EUROGEOGRAPHICS – QKEN WEBINAR, APRIL 20th 2023

# Analysing addresses: comparison of addresses from SGDC and CDAU (IECA)

Javier Luque – Spanish General Directorate of Cadastre

# Introduction

Postal addresses can be considered one of the most important geographic objects in a territory's spatial data infrastructure.

Allow:

- the management of public and private services, and also…
- geocode other data.

Generated with Dall.E

# Introduction

All national / regional / local agencies need reliable and complete address data to carry out their functions.

During 2022 → pilot project between IECA and DGC to analyze addresses available in both organizations and detect ways of improvement.

**Cooperation and collaboration.**

Generated with Dall.E

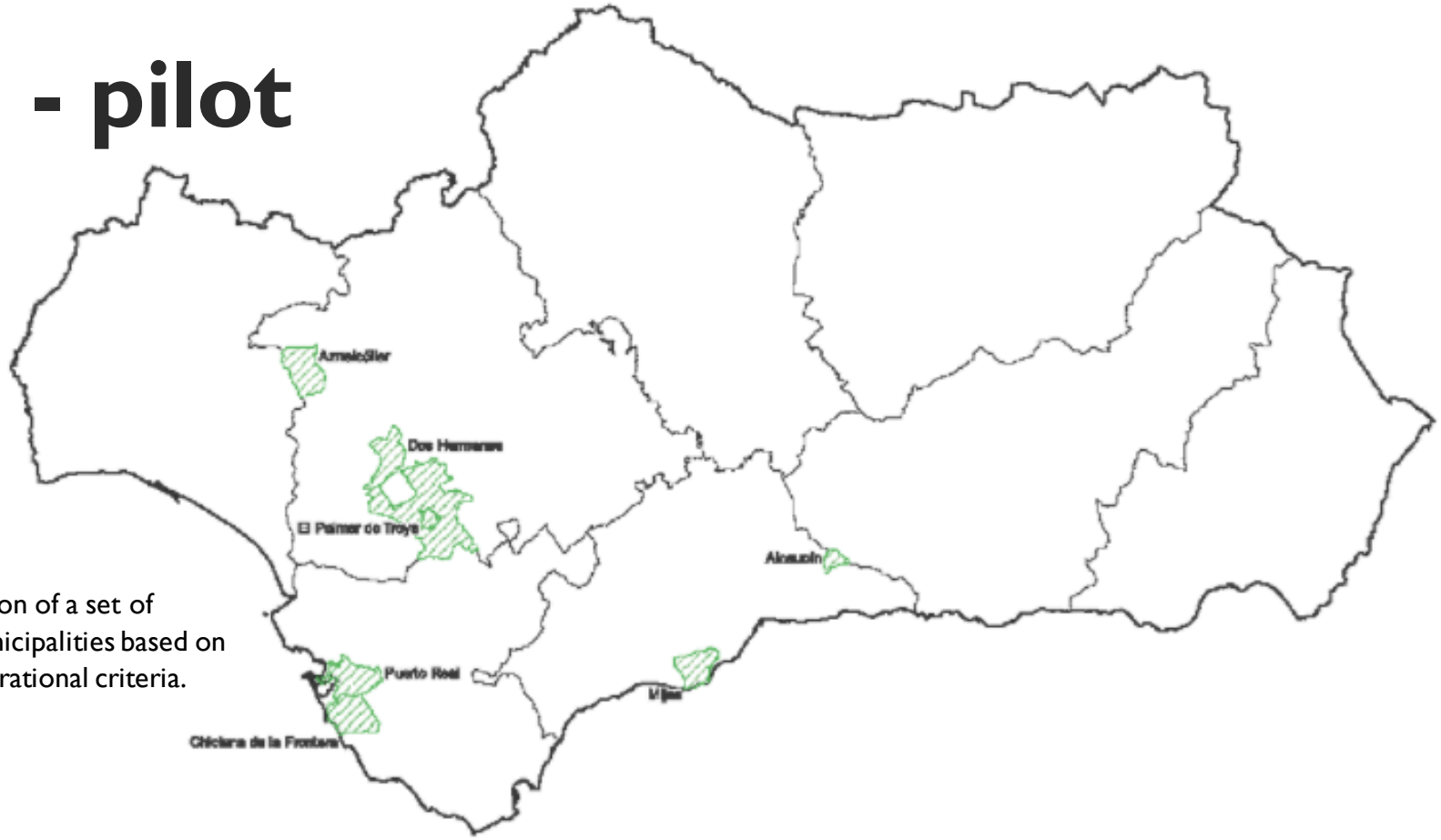
# Introduction - pilot

Municipalities

**Province of Cadiz**:
Chiclana de la Frontera
Puerto Real

**Provincia of Malaga**:
Alcaucín
Mijas

**Province of Sevilla**:
Aznalcóllar
Dos Hermanas
El Palmar de Troya
Utrera




Selection of a set of municipalities based on operational criteria.

