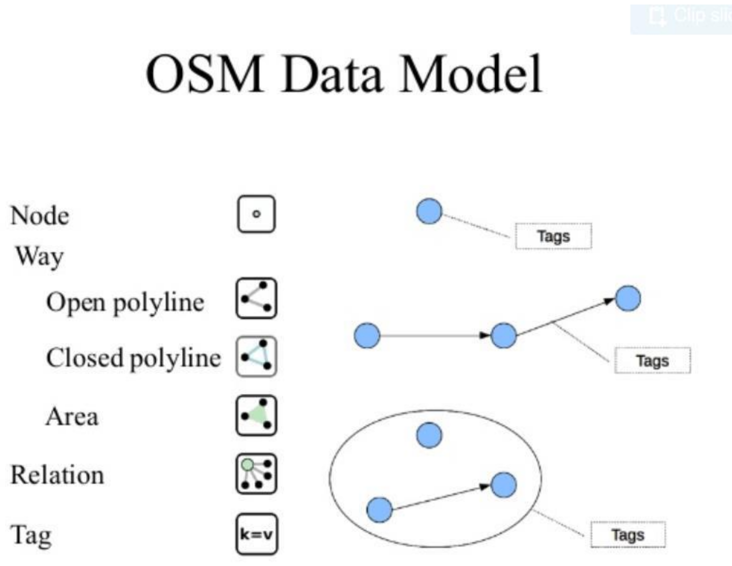

Figure 1 : OpenStreetMap Data Model (Mericskay 2017)

II.2.1.2. Formalised representation (H. Pruvost and P. Mooney, 2017)

We shall now provide a more formalized representation of OpenStreetMap.

- In OSM, a node N is a geographic coordinate with tags Nt where the set of tags can be empty.
- A way w is an ordered set of nodes (N0, N1, . . . Nn-1). If the way w represents a polyline feature such as a road or path, then N0 is different from Nn-1. However, if the way w represents a closed polygon feature such as a building or a lake then N0=Nn-1. For polyline ways, there should always be a minimum of two nodes while for closed polygon features there should always be a minimum of three nodes. A way can also have a set of tags Wt. The set Wt can be empty, but normally it is encouraged that this is not the case.
- Finally, in OSM a relation R is an element representing the logical relationship between objects. A relation r is represented by a set containing at least one object. A relation can contain any combination of existing nodes, ways and even relations themselves. These are called members. As is the case with nodes and ways relations can also have a set of tags Rt associated with them. As before this set can be empty but this is discouraged.

II.2.1.3. Types of relations

According to both the OSM Wiki and TagInfo, the five most frequently occurring types of relations (where Rt contains the tag for type) are outlined as follows:

- Multipolygon: a multipolygon relation can have any number of ways and these ways must somehow form valid rings to build a multipolygon from. Multipolygon relations are used to represent complex areas. Generally, the multipolygon relation can be used to build multipolygons in compliance with the OGC Simple Feature standard. Subsequently multipolygons allows for the expression of arbitrarily complex relations within OSM.
- Restriction: relations with the type=restriction key allows the modeling of different types of traffic flow restrictions at junctions, intersections, etc. A turn restriction at a junction is represented by a relation that has a set of tags describing the type of turn restriction. This turn restriction relation is not necessarily limited to turns. It can also be used in a number of other situations.
- Route: this type of relation models a regular known line or path of travel. According to the OSM Wiki routes can consist of paths taken repeatedly by people and vehicles: a ship on the predefined shipping route, a car on a numbered road, a bus on its route or a cyclist on a national route.
- Public Transport: relations with this type=public_transport allow for the description of relations used in the public transport tagging scheme in OSM. This relation corresponds

to the description of all types of public transport stops, stations, halts, areas or similar. As stated in the OSM Wiki pages “A stop area consists of everything regarding the embarkation and disembarkation of a specific public transport vehicle or service” (OSM Wiki, OpenStreetMap Relation Type-Public Transport). In the example of a specific railway line, this includes the adjacent platform, services on that platform, buildings, and information describing it (platform number, identification, etc.).

- Route Master: relations of type type=route_master contain all the direction and variant routes and information belonging to a whole route service. Routes or services are represented by vehicles that always run the same way with the same reference number. Each direction of a route should be tagged as a separate relation. If a route has several variants (e.g., different way at weekend), these variants should also be in separate relations.

In addition to relations with the type tag key assigned to one of the five values above, there are other frequently occurring types of relations in OSM including boundary, associatedStreet, site, waterway and bridge. However, these types of relations appear much less frequently within the global OSM database.

II.2.2. Geostandard

The CNIG Accessibility Standard, also known as geostandard, is a standard for collecting data on road accessibility for people with reduced mobility or disabilities. The accessibility information is particularly interested in:

- the path on the road and in the public space;
- Access routes to public transport stations and stops, whatever the mode, including connecting routes;
- Access to public administrative and social services and more generally to all establishments open to the public.

The CNIG Accessibility standard defines the data model and the catalogue of objects required to describe accessibility for the travel chain in the areas of roads and public spaces, and the built environment, here restricted to establishments open to the public (ERP).

It proposes a base of essential data to ensure a non-blocking path for disabled people or people with reduced mobility. It includes all the technical and organisational specifications for storing and exchanging the corresponding geographical data in digital format.

The geostandard defines 09 main types of objects:

- Paths: it describes physical paths to get from one point to another on foot (or in a wheelchair, etc.).
- Path_Trunk: a space opens to the public in which the person moves. The pathway section combines physical and PRM - DSP traffic-related characteristics.
- Node: extremities of a pathway section.
- Obstacle: an element in the pathway that may impede or prevent movement.
- Circulation: a "standard" path on a regular surface, without access equipment. For example: a path on a sidewalk, a place, a pedestrian area, etc.
- Access equipment: includes all access equipment:
    - Crossing: any marked area allowing pedestrians to cross at level the lanes reserved for road, cycle or public transport traffic.
    - Ramp: a sloping structure used to cross a difference in level or a change of level or floor.
    - Staircase: a structure for going up or down, consisting of a succession of steps.
    - Escalator
    - Conveyor belt: flat surface with a translational movement, used to transport people.
    - Lift: vertical transport system used to move people up and down.
    - Elevator: system for overcoming a difference in level, equipped with a platform or a nacelle.
    - Entrance: opening allowing passage.
    - Selective passage: device allowing the passage of pedestrians, but discouraging that of cycles and motorised vehicles.
    - Quai: equipment of a transport mode allowing access to the vehicle.
- PMR parking: vehicle parking space on a road reserved for people with reduced mobility.
- ERP: establishment receiving the public or installation open to the public (IOP).
- ERP pathway: pedestrian pathway within the site of an ERP, starting from an entrance or PRM parking space and ending at an entrance or reception area of the ERP.

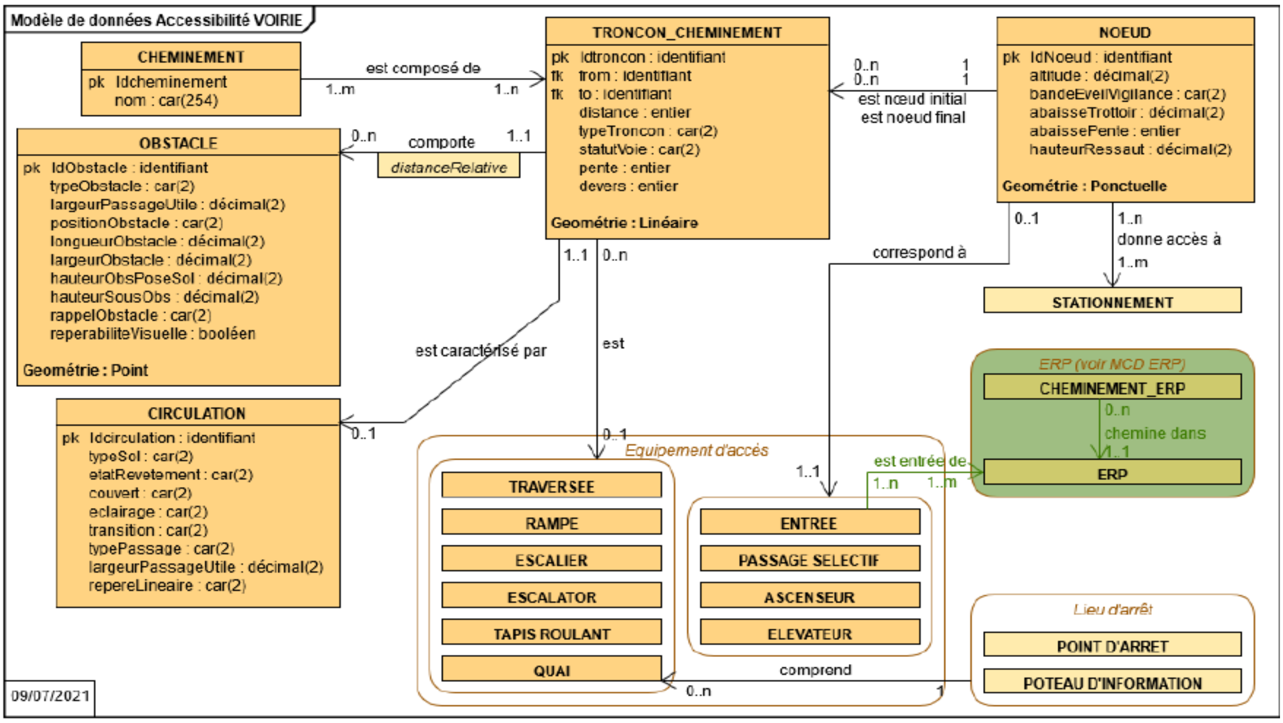

Figure 2 : Conceptual Data Model for road accessibility (CNIG Accessibility Standard 2021)

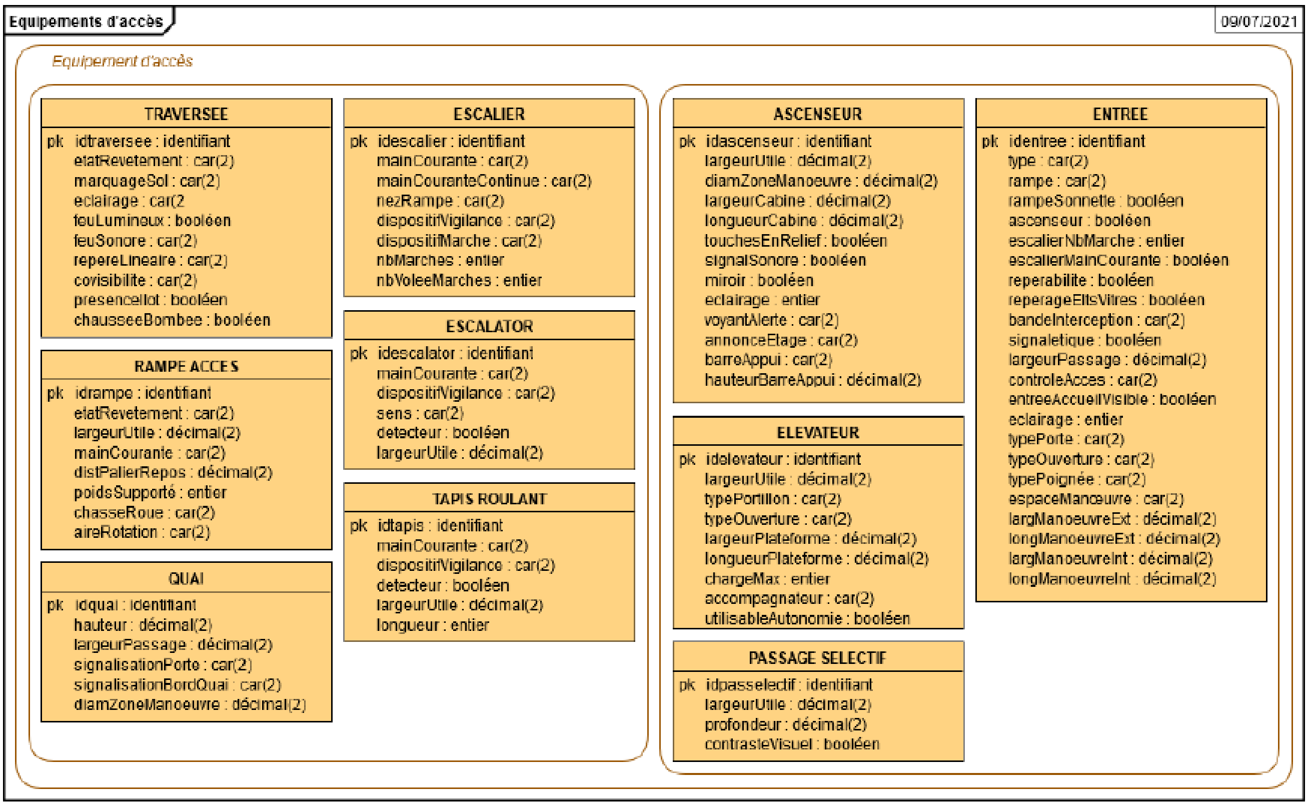

Figure 3 : Conceptual Data Model of access equipment (CNIG Accessibility Standard 2021)

