



Projecte NEO4J

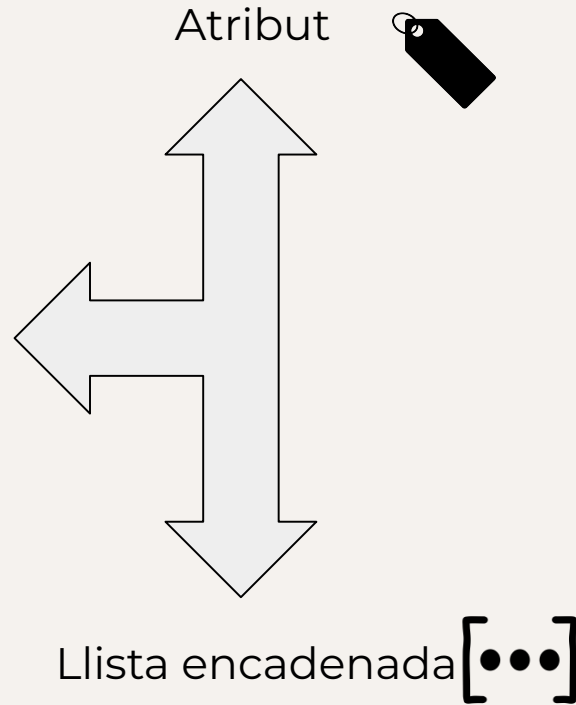
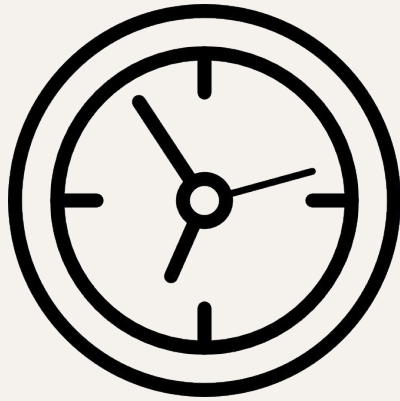
Iker Bolancé, Alejandro Rodríguez, Mariona Guisado & Marcel Coll



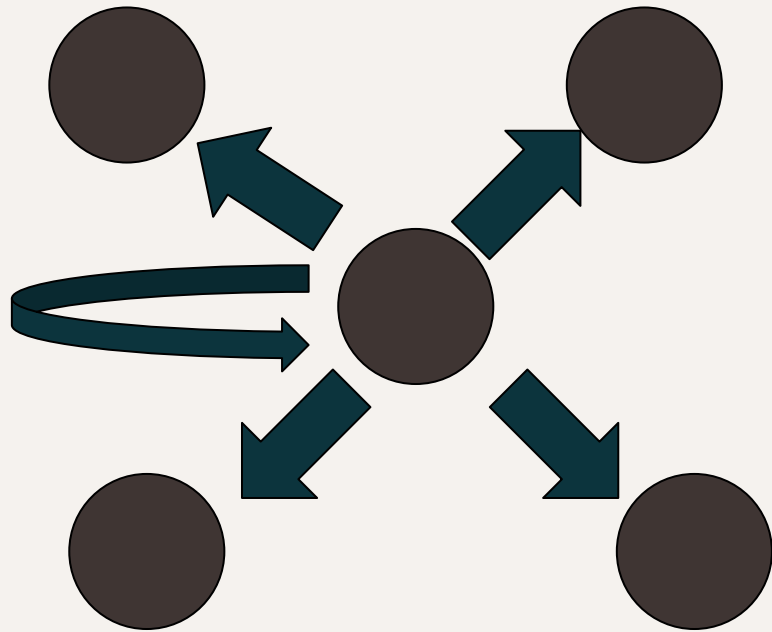
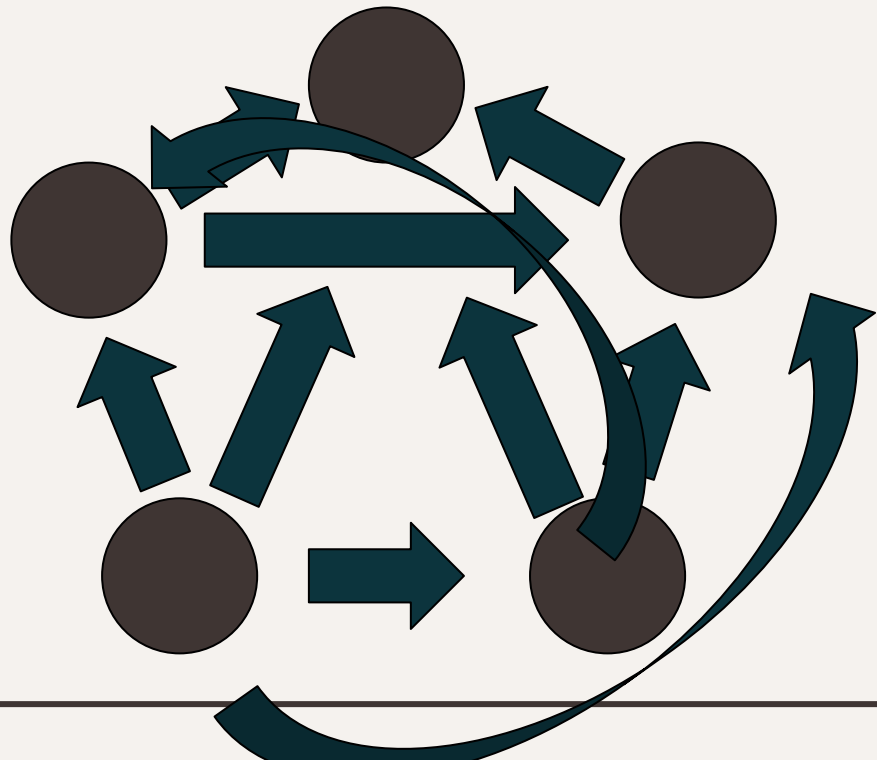