**After pilot project, analysis has been extended to whole Andalusia**

# Background

Complex panorama around addresses in Spain

- ✓ **INE (National Statistics Institute)**
  - → register of addresses not georreferenced
  - → register of street names
  - → register of municipalities and settlements names

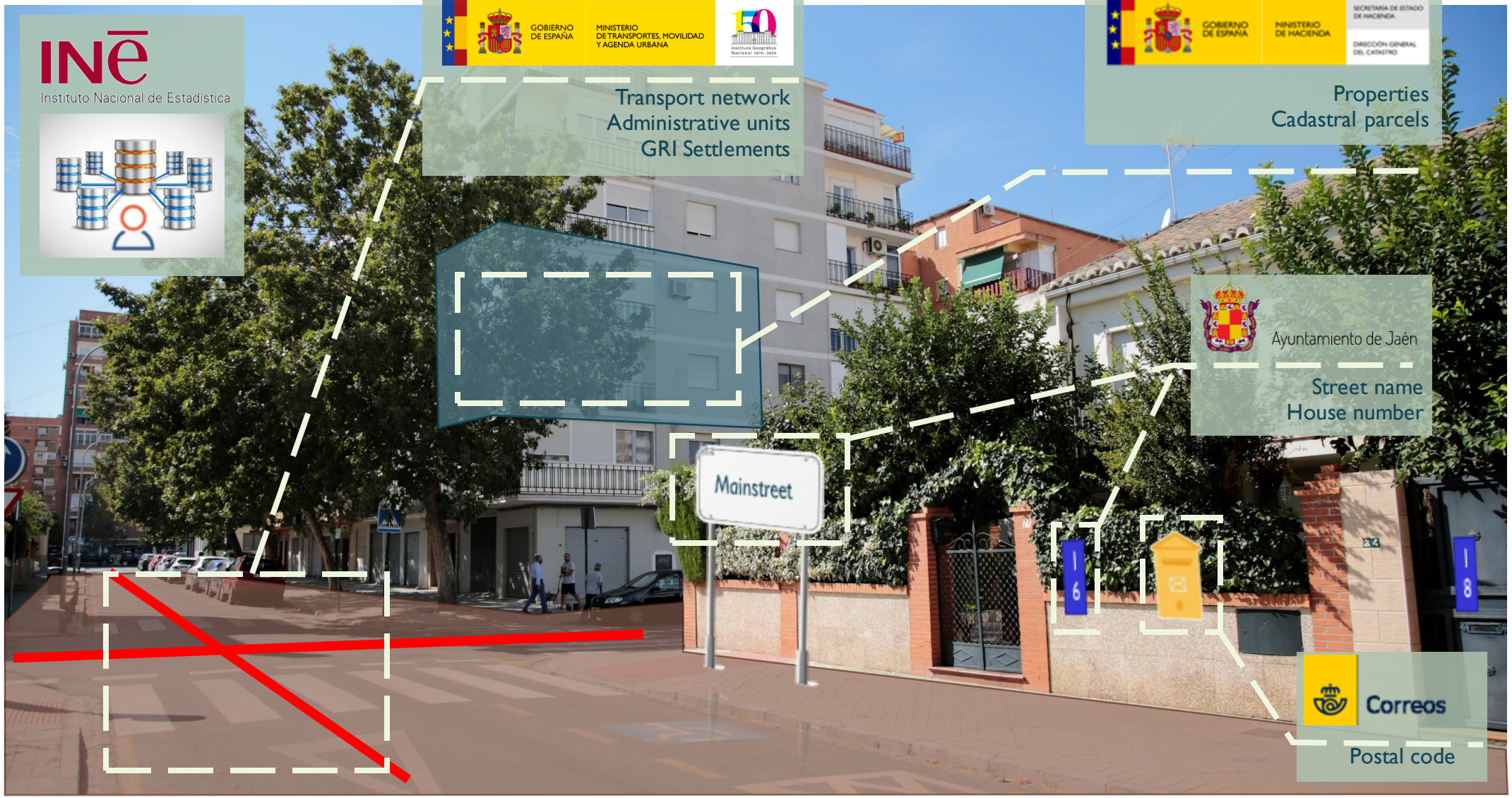
- ✓ **MUNICIPALITIES**
  - → set names to streets
  - → set house numbers

- ✓ **CORREOS (Postal Service)**
  - → maintain postal codes

- ✓ **GENERAL DIRECTORATE OF CADASTRE**
  - → **Georreferenced addresses from real state data (for its territorial scope of competence)**

- ✓ **IGN (National Geographic Institute)**
  - → Transport network
  - → Administrative units
  - → GRI Settlements