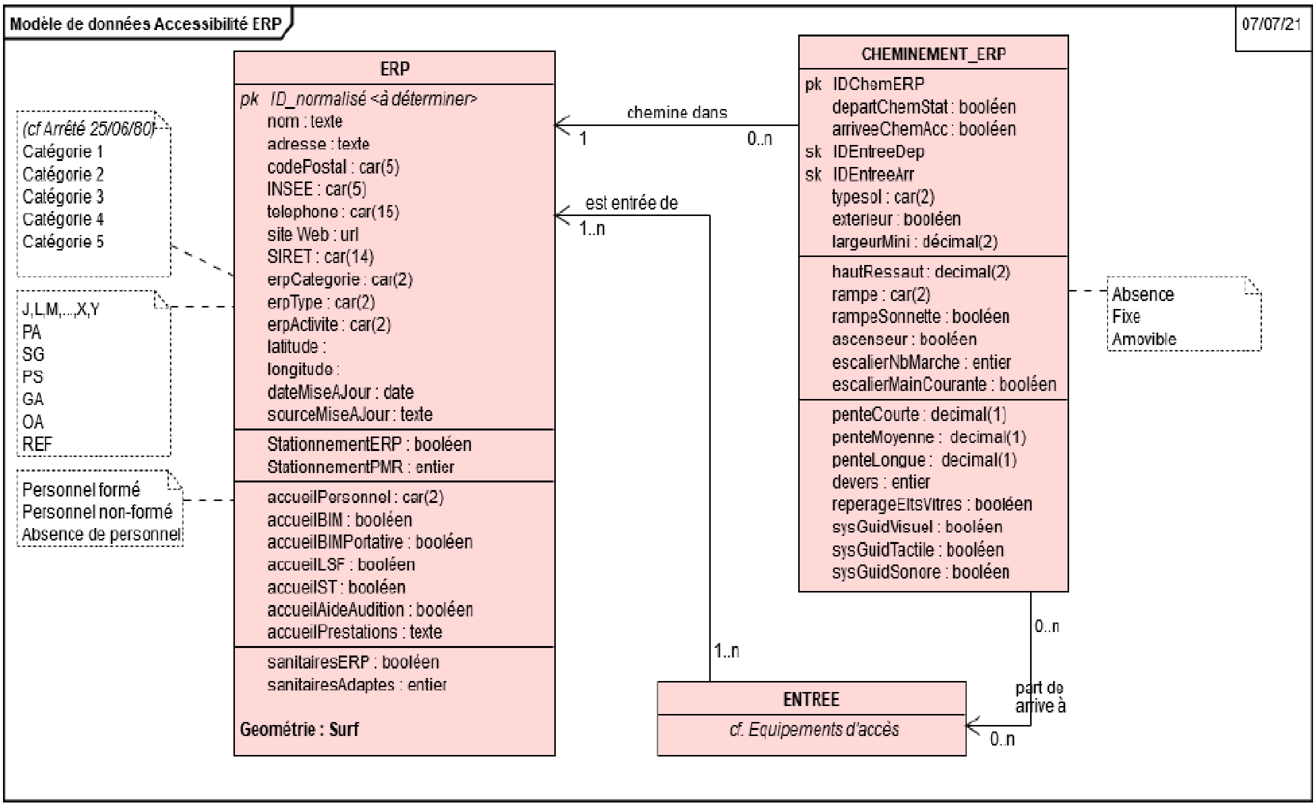

Figure 4 : Conceptual Data Model of ERP (CNIG Accessibility Standard 2021)

II.2.3. Data integration

Data are different in source, acquisition approach, data model, data format, coordinate and projection, semantic expression and other aspects. So the multi-source data can't be used to do data fusion directly, we have to perform data integration first.

The task of data integration is to eliminate the differences among multi-source spatial data. According to certain standards, those multi-source data are integrated into one uniform system, through which the direct management and structure of multisource data can be realized.

Content and techniques of multi-source data integration (chen et al. 2013)

- Unified Coordinates: different data use different coordinates and the current coordinates in using includes: old BJ54, new BJ54, CD80, WGS84 and CGCS2000 coordinate. The using of CGCS2000 would influence the geometry precision of data in old coordinates, so unified coordinates is necessary and the coordinates can be transferred through three parameters or seven parameters models. Geostandard data use Lambert 93 coordinates and OSM data use WGS84. In order to unify these coordinates, we will use the tool POSTGIS in which unified coordinates methods have already been implemented.
- Projection Transformation: according to the mapping purposes and locations, the existing multi-source data use different projections. It is necessary to set up the corresponding

relationship of the points from different planes through projection transformation. For vector data, the projection can be transferred through certain equation. But for raster data, the new image can't be built through point-by-point conversion, because point-by-point conversion may lead to slot and overlap in the new image. So the projection transformation of image is through reversed calculation (Shaomei Li, 2004). All these methods are theoretical, so we will use POSTGIS in order to perform projection transformation.

- Data Format Exchange: Different software and system use different data models to describe the world, which leads to different data formats. The common used data models are: Arc/InfoMapInfo, MapGIS, AutoCAD, SuperMap and so on. Each data model has its own data format and the exchange methods are ripe, but in our case we will use the OSM format, and the CNIG standard format which is completely new. Hence There is not yet an exchange method for these two formats.
- Data Compression: For the GPS data, the sampling interval is short, so large numbers of redundant points exist in GPS data, which makes the lines not smoothing after symbolization and hard to edit. So the redundant points need to be compressed, simplified and generalized according to the standards of certain scale. The common compression algorithms include: Ramer-Douglas-Peucker and vertical distance which is used to simplify a polygon or polyline by removing points. This method is theoretical and doesn't take in account the semantic associated to the graph, in our solution, we will combine semantics and geometry.

II.2.4. Data matching

Different data has different ways of expression, and data matching is just to identify the same feature in different datasets through similarity measurement of geometry, semantics and topology and build the corresponding relationships (Feng Xu, 2009).

In general, the matching process of vector data can be categorized in three ways: geometric, topological and semantic matching.

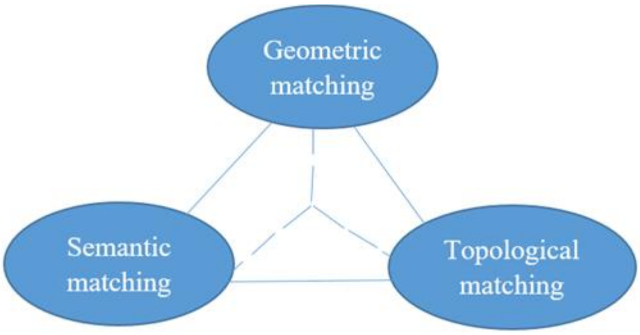

Figure 5 : Differentiation of matching

- Geometric similarity measurements are often based upon simple information like distance, length, angle or linearity (McMaster, 1986). Furthermore, combinations and indices, e.g. the Hausdorff-distance or Turning-Function, are used to describe and compare features to each other (Veltkamp and Hagedoorn, 2001).
- Topological similarity is often combined with geometric measurement and depends on the relationship of features. Therefore, approved approaches are the Spider-Function (Rosen and Saalfeld, 1985) and the Round-Trip-Walk (Filin and Doytsher, 2000), both depending on matched nodes and adjacent edges.
- Current research on similarity measurement often deals with semantic matching, usually based on thematic attributes, data structure or geo-ontologies (Weis and Naumann 2004, Giunchiglia et al., 2007).

All these methods are theoretical, hence our solution.

Furthermore, for different kinds of geographical entities, the corresponding matching algorithms are different. The point feature is comparative simple which is through location, structure and semantic similarity. For linear feature, the factors of location, shape, dimension, direction, and topology are considered synthetically (Xiaoya An 2011). While the areal feature seldom uses topology comparing to the linear feature. But the accurate rate of the matching algorithms at present is still has a long way to go, so we have to improve this in our project. Once the data matching is finished, the work of data fusion is relatively simple to develop.

II.2.5. Data fusion

Multi-source data integration eliminates differences and data matching builds the corresponding relationship of the same feature in different datasets, these processing don't change the existing data qualitatively and still keep their own characters as before. The data fusion is just the qualitative change step, which gathers the advantages of existing data extremely to obtain better new data through certain processing.

Spatial data fusion describes any kind of combination of spatial data from multiple sources. It targets either the generation of a new data source that is superior to any of the inputs or the extraction of meaningful information with regards to a specific application (J. Zhang 2010).

A spatial data relation is any kind of connection that exists between spatial data. This includes all associations between spatial datasets and between different feature views. Spatial data relations are essential to the fusion process, because they determine how data can be combined and which conclusions can be drawn from that combination.

The conventional way to solve a specific spatial data fusion task is to load the desired input data into a GIS software tool that offers corresponding processing capabilities. Basic fusion functionality is already provided by major GIS vendors, e.g. available in ArcGIS or the Feature Manipulation Engine (FME). More specialized tools include the Java Conflation Suite (Blasby et al. 2003), the Data Fusion System (Stankute and Asche 2012), FAGI-gis (Giannopoulos et al. 2015) or Hootenanny9. We also have QGIS which functions as Geographic Information System (GIS) software, allowing users to analyze and edit spatial information, in addition to compose and export graphical maps. Moreover, there is GDAL, a translator library for raster and vector geospatial data formats that is released under an X/MIT style Open Source License by the Open Source Geospatial Foundation. In our case, in order to perform the fusion of OSM and Geostandard data, we will use the tool POSTGIS which is a spatial database extender for PostgreSQL object-relational database. It adds support for geographic objects allowing location queries to be run in SQL.

II.2.6. Data coverage

A **coverage** is a mapping of one aspect of data in space. It represents a "domain" (the universe of extent) in terms of characteristics expressing a range of values.

"Coverage" is the term typically applied to the legacy ARC/INFO feature data format developed by Esri. The ARC/INFO coverage data model was a revolutionary concept, by storing graphical information with arc-node topology ("ARC") with features linked to attributes stored a simple, single-user RDBMS ("INFO"). In ARC/INFO, the BUILD and CLEAN commands and overlay tools such as INTERSECT and IDENTITY construct and maintain planar arc-node topology.

However, in our case, we don’t use the ARC/INFO feature data format, so we will study the data coverage between OpenStreetMap and the geostandard on different criteria:

II.2.6.1. Territory

OpenStreetMap has a global extent and covers all areas accessible to humans including airspace, aviation areas, borders, etc...

The geostandard focuses on collecting data for the mobility of people with disabilities and on the French territory. It covers only the areas and services accessible by PRM.

II.2.6.2. Degree of information

Geostandard takes into account more attributes and fields than OpenStreetMap. It includes fields for hearing aids, for the presence of visual, tactile and sound guidance systems, and even for PRM parking lots.

OpenStreetMap is less informative in this sense and just provides basic characteristics of places.

II.2.6.3. Degree of geographical accuracy

In OSM, the accuracy of the cadaster is of the order of one meter. Orthophotos used for plotting often have a precision of the order of 20 to 50 cm for a pixel.

The degree of geographical precision in the geostandard is 10 cm.

III. PROJECT PLANNING

III.1. Gantt diagram

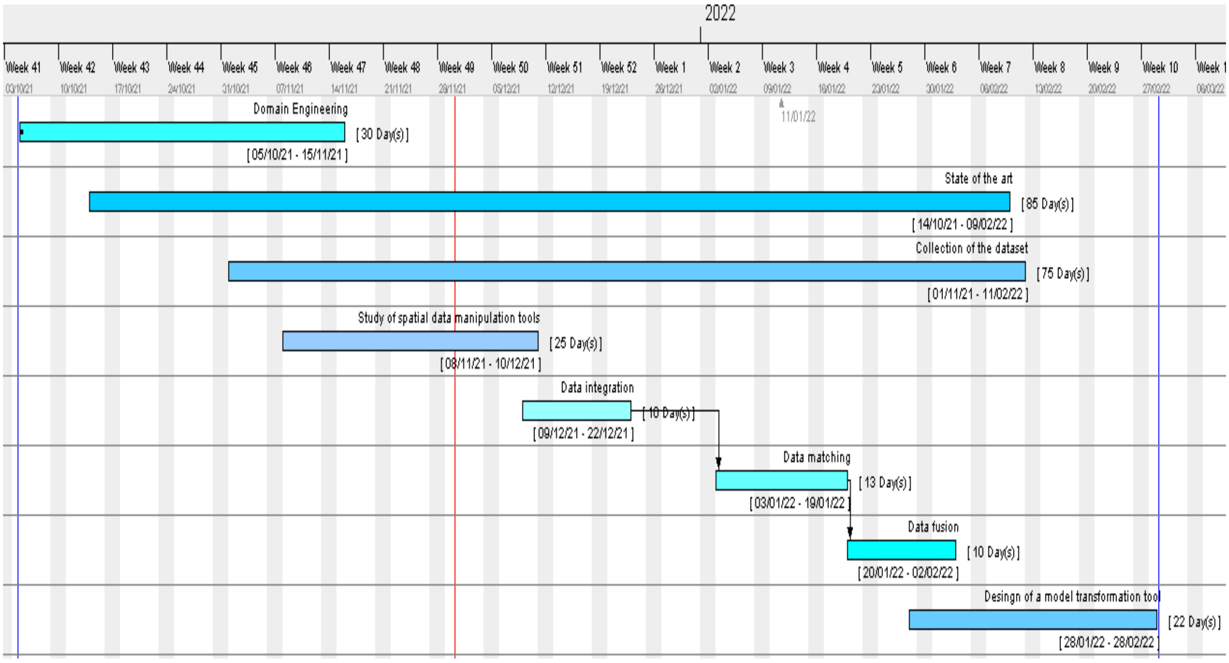

Figure 6 : *Gantt diagram*

III.2. Pert chart

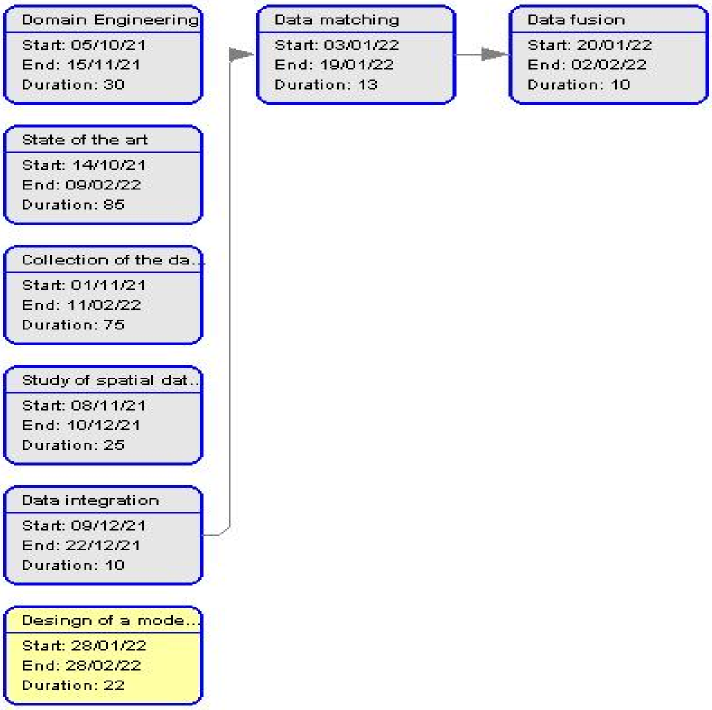

Figure 7 : *Pert chart*

IV. MATCHING STUDY BETWEEN OPENSTREETMAP AND CNIG

Here, we will study in detail the similarities and differences between OpenStreetMap data and CNIG data in terms of geometry and semantics.

IV.1. Geometry overview

Geometries, coordinates and references to coordinates, locate objects in space.

IV.1.1. Coordinate system

Geographers use coordinate systems to identify locations, and these systems use abscissa and ordinate coordinates. Points on the globe are often referenced by longitude (East-West) and latitude (North-South). For GPS receivers, the most common standard for determining the coordinate system and origin of a point is WGS84.

However, units of measure in degrees are not appropriate for measuring surfaces, shapes, distances and/or directions. Therefore, a projected coordinate system is generally used. This system allows to perform analyses on distances or angles and to obtain accurate measurements.

OpenStreetMap is based on GPS measurements and is therefore able to detect and read the WGS84 system. The projected coordinate system is WGS84/EPSG:4326 measured in meters.

Indeed, there are different types of projections that focus on different aspects of the world. Some projections are better suited to a small area, describing it more accurately, while others are more respectful of distances. Some projections preserve directions, while others better preserve shape. Distortion, or distorted representation of the earth, is always a problem. Because of concerns about distortion, particular projections, such as the Lambert Projection and the Universal Transverse Mercator (UTM) Projection, are used for small local areas, while other projections, such as Mercator, are better suited for global projections, while minimizing distortion.

OpenStreetMap is a world map, which justifies the use of the spherical Mercator projection for map rendering. The CNIG accessibility standard was created to represent the geographical data of local authorities, hence the use of the Lambert 93 projection.

Data from OpenStreetMap can be reprojected or converted to any projection system and vice versa. In this project, we will use QGIS software to bring both types of data to the EPSG:4326 coordinate system.

IV.1.2. Study of the geometries

In the data format, PostGIS allows several types of geometries. These types include point, line string, polygon, multipoint, multiline string, multipolygon and geometrycollection (L. Zhang and Yi 2010).

- POINT: a point represents a single location on the Earth. It is represented by a single coordinate (including either 2, 3 or 4 dimensions). Points are used to represent objects when exact contour detail is not important at a given scale.

Figure 8: *Point*

- LINESTRING: a line is a path between several points. It takes the form of an ordered array of two (or more) points. Roads and rivers are typically represented as lines. A line is said to be closed if it starts and ends at the same point. It is said to be simple if it does not intersect or touch itself (except at its ends if it is closed). A line can be both closed and simple.

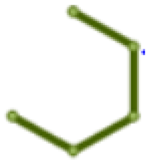

Figure 9: *Simple non-closed linestring*

- POLYGON: a polygon is represented as an area. The outer contour of the polygon is represented by a single closed line. Holes are represented in the same way.
  Polygons are used to represent objects that are large in size and shape. City limits, parks, buildings or rivers are usually represented by polygons when the scale is large enough to distinguish their areas. Roads and rivers can sometimes be represented as polygons.

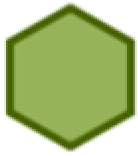

Figure 10: *Polygon defined by exterior ring*

- COLLECTION: there are four types of collections, which group together several simple geometries:
  - ✓ MultiPoint: a collection of points.

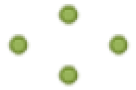

Figure 11: *Multipoint with 4 parts*

  - ✓ MultiLineString: a collection of lines.

Figure 12: *Simple multilinestring defined by 4 endpoints of 2 elements*

  - ✓ MultiPolygon: a collection of polygons

Figure 13: *Multipolygon consisting of two elements defined by exterior rings and three interior rings*

  - ✓ GeometryCollection: a heterogeneous collection of any geometry (including other collections).

In OpenStreetMap, geometries are defined by the data type (Node, Way, Relation). In the CNIG accessibility standard, geometries are explicitly specified for each type of object; there are three main types of geometry: point, linestring and surface.

The following table shows a correspondence study between the geometries of OpenStreetMap objects and those of the CNIG accessibility standard.

CNIG OBJECTS AND GEOMETRIES	OSM TAGS AND GEOMETRIES
Obstacle Geometry : Point	**Node** Geometry : Point
Troncon_cheminement Geometry : Linestring	**Ways** Geometry : Linestring
Nœud Geometry : Point	**Node** Geometry : Point
Circulation Geometry : Linestring	**Ways** Geometry : Linestring
Traversee Geometry : Linestring	**Ways** Geometry : Linestring
Rampe Geometry : Linestring	**Ways** Geometry : Linestring
Escalier Geometry : Linestring	**Ways** Geometry : Linestring
Escalator Geometry : Linestring	**Ways** Geometry : Linestring
Tapis roulant Geometry : Linestring	**Ways** Geometry : Linestring
Quai Geometry : Linestring	**Ways** Geometry : Linestring
Ascenceur Geometry : Point	**Node** Geometry : Point
Elévateur Geometry : Point	**Node** Geometry : Point
Entrée Geometry : Point	**Node** Geometry : Point
Passage sélectif Geometry : Point	**Absent**
Stationnement Pmr Geometry : Point	**Node** Geometry : Point
ERP Geometry : Surface	**Area** Geometry : Multipolygon

IV.2. Study of semantics of OSM and CNIG data

The aim here is to make a correspondence study on the semantic level between OpenStreetMap data and the CNIG accessibility standard data. To do so, we will first list the tags specific to disabled people or people with reduced mobility present in OpenStreetMap and their correspondences in the CNIG accessibility model. Then, for each object of the CNIG standard we will give the OSM tags equivalent to each of its attributes when they exist.

The appendixes used in this section were produced in collaboration with Cyril from the company Wegoto.

IV.2.1. OpenStreetMap accessibility model → CNIG accessibility standard

IV.2.1.1. Sidewalks

The presence of sidewalks in a street can be defined with a specific tag ("sidewalk = no/left/right/both") of the street: however, this simple way does not allow to store several attributes that are needed to quantify the accessibility of the sidewalks (Biagi, Brovelli, and Stucchi 2020).

A sidewalk can therefore be added as an independent way that is parallel to the street and intersect it and other streets in the crossings. The used tags are **"highway = footway"** combined with **"footway = sidewalk"**; it corresponds to the object ***"Circulation"*** in the CNIG accessibility standard. The tag **"width = m"**, in meters, provides the fundamental information on the sidewalk width; it corresponds to the attribute ***"largeurPassageUtile"*** in the CNIG accessibility standard.

The surface is an important aspect of sidewalk. it is described by two different keys: **"surface"** and **"smoothness"**. When referred to sidewalks, "surface" can assume the following values:

- "paved" just indicates that the way is paved with some material;
- "asphalt" means that is composed of asphalt concrete;
- "paving stones" means that it is covered by blocks well connected, and the surface is smooth;
- "sett" means that it is covered by natural stones, even with large gaps and a rougher surface than the previous one.

The tag surface corresponds to the attribute ***"typeSol"*** of the object ***"Circulation"*** in the CNIG accessibility standard.

Smoothness corresponds to the attribute ***"etatRevetement"*** of the object ***"Circulation"*** in the CNIG accessibility standard and really describes the accessibility of the sidewalks, since quantify the quality of the way. The values can be:

- “excellent”: for very new asphalt surface without holes and gaps; the way is suitable even for small wheels like skateboards;
- “good”: still an excellent accessibility for all the users, however, some imperfection can be present;
- “intermediate”: the quality of the paved is compromised, but the presence of some holes does not compromise the usability of a wheelchair user;
- “bad”: the way is not accessible to wheelchairs, due to the presence of wide or deep holes.

The longitudinal slope is another important to characterized the accessibility of sidewalks. The tagging in the OSM is possible by the keys **“incline”**. The tag has to be inserted on the way of the sidewalks or on the nodes of the kerbs. The value is expressed as a percentage “%”. it is positive or negative if the way is upward or downward respectively. The transverse inclination has also to be considered: it is almost imperceptible for ordinary people walking on the sidewalk but it can influence and create problems for wheelchairs, especially for electrics ones. It is quantified by the key **“incline:across”**; considering the direction of the way, it is positive for a downward slope toward the right, negative in the opposite case. The key “incline” corresponds to the attribute ***“pente”*** of the object ***“Troncon_Cheminement”*** in the CNIG accessibility standard.

IV.2.1.2. Crossing

A crossing is mapped as a separate way from the sidewalks that it connects: it can be represented with a new way with five nodes, as shown in Figure 14. This is the simplest case, but more complex cases are based on different combinations of it.

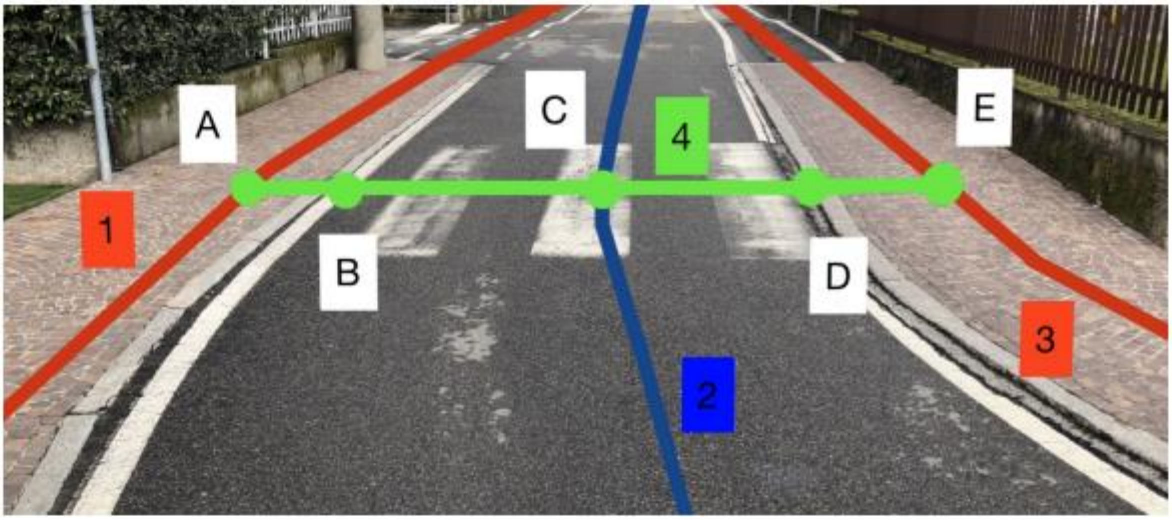

Figure 14: *Simple schema of the mapping of crossings (Biagi, Brovelli, and Stucchi 2020)*

In the white boxes, there are the letters of the nodes and in the red/blue boxes, the numbers of the ways.

The crossing (4) shares the first node (A) with the sidewalk (1); node (B) is simply the kerb from sidewalk (1) to the street, the node (C) is the connection between the crossing and the street (2). The node (D) is the kerb between the street and the other sidewalk (3), and finally (E) is the connection between the crossing and the sidewalk (3).

The different nodes require different tagging schema accordingly to their meaning. B and D are the nodes that represent the kerbs. Note that in OSM standards, kerbs with steps bigger than 3 cm are considered non accessible to wheelchairs and bicycles. They need the use of the key **"kerb"** with a value that can be:

- "raised": the step is more than 3 cm and does not guarantee accessibility;
- "lowered": the step is approximately 3 centimeters or less, and a structure (ramp) exist to ease the passage from the sidewalk to the street;
- "flush": there is not a significant step between the street and the sidewalk (Biagi, Brovelli, and Stucchi 2020).