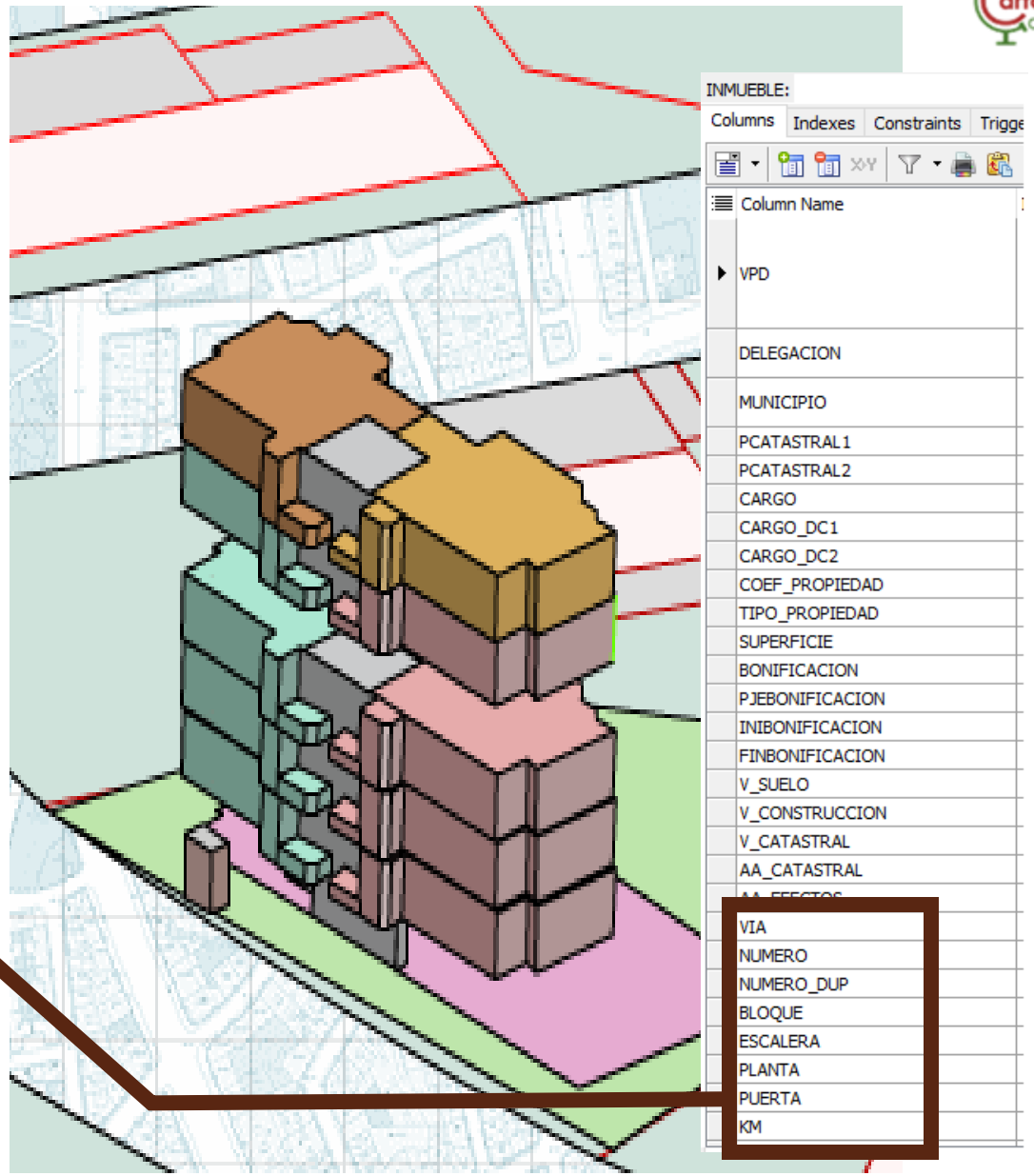
- ✓ **… and OTHERS** like regional statistics and/or mapping organizations


INE
Instituto Nacional de Estadística
GOBIERNO DE ESPAÑA
MINISTERIO DE TRANSPORTES, MOVILIDAD Y AGENDA URBANA
Transport network
Administrative units
GRI Settlements
GOBIERNO DE ESPAÑA
MINISTERIO DE HACIENDA
SECRETARÍA DE ESTADO DE HACIENDA
DIRECCIÓN GENERAL DEL CATASTRO
Properties
Cadastral parcels
Ayuntamiento de Jaén
Street name
House number
Mainstreet
Correos
Postal code

# Address in cadastral database

Address information is part of the cadastral objects: cadastral parcels, apartments, etc.

Address stored as attributes in tables that represent cadastral elements.