In the CNIG standard, the OSM tag kerb is equivalent to the attribute ***"hauteurRessaut"*** of the object ***"Noeud"***. The CNIG accessibility standard takes into account only crossing for pedestrian mapped with the combination of the tags **"highway = footway"** and **"footway = crossing"** and it corresponds to the CNIG object ***"Traversée"***.

Appendix 1 shows the PAM tags which are attributes for people with visual impairments or other disabilities and their equivalents in the CNIG accessibility standard. In this case the elements of interest are the tactile paving and the sound traffic lights.

- The **tactile paving** is represented with the tag "tactile paving = yes": when present, it should be added on the sidewalks and on the nodes that represent the kerbs. It can take different values (yes, no, incorrect) indicating respectively the presence, absence or anomaly of the podotactile surfaces on either side of the crosswalk. This tag corresponds to the **"*bandeEveilVigilance*"** attribute of the object ***"Noeud"*** in the CNIG accessibility standard, when it indicates the presence or absence of a podotactile surface; and to the ***"controleBEV"*** attribute when the podotactile surface has an anomaly.
- For the sound traffic lights, the tag to be used is "traffic signals:sound = yes" on the way where the tag **"crossing = traffic signals"** is present. In some cases, the signals do not start automatically, but only by pushing a button: this situation can be mapped with the

tag “button operated = yes”. The different values of the crossing tag define the characteristics of pedestrian crossings specific to blind, autistic or reduced mobility people, and correspond to different attributes of the object ***“Traversée”*** in the CNIG accessibility standard.

IV.2.1.3. Obstacles

Barriers or obstacles are mapped in OSM as single nodes. It is better to do not add this node as part of the way of the sidewalks, but instead, a node closer to it since they are external elements and part of the sidewalks. The different types of obstacles have different tags:

- Traffic signals: use the key “traffic sign” with different values according to the signal, and the tag **“support = pole”** to indicate the presence of the pole;

  The height of traffic signals or advertising panels on sidewalks are also an important element, in particular for **blind people**. The tagging schema in OSM starts by creating the obstacle. Height can be tagged by the combination of **“support = pole”** and **“height = m”**, as shown in Figure 15. It corresponds to the attribute ***“hauteurObsPoseSol”*** of the object ***“Obstacle”*** in the CNIG model.

Figure 15: *Example of mapping the height of a traffic sign (Biagi, Brovelli, and Stucchi 2020)*

- Streetlight: use the tag **“highway = street lamp”** plus the tag **“support = pole”**;
- Storm drain: using the tag **“manhole = drain”** with the possibility of adding the size of the holes, if they are larger than 2 cm: this can be tagged with the key **“obstacle: description”**;

➢ Tree: described with the tag **“natural = tree”**.

All obstacles are modeled in the CNIG accessibility standard thanks to the object **“Obstacle”**.

Obstacles can represent a problem just for some specific type of users. This can be tagged with the key “obstacle: <type of transport>”, one example can be **“obstacle: wheelchair = yes”** for obstacles that stop wheelchairs (Biagi, Brovelli, and Stucchi 2020).

Appendix 2 presents the different possible values (yes, no, limited or bad) of the **wheelchair** tag, that help to define whether a crosswalk is accessible to people in wheelchairs or not. It corresponds to the attributes ***“hauteurRessaut”*** and ***“abaisseTrottoir”*** of the object ***“Nœud”*** in the CNIG accessibility standard. For each value of the wheelchair tag, the two nodes of the crossing in the CNIG model must be taken into account.

IV.2.1.4. Parkings

A parking for cars is mapped in OSM as an area with the tag “amenity = parking”. An important tag associated to it is the tag “capacity” and a particular subkey is “capacity: disabled” to specify the number of lots reserved for people with disability. Generally, this tagging schema is not sufficient to identify the reserved lots, in particular in huge parking areas. This issue can be solved inserting individually the reserved lots inside a parking area. A closed area is created and the tag **“amenity = parking space”** and **“parking space = disabled”** are associated; This corresponds to the object **“Stationnement_PMR**” in the CNIG model. Figure 16 shows how to map PMR parking spaces.

Figure 16: *Example of mapping the parking spaces reserved for people with disability(Biagi, Brovelli, and Stucchi 2020)*

IV.2.2. CNIG accessibility standard → OpenStreetMap tags

In this part, we carry out a study of correspondences between the objects of the CNIG standard and the OSM tags when they exist.

Appendix 3 associates to each attribute of the CNIG model objects an OSM tag when it exists with a quality of conversion, according to whether the CNIG attributes are easily converted into OSM tags or not. This appendix clearly shows the lack of equivalence of several CNIG attributes in the OpenStreetMap model and the presence of several attributes that are difficult to convert into OSM tags.

V. IMPLEMENTATION OF TRANSFORMATION MODELS

V.1. Study environment and tools

PostgreSQL is an enterprise-level database that is able to scale efficiently upon the demands of a single user or multiple users. By enabling spatial data storage with the PostGIS extension, PostgreSQL can store a wide variety of geometric objects, such as point, line polygons, multipoint, multiline, multipolygon, and geometric collections (Markieta 2012).

It is a powerful, open source object-relational database system with over 30 years of active development that has earned it a strong reputation for reliability, feature robustness, and performance.

PostGIS is an extension of the PostgreSQL DBMS, which enables the manipulation of geographic information in the form of geometries, in accordance with the standards established by the Open Geospatial Consortium. It allows PostgreSQL to be a spatial DBMS for use in geographic information systems.

PostGIS can also handle the reprojection of data as it is retrieved from the database, such that data can be stored in one common projection, but retrieved in a user-specified coordinate system (Markieta 2012).

DBeaver is a software that allows you to manage several different databases in a single application, and thus to manipulate data from different sources very quickly. It is based on the Eclipse framework, it is open source and it supports several types of database servers like: MySQL, SQLite, DB2, PostgreSQL, Oracle...

This software is used to connect to the GNIG database created in PostgreSQL and to structure it

QGIS is an open-source GIS software licensed under the GNU public license that can be used on Windows, Mac and Linux computer systems.

we used it to query, extract, and visualize the OpenStreetMap data and try to merge them with the PostgreSQL database. It is a true desktop GIS that can perform advanced geospatial analysis and has a wide range of plugins (>300).

QGIS provides a vast array of functionality which makes it a competitor of commercial solutions. Also, due to its open-source roots, users will find direct compatibility with many other open-source projects such as the PostgreSQL database. QGIS also provides powerful Python scripting capabilities, making it a viable option for automating spatial workflows. Lastly, the QGIS community has a large and growing plugin repository where users can find useful tools or scripts that enhance the usability and efficiency of QGIS (Markieta 2012).

JOSM is an editor for OpenStreetMap (OSM) written in Java. The current version supports standalone GPX tracks, GPX track data from OSM database and existing nodes, line segments and metadata tags from the OSM database.

we used this tool to study the geometry of OSM data.

V.2. Creation of the CNIG database

- The CNIG database geosm1 is created in command line in Psql with the command:

```
create database geosm1;
```

- Later on, the POSTGIS extension is added for the manipulation of spatial data with the commands:

```
\c geosm1
create extension postgis;
```

- The creation of tables and constraints is done on dbeaver: to do this, you have to connect to the database through the dbeaver interface and create the tables either by writing SQL queries or in graphical mode. This is how we obtain the CNIG database show in Figure 17.

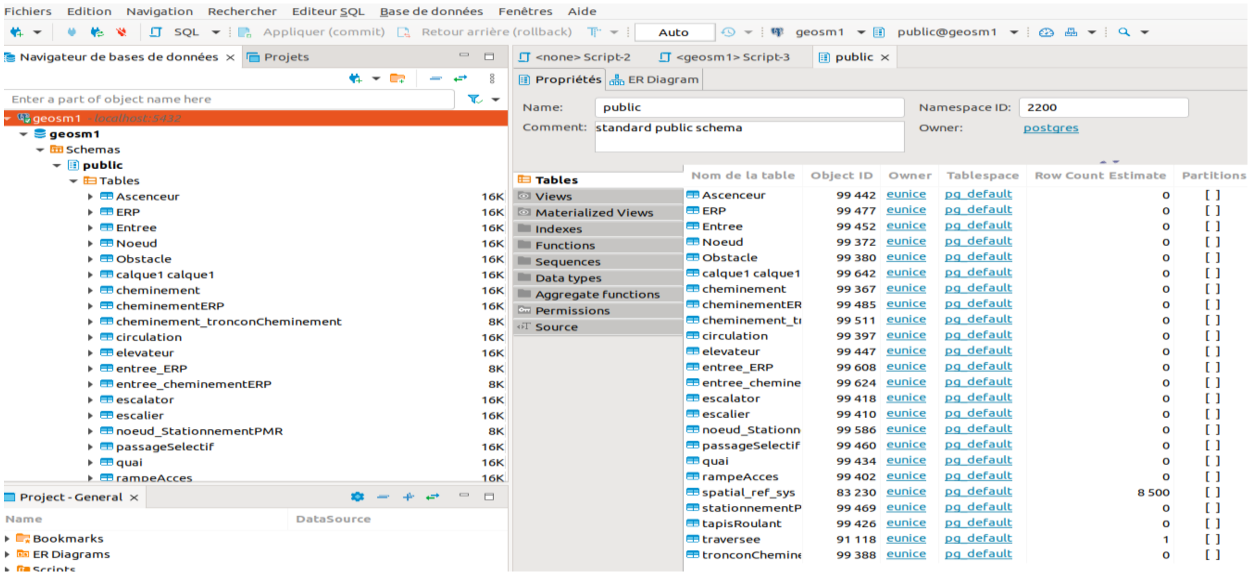

Figure 17: CNIG database

V.3. Fusion of the CNIG database with OSM data

V.3.1. Connection to the CNIG database and loading of OSM accessibility data

The insertion of OpenStreetMap data in the CNIG database is done through the following steps:

- Connect to the CNIG database geosm1 on QGIS;
- Load the OpenStreetMap layer and thanks to the Gban extension (previously installed), search for the area you are interested in as shown in Figure 18 and zoom on the red dot that appears. Here we want to download the OSM data from Orange.

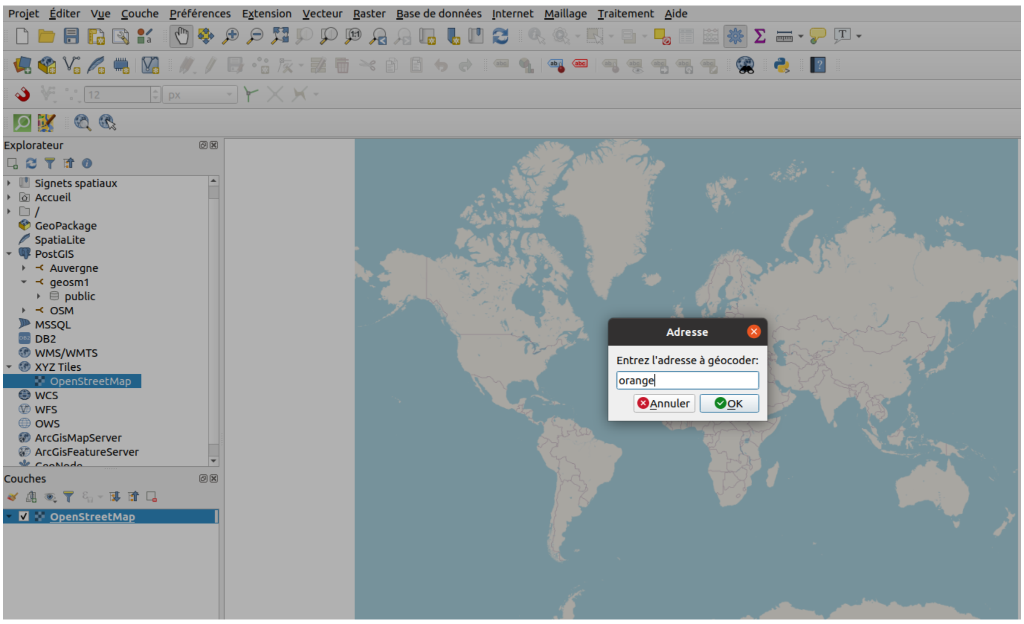

Figure 18: Loading OSM data from Orange

➢ Click on the magnifying glass of the QuickOSM extension (previously installed), then, in the Query tab write and execute the overpass query as shown in Figure 19 to retrieve the specific OpenStreetMap data you are interested in on the territory previously loaded: this will create new layers. Here we want to download the Orange footways and their nodes.

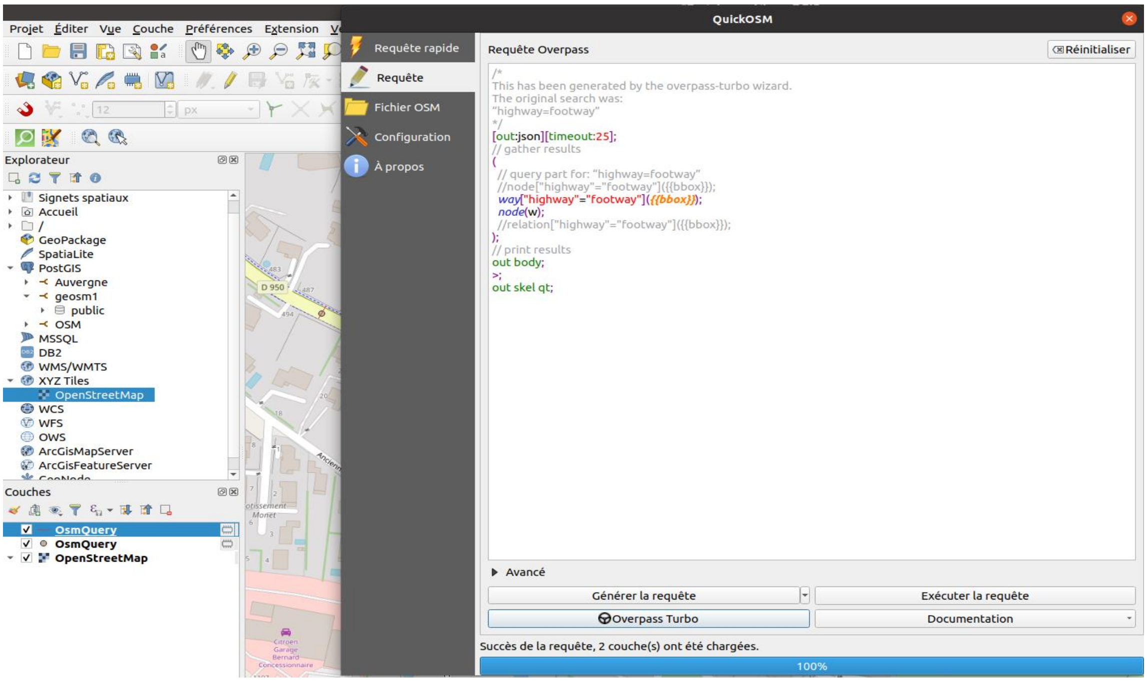

Figure 19: Download of Orange's footways

V.3.2. Design of OpenStreetMap data transformation models

The purpose here is to present the algorithms for processing and transforming OpenStreetMap data in order to prepare their fusion with CNIG data.

To transform and integrate the OSM data, in particular the path sections, their nodes and their obstacles in the CNIG database, the following 03 models must be executed in order.

V.3.2.1. Model for the generation of pedestrian paths

V.3.2.1.1. Presentation of the model

This model allows to obtain path sections and their start and end nodes as modeled in the CNIG accessibility standard from the raw data of OSM. Figure 20 illustrates the different stages of data transformation.

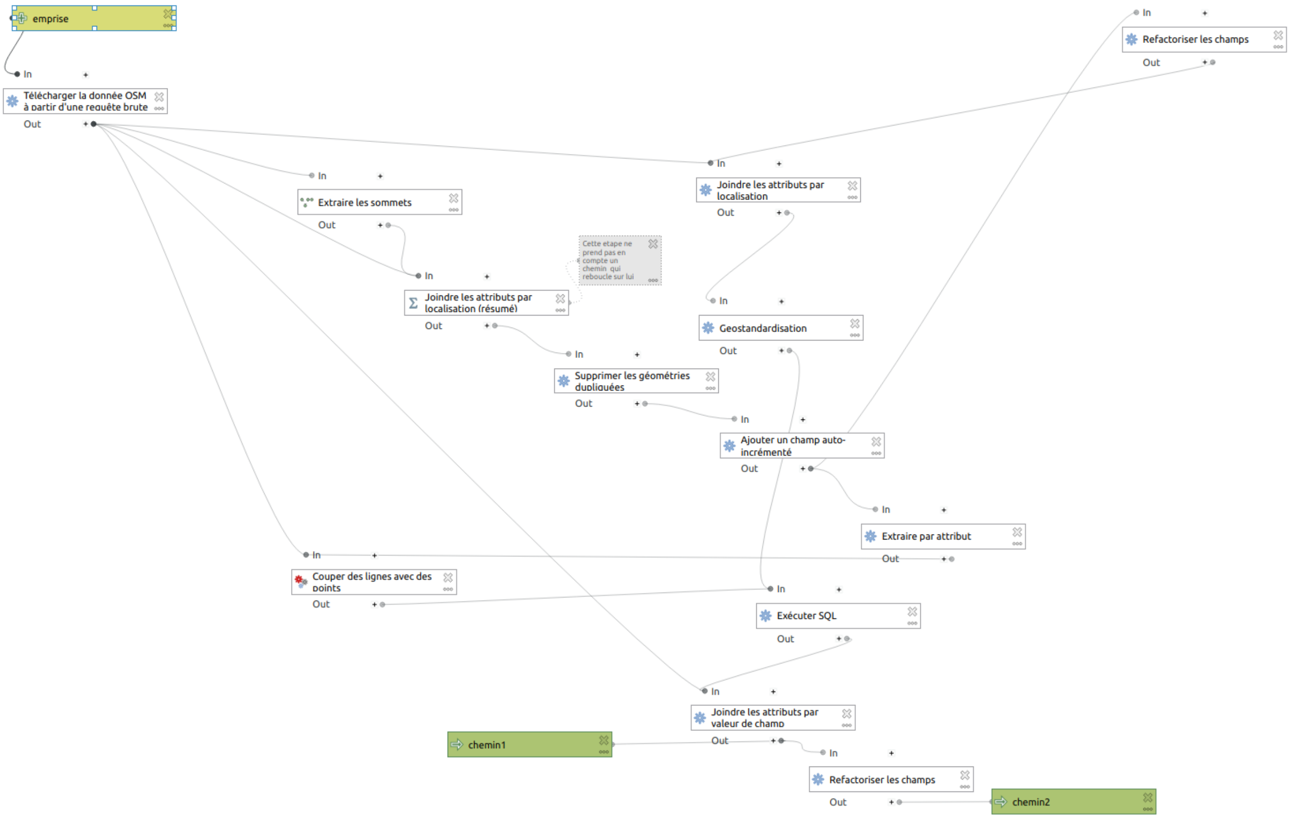

Figure 20: *Generation of pathway trunks*

V.3.2.1.1. Description of the model

- ➢ **Step 1:** we start by defining the OSM territory on which the model will be applied thanks to the "emprise" entry which allows to take as input of the model the canevas's emprise on the OSM territory previously loaded in QGIS;
- ➢ **Step 2: download of the OSM pedestrian paths**

To do this, we apply the algorithm download OSM data from a raw query. It allows to insert the Overpass query to download the areas corresponding to the footpaths ("highway"="footway") as shown in figure 21.

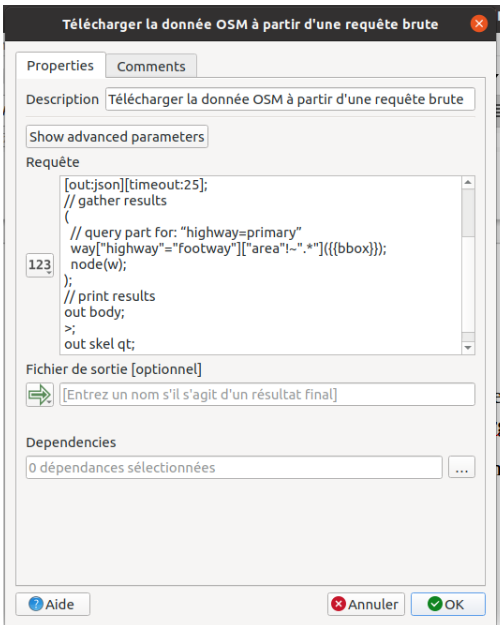

Figure 21: *Downloading pedestrian paths on OSM*

➢ **Step 3: extraction of path nodes**

Once the pedestrian paths are downloaded, we apply the Extract vertices algorithm which takes in these paths and extracts all the path nodes.

In fact, the treatment Extract vertices takes a vector layer and generates a point layer with points representing the vertices in the input geometries. The attributes associated to each point are the same ones associated to the feature that the vertex belongs to. Additional fields are added to the vertices indicating the vertex index (beginning at 0), the feature's part and its index within the part (as well as its ring for polygons), distance along original geometry and bisector angle of vertex for original geometry. Figure 22 shows how the vertices are extracted from a line and a polygon layer.

Figure 22*: Vertices extracted for line and polygon layer ("Géométrie Vectorielle — Documentation QGIS Documentation" n.d.)*

- **Step 4 & 5: join nodes and remove duplicates**

Now we need to join the paths from the Download OSM data from a raw query algorithm and their vertices previously extracted with the Extract vertices algorithm based on their OSM id. Then apply the Delete duplicate geometries algorithm to remove duplicates.

- **Step 6: add an automatic increment field**

It allows to add a new integer field to the vector layer resulting from the previous step, with a sequential value for each entity. This field will be used as a unique ID for the entities of the layer.

- **Step 7: extract by attribute**

Create two vector layers from an input layer: in our case the layer from the Add auto incremented field algorithm; one will contain only entities matching the selection attribute 'osm_id_count > 1' while the second will contain all non-matching entities.

- **Step 8: make path sections**

Once the relevant path nodes are obtained, we can now make path segments by cutting the pedestrian paths downloaded from OSM at the extracted nodes obtained in the previous step. To do this, we apply the Cut lines with points algorithm.

The Cut lines with points algorithm is a processing model that takes as input a layer of lines and a layer of points and produces as output cut lines as shown in figure 23.

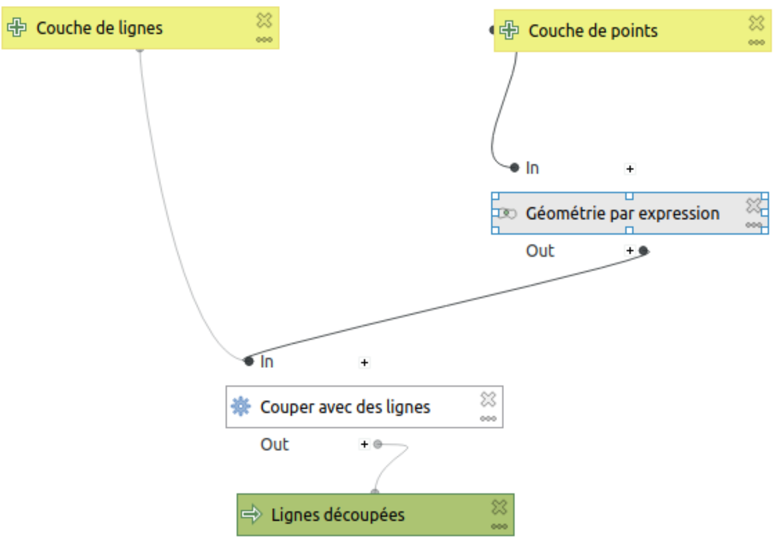

***Figure 23:** Model cut lines with points*

To be able to cut the lines, we have a QGIS treatment Cut with lines, so it is necessary to transform the layer of points into a layer of lines. Therefore, the model first transforms the layer of points into small lines similar to the points by means of a QGIS geometry expression provided by the Geometry by expression algorithm as can be seen in figure 24.

Figure 24: *Transforming points into lines*

Now that the points have been transformed into thin lines, we can unambiguously apply the QGIS processing Cut with lines as seen in figure 25.

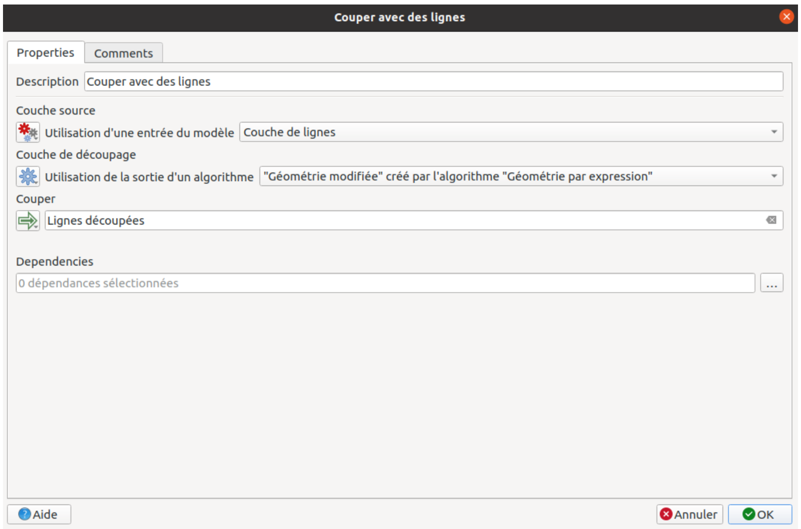

Figure 25: *Cut with lines*

- **Step 9: refactor the fields**

Allows editing the structure of the attribute table of a vector layer. Fields can be modified in their type and name, using a fields mapping. The original layer is not modified. A new layer is generated, which contains a modified attribute table, according to the provided fields mapping.

Here we refactor the fields of the vector layer from the Add an auto-incremented field algorithm in order to clear all the fields to join them later with the footpath nodes downloaded from OSM.

- **Step 10: join attributes by location**

Takes the input vector layer from the Refactor fields algorithm and creates a new vector layer that is an extended version of the input one, with additional attributes in its attribute table.

The additional attributes and their values come from a second vector layer: those of the nodes from the Download OSM data from a raw query algorithm. A spatial criterion is applied to select the values from the second layer that are added to each feature in the first layer.

- **Step 11: refactor the fields to model them as the path nodes of the CNIG accessibility standard**

Here, we refactor the fields of the vector layer from the Join attributes by summary location algorithm by adding the fields of the object *"Noeud"* of the CNIG accessibility standard and making correspondences from the OSM tag values as shown in figure 26. By doing this we get as output nodes as represented in the CNIG accessibility standard in the layer nodes.