# Address in cadastral database

Addresses geo-referenced:

- to entraces → usually as point located in street, out of cadastral parcel
- to features → centroid of cadastral parcels

Level of detail

- SGDC produce INSPIRE dataset with information up to house number + extension (if necessary) in most cases.
- Secondary accesses are not georeferenced → only main address of properties

Special cases:

- some buildings / locations have 2 house numbers
- some addresses improved up to block level (one step further)

# Callejero Digital de Andalucía Unificado (CDAU)

Andalucía se mueve con Europa

**C**allejero: Streetmap. Map representing streets and other relevant elements of the territory to make easier the interpretation

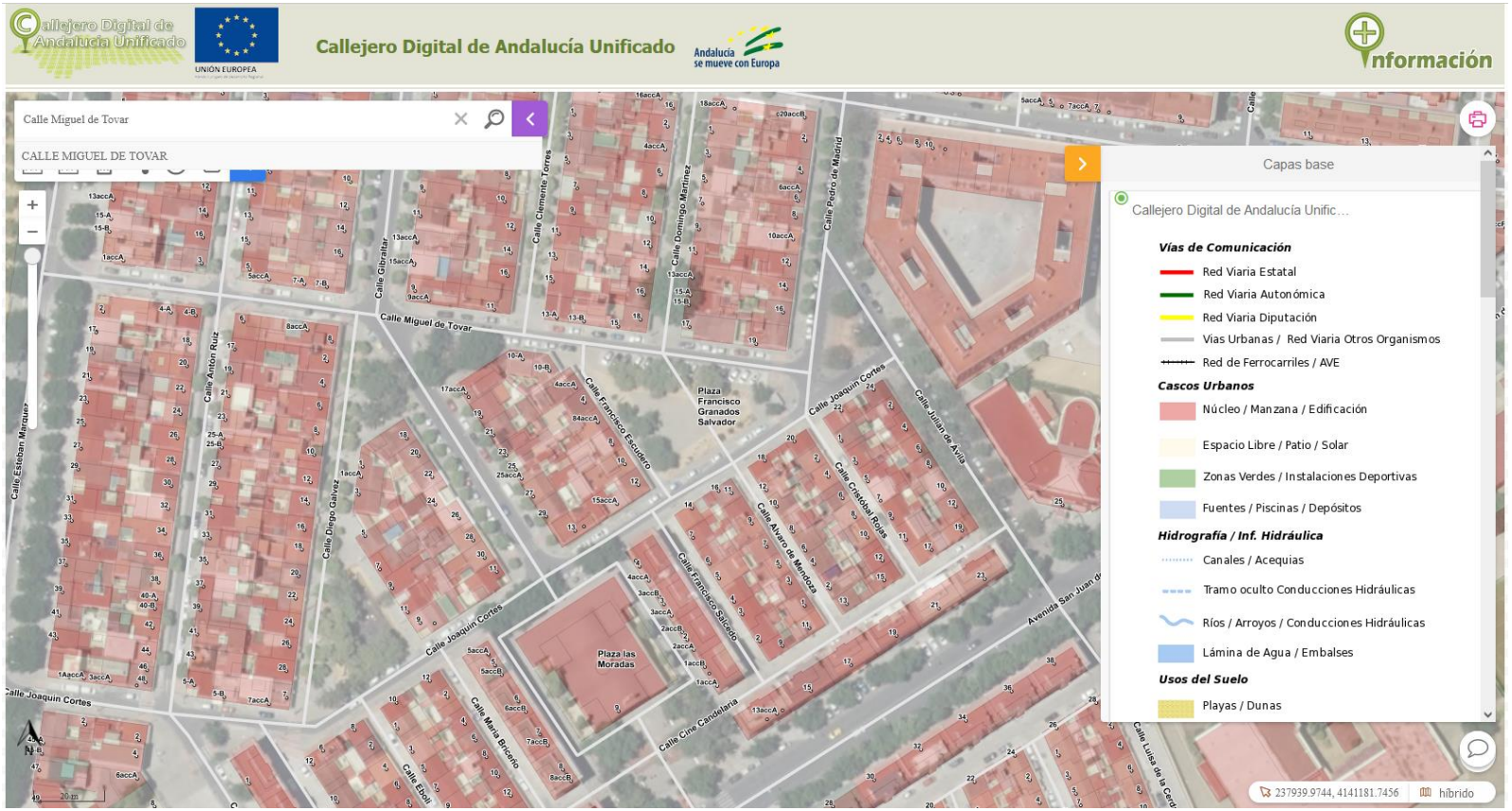
**D**igital: Digital. Accessible via internet.

De **A**ndalucía: Covers all municipalities of Andalusia

**U**nificado: unify data from different sources about streets and entrances in Andalusia and allow to create a unique institutional data with de-centralized maintenance.