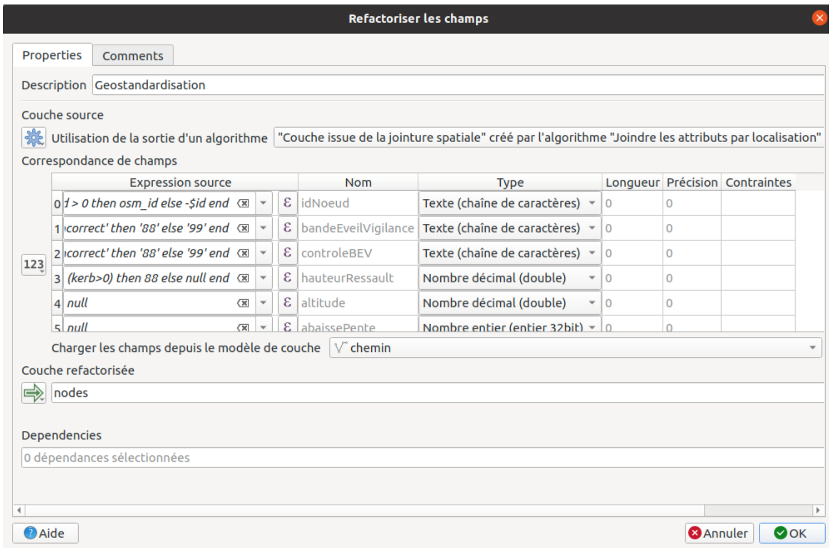

Figure 26: *Geostandardization of nodes*

- **Step 12: select initial and final path nodes**

Here we need to run an sql script to select the initial and final nodes from the vector layers resulting from the Geostandization and the Cut lines with points algorithm as shown in figure 27.

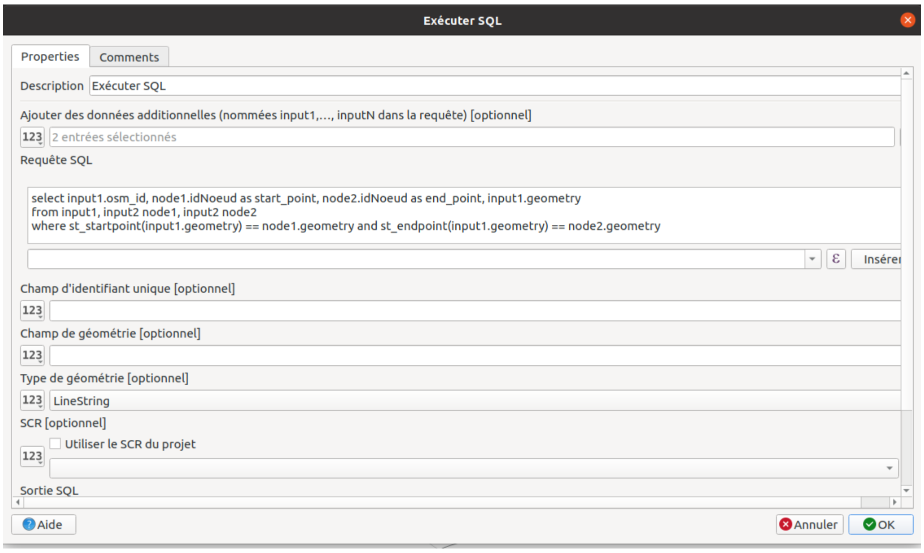

Figure 27: Selection of initial and final path nodes

➢ **Step 13: create path sections of the CNIG accessibility standard**

Now that we have the OSM pedestrian paths and the initial and final path nodes, we can generate the path sections corresponding to the CNIG accessibility standard; to do this, we first perform a join between the line vector layer obtained from the Download OSM data from a raw query algorithm and the one obtained by the Execute SQL algorithm, as shown in figure 28. Then, we refactor the fields to add the fields corresponding to the path sections of the CNIG accessibility standard as shown in figure 29.

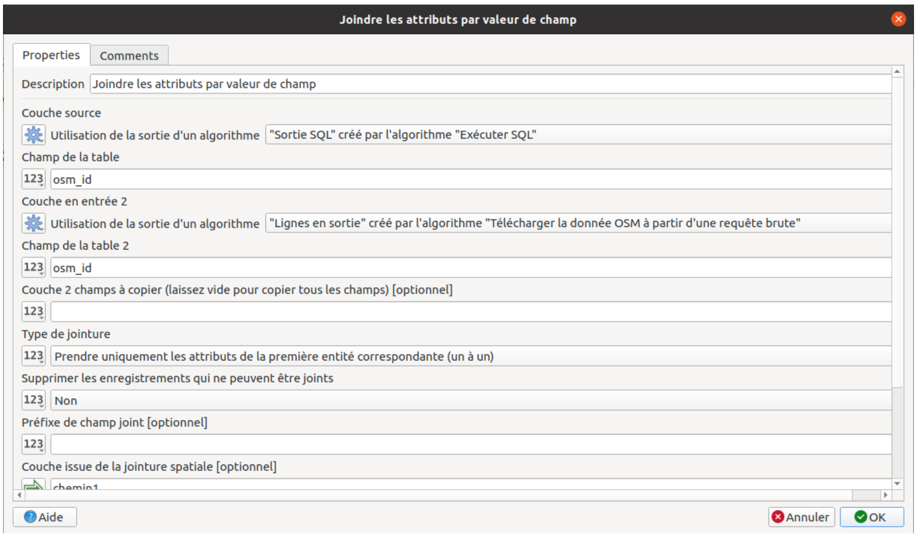

Figure 28: Joining of pedestrian paths with initial and final nodes

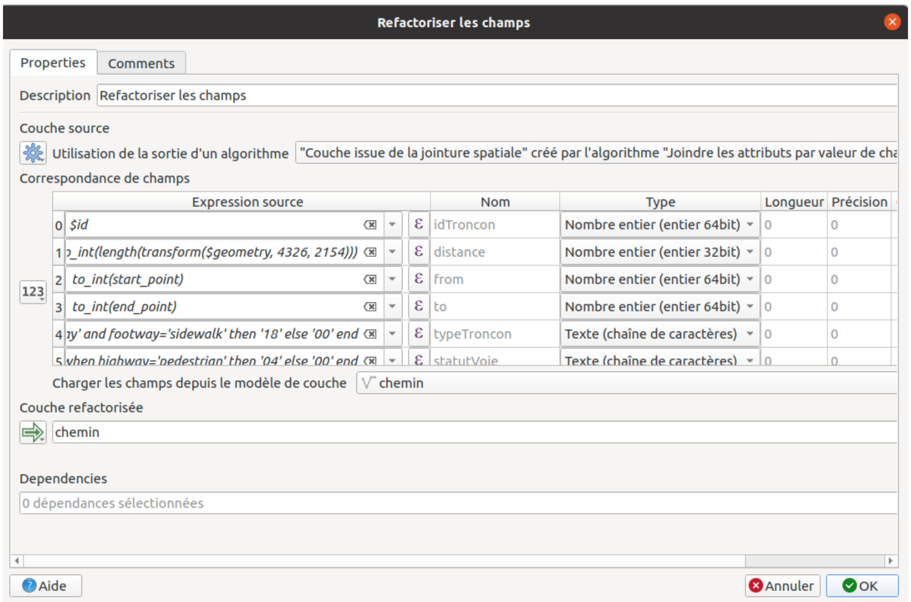

Figure 29: *Creation of the geostandard path trunks*

V.3.2.2. Model for the generation and attachment of obstacles

V.3.2.2.1. Presentation of the model

This model allows to download obstacles from OSM and link the ones located at a distance of 50 cm or less from the path trunks to the trunks. After, it applies some treatments in order to get as output obstacles as represented in the CNIG accessibility standard as shown in figure 30.

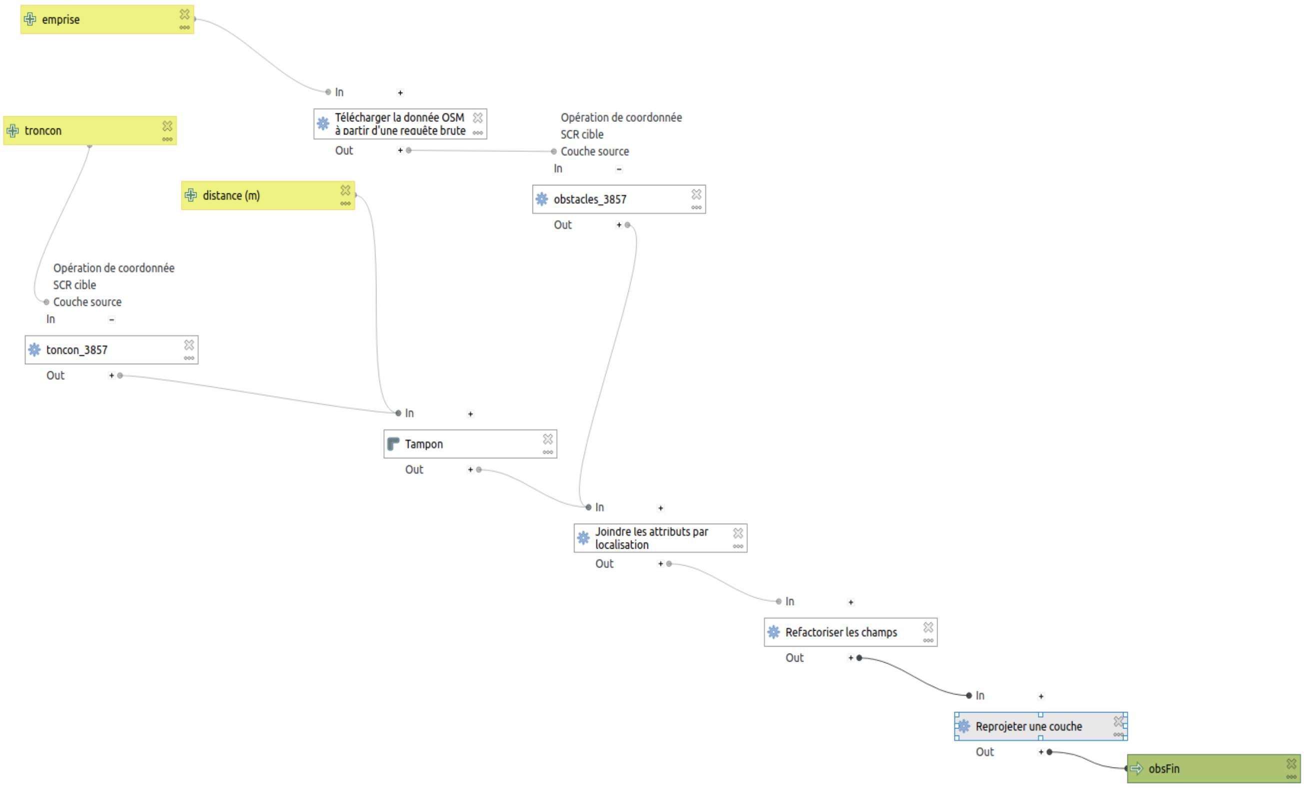

Figure 30: *Generate obstacles of pathways in the CNIG format*

V.3.2.2.2. Description of the model

- First we download possible obstacles from OSM by using the Download OSM data from a raw query algorithm as shown in figure 31.

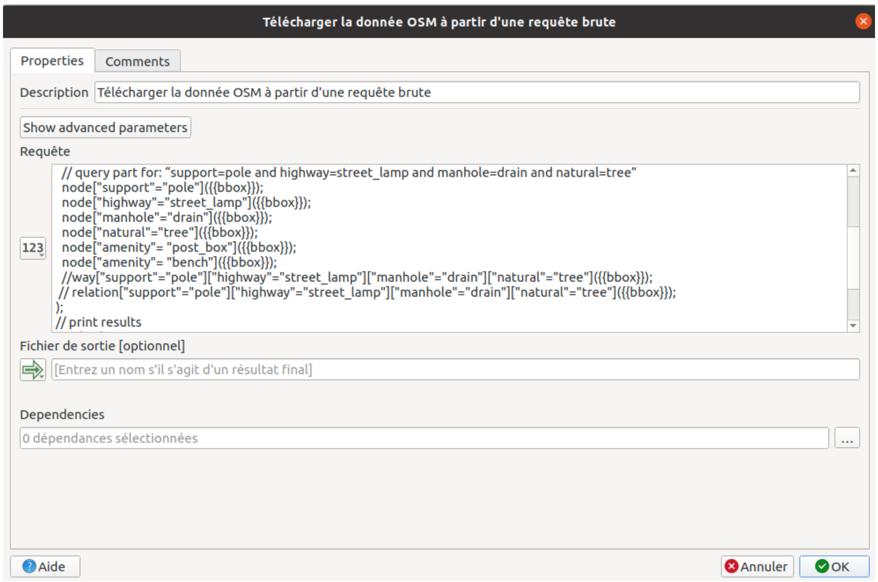

Figure 31: *Downloading of obstacles in OSM*

- Then vector layers *"troncon"* and the one obtained by the previous algorithm are re-projected into the SCR EPSG 3857 using the Reproject layer algorithm. The re-projected layers will have the same characteristics and attributes as the input layers. We needed to reproject them in order to apply the Buffer algorithm.
- The Buffer algorithm creates a buffer of fixed width (50 cm) for each feature in the *"troncon"* layer.
- After, the Join attributes by location treatment is applied to attach the obstacles which are located at a distance of 50 cm or less from the path trunks to the created buffer.
- Now that we got relevant obstacles, we can apply the Refactor fields algorithm to add fields corresponding to the obstacles in the CNIG accessibility standard as shown in figure 32.

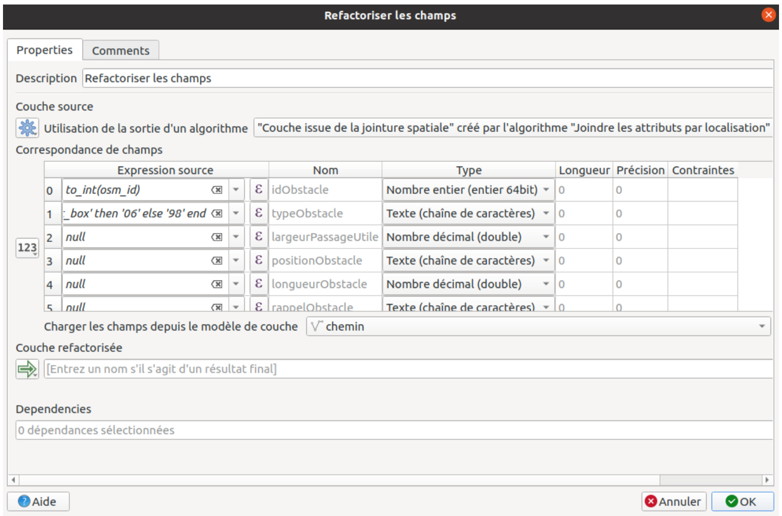

Figure 32: *Geostandardization of obstacles*

- ➢ The final step is to re-project the layer into the SCR EPSG:4326 to get the final result.

V.3.2.3. Model for integrating OSM data into the CNIG database

V.3.2.3.1. Presentation of the model

This model allows the integration of nodes, obstacles and path sections in the QGIS geosm1 database of the CNIG accessibility standard as illustrated in the figure 33.

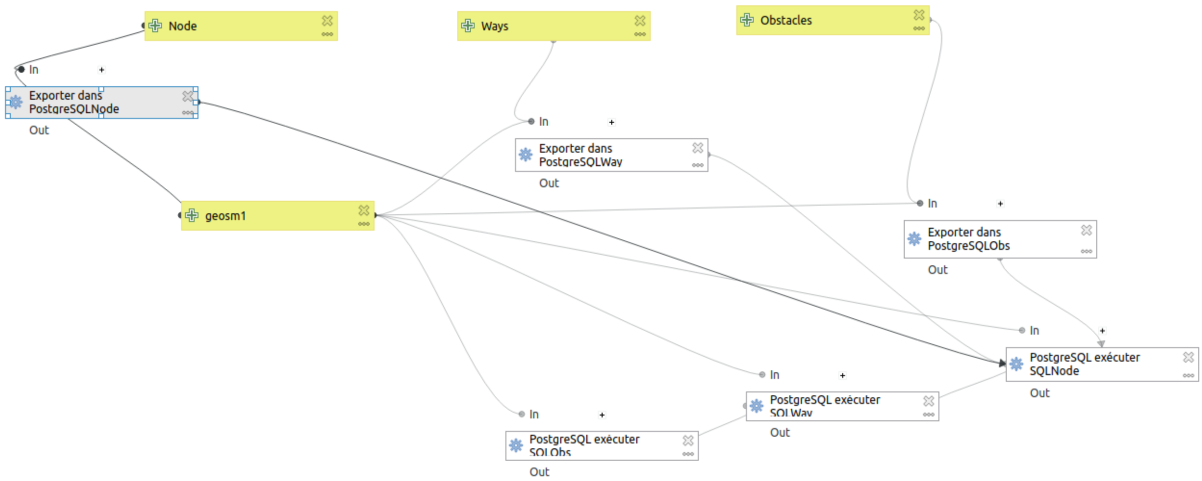

Figure 33: *Model for integrating OSM data into the CNIG database*

V.3.2.3.2. Description of the model

The model first exports the Node, Ways and Obstacles vector layers to the geosm1 database by applying the algorithms Export to PostgreSQLNode, Export to PostgreSQLWays, and Export to PostgreSQLObs. These algorithms create the node_tmp, ways_tmp, and obstacles_tmp tables respectively.

Figure 34 shows the Export to PostgreSQLNode algorithm with the creation of the new node_tmp table.

Figure 34: Exportation of nodes into the node_tmp table

Then, the PostgreSQL execute SQLNode, PostgreSQL execute SQLWay and PostgreSQL execute SQLObs algorithms are applied to insert the data exported in the node_tmp, ways_tmp and obstacles_tmp tables, respectively, into the *"Node"*, *" tronconCheminement"* and *"Obstacle"* tables, respectively, of the CNIG geosm1 database.

Figure 35 illustrates the PostgreSQL running SQLNode algorithm, with the SQL query used to integrate the data from the node_tmp table into the *" Nœud "* table of the geosm1 database.

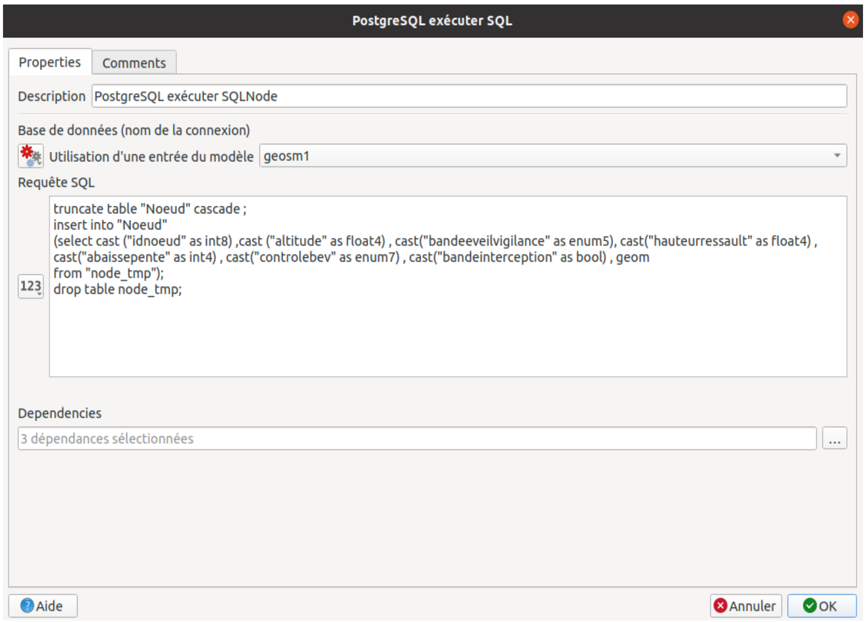

Figure 35: Insertion of nodes data into the "Noeud" table of the CNIG database

V.3.3. Presentation of the results

To simulate our models, we used an area of OSM where the pedestrian paths are well documented, notably a part of the Orange territory in France. However, for the values that lacked precision in OpenStreetMap, we created a value notably "88" to stipulate that the OSM tag lacks precision.

The models generated 225 pedestrian nodes, 03 obstacles and 110 ways as you will see in figures 36, 38, and 40.

By the way, all the work is available on Github ("Commits - Eunicewadjom/OSMCNIGProject" n.d.), including the CNIG database created and the processing models.

V.3.3.1. Path nodes

V.3.3.1.1. Graphical rendering

Figure 36: *Pathway nodes represented according to the CNIG accessibility standard*

V.3.3.1.2. Table of attributes

	idNoeud	altitude	ndeEveilVigilar	auteurRessaul	abaissePente	controleBEV	ndeInterceptio
1	884891261	NULL	01	NULL	NULL	99	false
2	884891290	NULL	01	NULL	NULL	99	false
3	884891375	NULL	88	NULL	NULL	88	false
4	1362811824	NULL	99	NULL	NULL	99	false
5	1362811826	NULL	99	NULL	NULL	99	false
6	1362811829	NULL	99	NULL	NULL	99	false
7	1425163601	NULL	99	NULL	NULL	99	false
8	1425163615	NULL	99	NULL	NULL	99	false
9	1425163628	NULL	99	NULL	NULL	99	false
10	1548644675	NULL	99	NULL	NULL	99	false
11	1548645482	NULL	99	NULL	NULL	99	false
12	1557164030	NULL	02	NULL	NULL	01	false
13	1557164060	NULL	01	NULL	NULL	99	false
14	1557164104	NULL	99	NULL	NULL	99	false
15	1557164107	NULL	01	NULL	NULL	99	false
16	1557164113	NULL	01	NULL	NULL	99	false
17	1557164114	NULL	01	NULL	NULL	99	false
18	2178519966	NULL	01	NULL	NULL	99	false
19	2282204510	NULL	99	NULL	NULL	99	false
20	2705440388	NULL	01	NULL	NULL	99	false
21	2705440392	NULL	01	NULL	NULL	99	false
22	3012267989	NULL	01	NULL	NULL	99	false
23	3012267997	NULL	02	NULL	NULL	01	false
24	3012268009	NULL	01	88	NULL	99	false

Figure 37: *Attribute values of pathway nodes according to the CNIG accessibility standard*

V.3.3.2. Obstacles

V.3.3.2.1. Graphical rendering

Figure 38: *Obstacles represented according to the CNIG accessibility standard*

V.3.3.2.2. Table of attributes

	idObstacle	typeObstacle	jeurPassageU	ositionObstacl	ngueurObstac	appelObstacle	erabiliteVisue	argeurObstacl	uteurObsPose	auteurSousOb	idtroncon
1	2657856639	07	NULL	NULL	NULL	NULL	false	NULL	NULL	NULL	95
2	8309711922	07	NULL	NULL	NULL	NULL	vrai	NULL	NULL	NULL	94
3	8711524066	07	NULL	NULL	NULL	NULL	false	NULL	NULL	NULL	96

Figure 39: *Attribute values of obstacles according to the CNIG accessibility standard*

V.3.3.3. Pathway sections

V.3.3.3.1. Graphical rendering

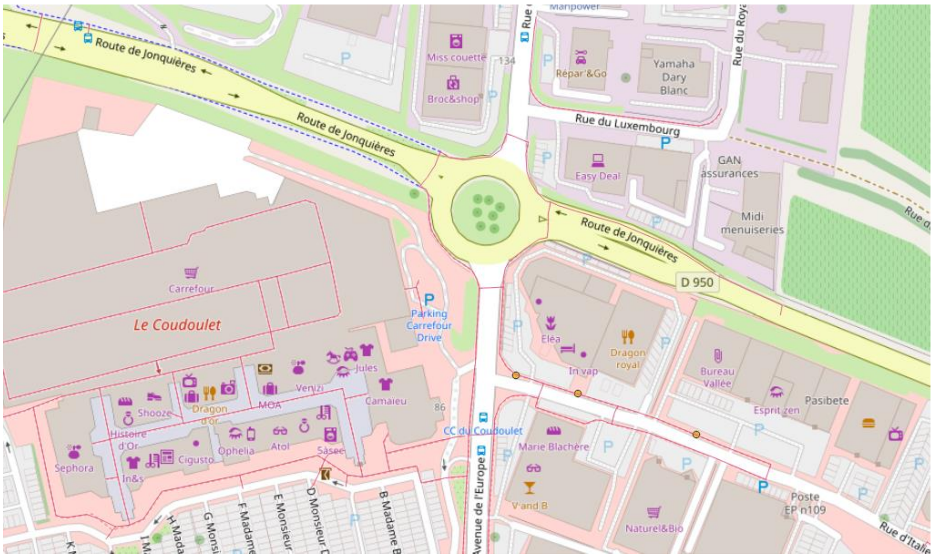

Figure 40: *Pathway sections represented according to the CNIG accessibility standard*

V.3.3.3.2. Table of attributes

	idTroncon	from	to	distance	typeTroncon	statutVoie	pente	devers	essibiliteGlob
1	1	-93	1362811824	0	00	00	NULL	NULL	NULL
2	2	-102	-84	0	00	00	NULL	NULL	NULL
3	3	-84	-75	0	00	00	NULL	NULL	NULL
4	4	-75	-79	0	00	00	NULL	NULL	NULL
5	5	-80	-102	0	00	00	NULL	NULL	NULL
6	6	-82	-79	0	00	00	NULL	NULL	NULL
7	7	-86	-88	0	00	00	NULL	NULL	NULL
8	8	-88	1362811829	0	00	00	NULL	NULL	NULL
9	9	-56	-50	0	00	00	NULL	NULL	NULL
10	10	-50	-75	0	00	00	NULL	NULL	NULL
11	11	-91	-103	0	00	00	NULL	NULL	NULL
12	12	-109	-107	0	00	00	NULL	NULL	NULL
13	13	-103	-93	0	00	00	NULL	NULL	NULL
14	14	-93	-109	0	00	00	NULL	NULL	NULL
15	15	-64	-54	0	00	00	NULL	NULL	NULL
16	16	-54	-95	0	00	00	NULL	NULL	NULL
17	17	-95	-89	0	00	00	NULL	NULL	NULL
18	18	-89	-91	0	00	00	NULL	NULL	NULL

Figure 41*: Attribute values of pathway sections according to the CNIG accessibility standard*

CONCLUSION AND PERSPECTIVES

The movement of people with reduced mobility or in a situation of disability is, until now, a real obstacle course. The solutions developed so far do not take into account all the territories and are proprietary. The only way to guarantee accessibility to all is to improve the level of accuracy of OpenStreetMap which is public by finding a way to merge OSM data with CNIG data that will be produced by each community in application of the LOM (Mobility Orientation Law).