## Unified Digital Street Map of Andalusia

# Callejero Digital de Andalucía Unificado (CDAU)



# Callejero Digital de Andalucía Unificado (CDAU)

CDAU is a living and collaborative project

| | ¿How to participate? | |
|---|---|---|
| | **Detect and communicate changes in the street map** | **Edit and correct data** |
| **Instituto de Estadística y Cartografía de Andalucía** | ✓ | ✓ |
| **Local entities** (Councils y Provincial Councils) | ✓ | ✓ |
| **Advanced users from public organizations** (112 Emergencies - Andalusia) | ✓ | ✗ |
| **Citizens** | ✓ | ✗ |

# Callejero Digital de Andalucía Unificado (CDAU)

Andalucía se mueve con Europa

Addresses geo-referenced:

- to entraces → point located inside cadastral parcel, in its façade

Level of detail

- dataset with information up to house number + extension (if necessary) in most cases.
- Secondary accesses are georeferenced
- Include kilometric points in roads

Special cases:

- some buildings / locations have 2 house numbers

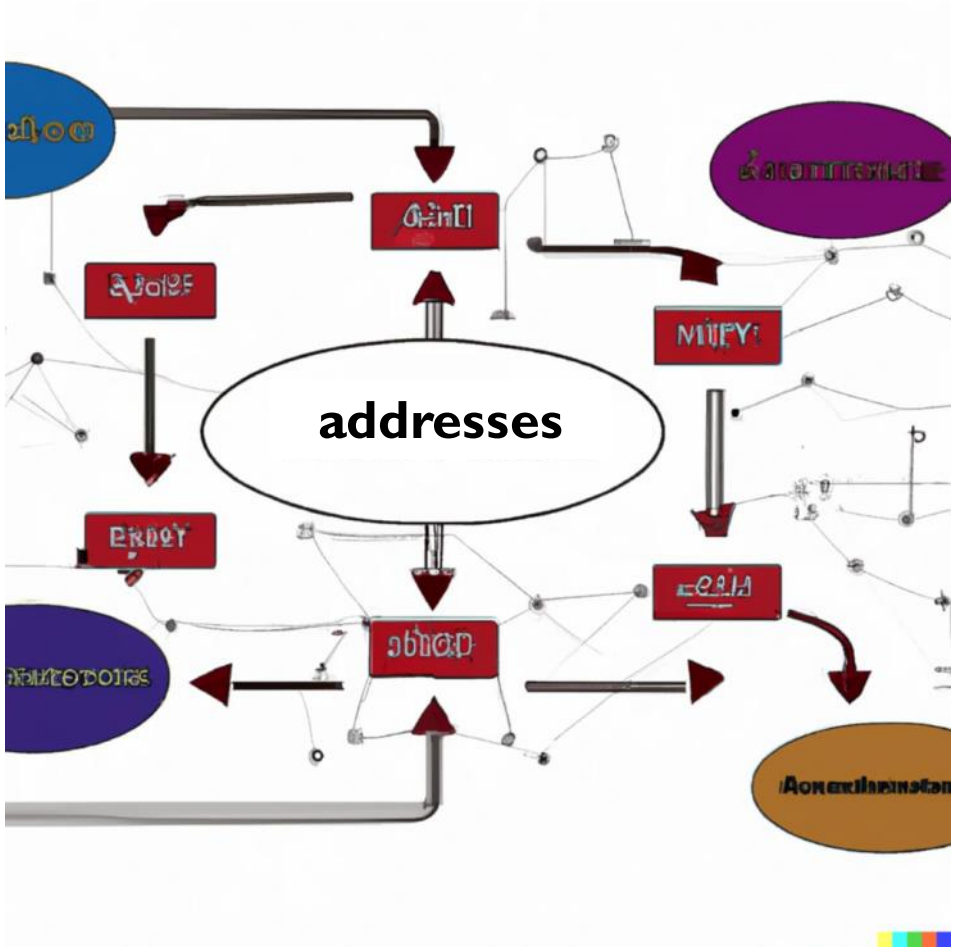
# Analysis process

Inputs:

- Addresses from CDAU (points) and roads / streets (lines).
- Adresses from SGDC (points) and roads / streets as table (no geometry).