In this project, we provided models for the transformation of OSM data into CNIG data. These models can help communities when collecting CNIG data, by providing them with CNIG data already represented in OpenStreetMap and are a good basis for merging OSM data with CNIG data.

However, it should be pointed out that there are areas on OpenStreetMap where pedestrian paths are not represented, so the question arises as how to make the paths in this case? One answer would be to create pedestrian paths parallel to a highway; this would be an excellent perspective to our work.

REFERENCES

An Xiaoya, Sun Qun, Xiao Qiang, 2011. A shape multilevel description method and applicationin measuring geometry similarity of multi-scale spatial data. *Acta Geodaetica et Cartographica Sinica*, 40(4), pp. 495-501.

Antu and Jilin 2013, "International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences", Volume XL-7/W1, 3rd ISPRS IWIDF 2013.

Blasby, D., Davis, M., Kim, D., and Ramsey, P. (2003). GIS Conflation using Open Source Tools. Tech. rep. The Jump Project, p. 32.

Biagi, L., M. A. Brovelli, and L. Stucchi. 2020. "MAPPING THE ACCESSIBILITY IN OPENSTREETMAP: A COMPARISON OF DIFFERENT TECHNIQUES." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B4-2020 (August): 229–36. https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-229-2020.

Brovelli, Maria Antonia, Maria Ilie Codrina, and Serena Coetzee. 2019. "Openness and Community Geospatial Science for Monitoring SDGs – An Example From Tanzania." In *Sustainable Development Goals Connectivity Dilemma*. CRC Press.

Chen, H., Q. Sun, L. Xu, and Z. Xiong. 2013. "Application of data fusion in the production and updating of spatial data." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-7/W1 (July): 5–10. https://doi.org/10.5194/isprsarchives-XL-7-W1-5-2013.

CEREMA. (2021). "CNIG Accessibility Standard "

"Commits · Eunicewadjom/OSMCNIGProject." n.d. GitHub. Accessed February 27, 2022. https://github.com/eunicewadjom/OSMCNIGProject.

Filin, S., Doytsher, Y. (2000). The Detection of Corresponding Objects in a Linear-Based Map Conflation. Surveying and Land Information Systems, 60 (2000), Nr. 2. 117–128.

"Géométrie Vectorielle — Documentation QGIS Documentation." n.d. Accessed February 24, 2022. https://docs.qgis.org/3.4/fr/docs/user_manual/processing_algs/qgis/vectorgeometry.html#extract-vertices.

Giannopoulos, G. et al. (2015). "FAGI-gis: A Tool for Fusing Geospatial RDF Data". In: The Semantic Web: ESWC 2015 Satellite Events. Ed. by F. Gandon et al. Vol. 7540. Lecture Notes in Computer Science. Springer International Publishing, pp. 51–57. doi:10.1007/978-3-319-

25639-9_10.

Giunchiglia, F., Yatskevich, M., Shvaiko, P. (2007). Semantic Matching: Algorithms and Implementation. In Spaccapietra, S. (ed.). Journal on Data Semantics IX, Springer-Verlag, Berlin - Heidelberg. 1–38.

McMaster, R. B. (1986). A Statistical Analysis of Mathematical Measures for Linear Simplification. The American Cartographer, 13 (1986), Nr. 2. 103–116.

Mericskay, Boris. 2017. “Introduction Aux Données OpenStreetMap (Structuration, Interrogation, Extraction et Édition).” Master. France. https://hal.archives-ouvertes.fr/cel-01660629.

OSM Wiki. OpenStreetMap Relation Type-Public Transport. Availableonline: http://wiki.openstreetmap.org/wiki/Relation:public_transport (accessed on 05th November 2021).

Pruvost, Hippolyte, and Peter Mooney. 2017. “Exploring Data Model Relations in OpenStreetMap.” *Future Internet* 9 (4): 70. https://doi.org/10.3390/fi9040070.

Rosen, B., Saalfeld, A. (1985). Match criteria for automatic alignment. Proceedings of AUTOCARTO 7, March 1985. 456–462.

Schmitz, Bernhard, and Thomas Ertl. 2013. “Creating Task-Specific Maps with Map Content Transformations.” In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on MapInteraction*, 84–90. MapInteract ’13. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2534931.2534945.

Stankute, S. and Asche, H. (2012). “A data fusion system for spatial data mining, analysis and improvement”. In: Computational Science and Its Applications – ICCSA. Ed. by B. Murgante et al. Vol. 7334. Lecture Notes in Computer Science. Springer Berlin Heidelberg,pp. 439–449.

Veltkamp, R. C.; Hagedoorn, M. (2001). State of the Art in Shape Matching. In Lew, M. S. (ed.). Principles of visual information retrieval. Springer Verlag, London. 87–119.

Weis, M., Naumann, F. (2004). Detecting Duplicate Objects in XML Documents. Proceedingsof the 2004 international workshop on Information quality in information systems, Paris. 10–19.

Xu Feng, Deng Min, Zhao Binbin, Chen Jian-jun, 2009.A detailed investigation on the methodsof object matching. *Journal of Geo-Information Science*, 11(5), pp. 657-663.Biagi, L., M. A. Brovelli, and L. Stucchi. 2020. “MAPPING THE ACCESSIBILITY IN OPENSTREETMAP: A COMPARISON OF DIFFERENT TECHNIQUES.” *The International Archives of the*

Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B4-2020 (August): 229–36. https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-229-2020.

Zhang, Jixian. 2010. "Multi-Source Remote Sensing Data Fusion: Status and Trends." *International Journal of Image and Data Fusion* 1 (1): 5–24. https://doi.org/10.1080/19479830903561035.

Zhang, Lijing, and Jing Yi. 2010. "Management Methods of Spatial Data Based on PostGIS." In *2010 Second Pacific-Asia Conference on Circuits, Communications and System*, 1:410–13. https://doi.org/10.1109/PACCS.2010.5626962.

APPENDIXES

Appendix 1: PAM OpenStreetMap TAGS → CNIG Objects attributestactile_paving 51
Appendix 2: Wheelchair accessibility tag → CNIG Objects attributes 53
Appendix 3: CNIG Objects attributes → OSM tags 54

Appendix 1: PAM OpenStreetMap TAGS → CNIG Objects attributestactile_paving

crossing

	PAM TAGS	Values	Description	Pictures	CNIG
	tactile_paving	Yes	Tactile surfaces (or tactile strips) present on either side of the pedestrian crossing.		**NOEUD_CHEMINEMENT.**bandeEveilVigilance code 02/03/04/05
	tactile_paving	No	Lack of tactile surfaces (tactile bands).		**NOEUD_CHEMINEMENT.**bandeEveilVigilance code 01
	tactile_paving	Incorrect	Tactile surface (tactile strip) incorrectly placed, or unsecured, or missing on one side of the pedestrian crossing.		**NOEUD_CHEMINEMENT.**controleBEV code different from 01

	crossing	Island	Passage with a small central island for pedestrians. An island is a space laid out in the middle of a road, whose role is to physically separate traffic flows.		**TRONCON_CHEMINEMENT.TRAVERSEE.** presenceIlot
 	Lit	Yes	Illuminated pedestrian crossing, useful for everyone, and especially for people with autism.		**TRONCON_CHEMINEMENT.TRAVERSEE.** eclairage
 	crossing	traffic_signals	Presence of a pedestrian crossing where the crossing of pedestrians and cyclists is regulated by traffic lights.		**TRONCON_CHEMINEMENT.TRAVERSEE.** feuLumineux

 					
	traffic_signals:sound	Yes/no	Pedestrian crossing equipped with an audible signal for people with visual impairments.		**TRONCON_CHEMINEMENT.TRAVERSEE.** feuSonore
	crossing	traffic_signals	traffic_signals:floor_vibration to indicate that vibrations, occur when crossing is permitted.		**ABSENT**
	crossing	traffic_signals	traffic_signals:minimap to indicate that there is a small tactile map at the traffic light pole to indicate the crossing layout for blind persons.		**ABSENT**

Appendix 2: Wheelchair accessibility tag → CNIG Objects attributes

wheelchair

					CNIG
	wheelchair	yes	Presence of boats on either side of the pedestrian crossing. (lowered areas to allow wheelchairs to pass)		**NOEUD_CHEMINEMENT.**hauteurRessaut + **NOEUD_CHEMINEMENT**.abaisseTrottoir
	wheelchair	no	Sidewalk not lowered, preventing any crossing of the pedestrian crossing for people moving in wheelchairs.		
	wheelchair	limited	Partially accessible, one of the 2 boats is not accessible or requires the assistance of a person to push the wheelchair on a fairly high step.		
	wheelchair	bad	Boat not sufficiently marked (or with an awkward gutter slope) preventing anyone in a wheelchair from crossing easily.		

***Appendix 3:** CNIG Objects attributes → OSM tags*

<u>CONVERSION QUALITY</u>

1	Existing attribute and easily convertible value
2	Existing attribute and value not easily convertible
3	Attribute not existing in the model

CNIG Objects		**OSM Tags**	**Conversion quality**
Name:	**CHEMINEMENT**	/	3
	Idcheminement	id	
Attributes :	nom	name	

Name: Attributes :	**TRONCON_CHEMINEMENT**	**WAYS**	1
	Idtroncon	id	2
	from	id du premier nœud	2
	to	id du deuxième nœud	2
	distance	/	3
	typeTroncon	highway footway	2
	ascenseur	highway=elevator	1
	tapis roulant	highway=footway + conveying	1
	escaliers	highway=steps	1
	espace ouvert	covered	2
	chemin piéton	highway= footway	1
	escalator	highway=steps + conveying	1
	rampe	ramp	1
	traversée	highway=crossing	1
	trottoir	footway=sidewalk	1

	statutVoie	highway=living_street highway=pedestrian	2
	pente	incline	1
	devers	/	3
	accessibiliteGlobale	/	3

Name: Attributes :	**CIRCULATION**	**WAYS**	2
	idcirculation	id	2
	typesol	surface	1
	largeurPassageUtile	width	1
	etatRevetement	smoothness	2
	eclairage	lit	2
	transition	/	3
	typepassage	footway = sidewalk	2
	repereLineaire	tactile_paving	2
	couvert	/	3

Name:	**TRAVERSEE**	**WAYS** **footway=crossing**	1
Attributes :	idtraversee	id	2
	etatRevetement	smoothness	2
	marquageSol	/	3
	eclairage	lit	2
	feuLumineux	crossing= traffic_signals	1
	feuSonore	traffic_signals : sound	1
	repereLineaire	tactile_paving	2
	presenceIlot	island	3
	chausseeBombee	/	3
	covisibilite	/	3

CNIG Objects		OSM Tags	
Name:	**ESCALIER**	**WAY** **highway=steps**	1
Attributes :	idescalier	id	2
	etatRevetement	smoothness	2
	mainCourante	handrail	1
	dispositifVigilance	tactile_paving	1
	dispositifMarche	/	3
	largeurUtile	width	1
	mainCouranteContinue	/	3
	prolongMainCourante	/	3
	nbMarches	step_count	1
	nbVoleeMarches	/	3
	hauteurMarche	/	3

Name: Attributes :	**ESCALATOR**	highway=steps + conveyor=yes	1
	idescalator	id	2
	sens	Incline = up / down	1
	dispositifVigilance	tactile_paving	2
	largeurUtile	width	1
	detecteur	/	3
	supervision	/	3

Name:	**ASCENCEUR**	highway = elevator	1
Attributes :	idascenseur	id	2
	largeurUtile	width	1
	diamZoneManoeuvre	wheelchair = yes/no	2
	largeurCabine	/	3
	longueurCabine	/	3
	touchesEnRelief	tactil_writting :brail :lg = yes/no	2
	signalSonore	/	3
	equipBim	/	3

	miroir	/	3
	eclairage	/	3
	voyantAlerte	/	3
	annonceEtage	/	3
	typeOuverture	/	3
	barreAppui	/	3
	hauteurBarreAppui	/	3
	etatRevetement	/	3
	supervision	/	3
	autrePorteSortie	/	3

Name:	**NOEUD**	**NODES**	**1**
Attributes :	idnoeud	id	2
	altitude	ele=* – Used to store the elevation relative to a EGM96 geoid	2
	bandeEveilVigilance	kerb:tactile_paving	2
	hauteurRessaut	kerb:height	1
	abaissePente	/	3
	abaisseTrottoir	/	3
	controleBEV	/	3
	bandeInterception	/	3

Name:	**OBSTACLE**	**NODES**	2
Attributes :	idobstacle	Id	2
	typeObstacle	amenity / barrier	2
	largeurPassageUtile	/	3
	positionObstacle	/	3
	longueurObstacle	/	3
	rappelObstacle	/	3
	reperabiliteVisuelle	covered	3
	largeurObstacle	/	3
	hauteurObsPoseSol	/	3

	hauteurSousObs	/	3

Name: Attributes :	**STATIONNEMENT PMR**	amenity=parking_space parking_space=disabled	1
	idstationnement	ref	2
	typeStationnement	/	3
	etatRevetement	smoothness	1
	largeurStat	/	3
	longueurStat	/	3
	bandLatSecurite	/	3
	surLongueur	/	3
	signalPMR	/	3
	marquageSol	/	3
	pente	incline	2
	devers	/	3
	typesol	surface	1