Comparative analysis quite complex due to some reasons…

… let's have a look over the data




Generated with Dall.E

# Issues to consider in data analysis

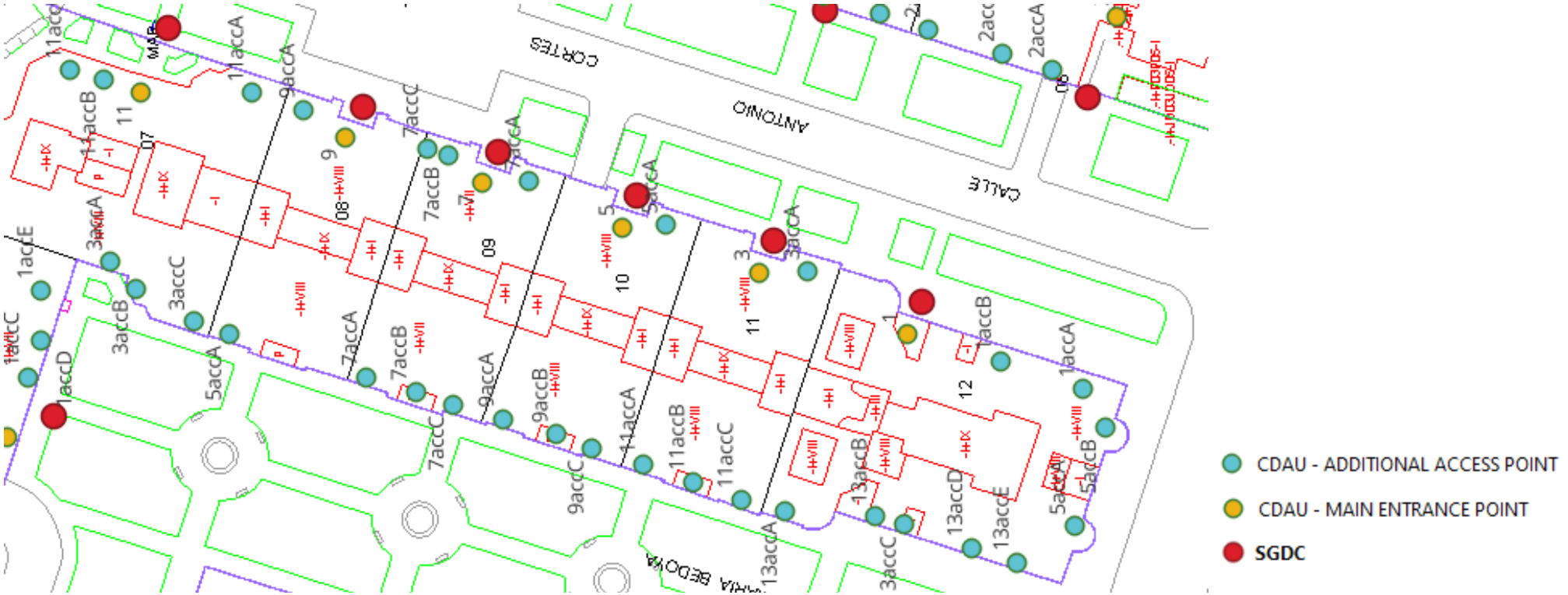
Different locations

# Issues to consider in data analysis

Different expression at number extension / block identification

# Issues to consider in data analysis

Different level of granularity



# Analysis process

Process improved in lasts weeks → previous process used a stepwise join approach using fewer key attributes at each stage

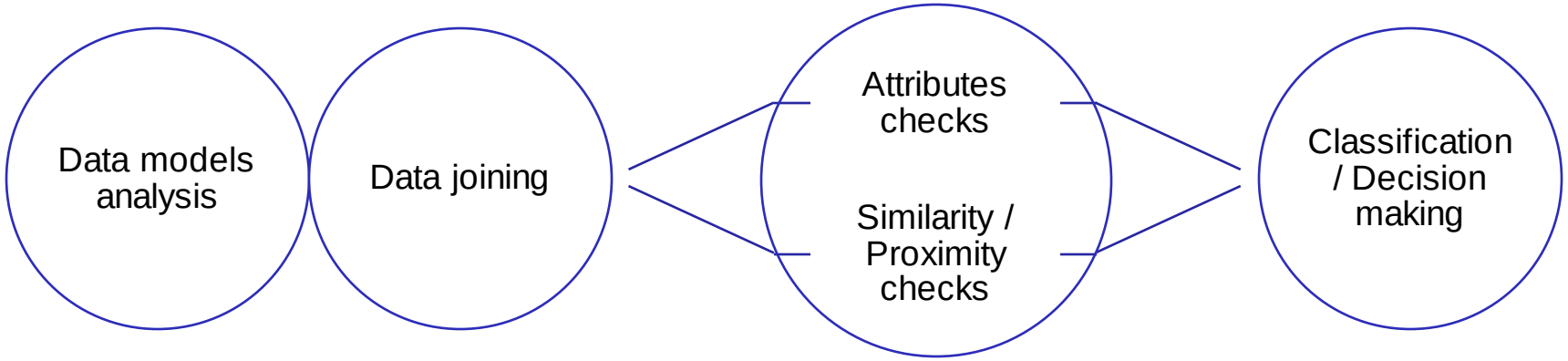
Stages:

1) Analysis of data models, establishment of equivalence of attributes between data sets.

2) Join of data sets through common attributes (municipality + cadastral reference when possible). Check codes of streets, house numbers and house number extension.

3) Spatial join by proximity for un-joined data

Analysis of proximity and similarity between joined data and decision-making in data pairings transversally to the process.



# Resuls

| | FULL JOIN | UP TO NUMBER EXT. | UP TO HOUSE NUMBER | SPATIAL JOIN - REQUIRES ATTRIBUTES CHECKS | TOO FAR - FULL JOIN | TOO FAR - UP TO NUMBER EXT. | TOO FAR - UP TO HOUSE NUMBER | TOO FAR - SPATIAL JOIN | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| CADIZ | 52,5% | 0,8% | 3,8% | 29,4% | 0,6% | 0,6% | 0,0% | 12,2% | 100,0% |
| CORDOBA | 67,9% | 0,3% | 2,4% | 28,6% | 0,1% | 0,0% | 0,0% | 0,7% | 100,0% |
| GRANADA | 64,5% | 0,1% | 2,3% | 32,5% | 0,1% | 0,0% | 0,0% | 0,4% | 100,0% |
| HUELVA | 57,1% | 0,4% | 3,7% | 35,8% | 0,3% | 0,0% | 0,0% | 2,8% | 100,0% |
| JAEN | 75,5% | 0,3% | 1,4% | 22,4% | 0,0% | 0,0% | 0,0% | 0,4% | 100,0% |
| MALAGA | 60,7% | 0,5% | 2,4% | 34,2% | 0,3% | 0,0% | 0,0% | 2,0% | 100,0% |
| ALMERIA | 60,1% | 0,2% | 3,1% | 35,3% | 0,3% | 0,0% | 0,5% | 0,5% | 100,0% |
| SEVILLA | 70,8% | 0,4% | 4,0% | 22,3% | 0,3% | 0,0% | 0,1% | 2,0% | 100,0% |
| TOTAL | 64,3% | 0,4% | 3,1% | 28,9% | 0,3% | 0,1% | 0,1% | 2,9% | 100,0% |

# Results - summary

~68% postal addresses established direct correspondence.

+ 4% additional joins due to spatial proximity with good matching.

Initial estimated percentage of agreement ~72%

But…
~3,5% of joins addresses considered incorrect…
… and ~24% of joins cannot be automatically confirmed, need extra manual check


Not joined:

- 16% DGC addresses
- 11% addresses CDAU-IECA.

Reasons:

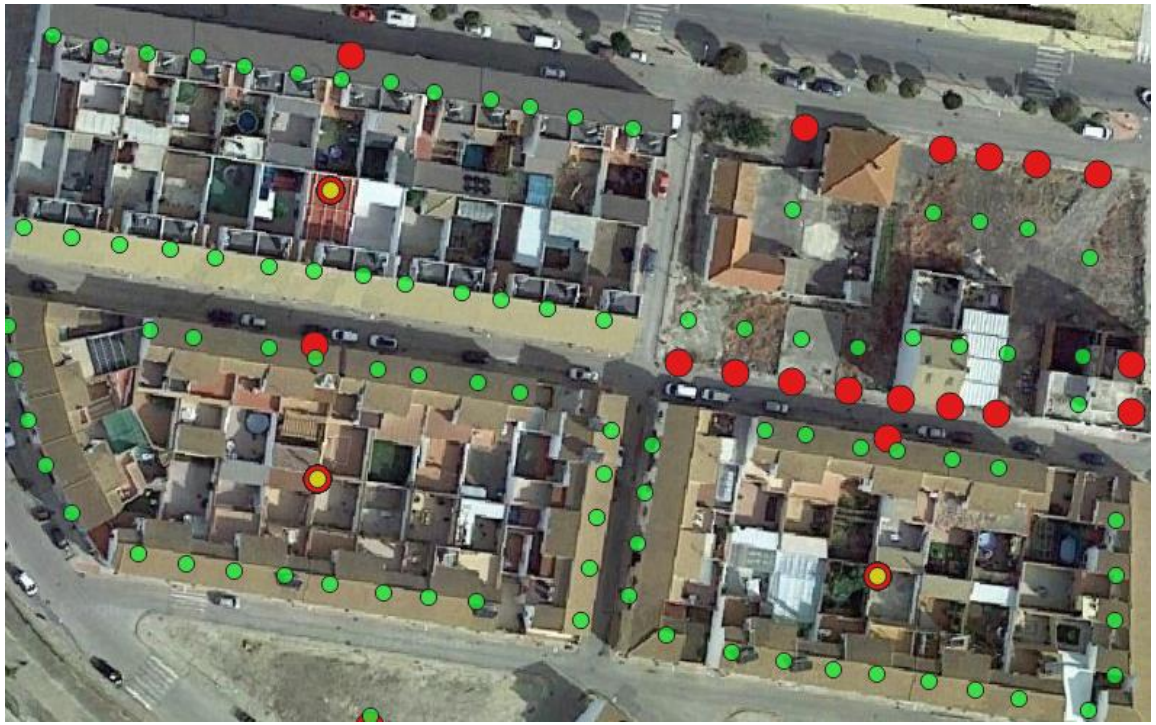
- Omissions of addresses in datasets
- Omission of attribute or incorrect values that prevent join.

# Results - examples

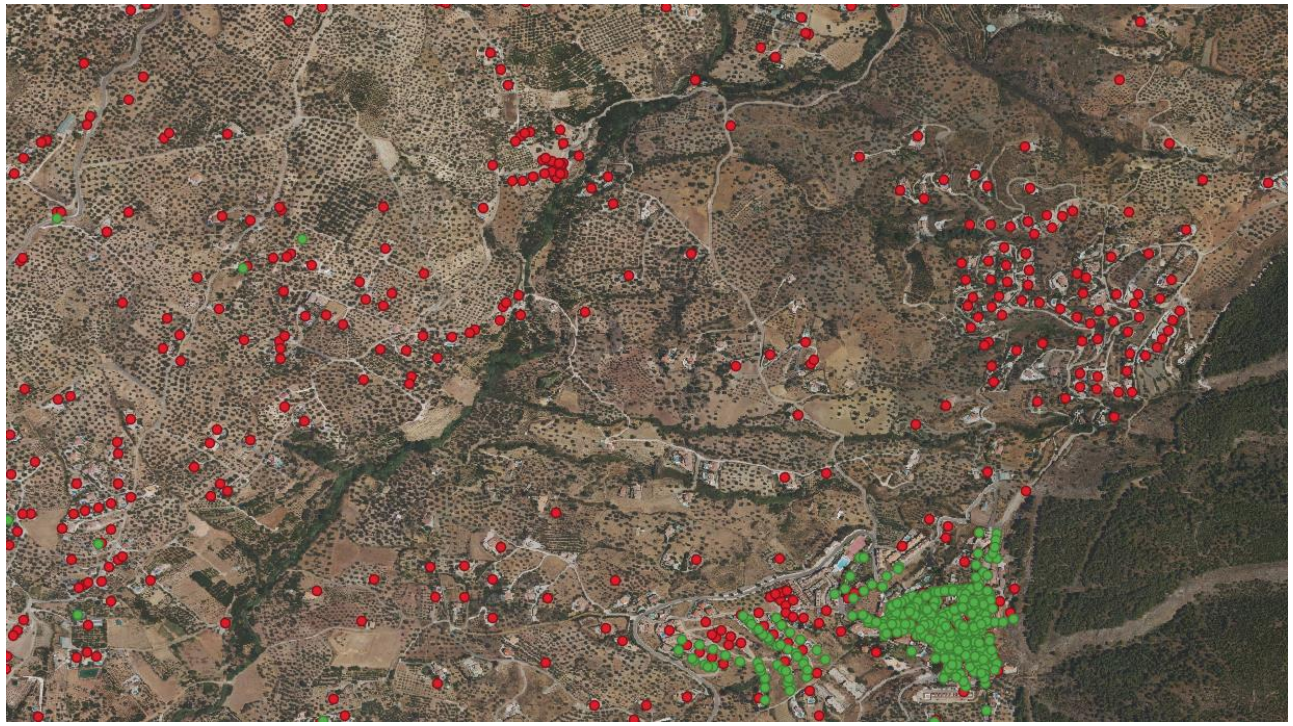

← **Mijas.** High relation between both dastasets.


**El Palmar de Troya.** In green addresses connected and verified correct by street name. In pink, addresses for which complete verification has not been possible, but there is a high probability of being a correct match.

# Results - examples




**El Palmar de Troya.** Area with correspondence between addresses in both sources, with georeferencing to centroid in Cadastre (red color) and georeferencing to building entrance in CDAU (green color). Allows improvement of georeferencing of this data set.




**Alcaucín.** Area with Cadastre addresses (red color) in "disseminated" area not collected in CDAU (green color).

# Conclusions

- The differences between data models are appreciable but allow analysis and combination of information.
- The absence of certain attributes (INE road code, cadastral reference,...) in some objects make it difficult to join the information.
- Need to advance in the completion of said attributes.
- The systematic exchange of information is feasible and will allow the evolution and improvement of both sets of data.
- It lays the groundwork for future collaboration. It shows the potential of inter-administrative collaboration.

# Acknowledgments


**Instituto de Estadística y Cartografía de Andalucía**

Luis Copano
José Antonio Moreno
Paloma López
Daniel José Montero

**Dirección General del Catastro**

José Miguel Olivares

# Thanks!

Javier Luque

javier.luque@catastro.hacienda.gob.es