Problema

01 Canvis realitzats

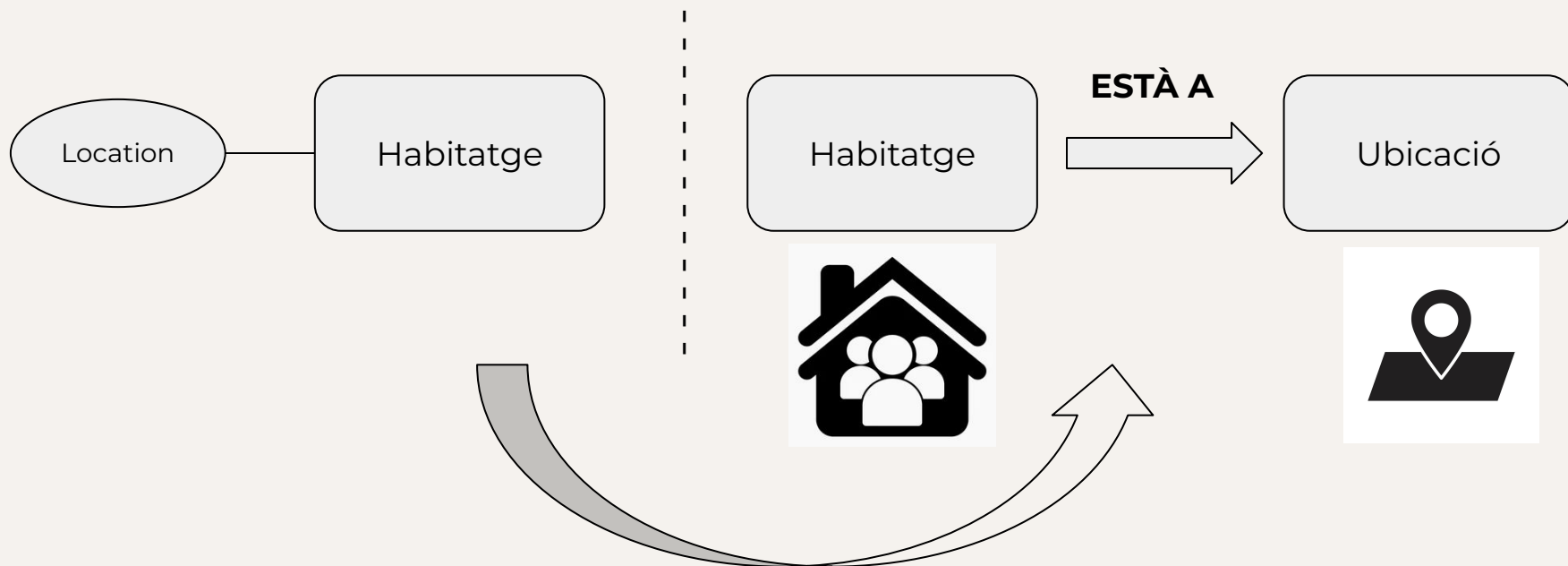


02 Canvis realitzats



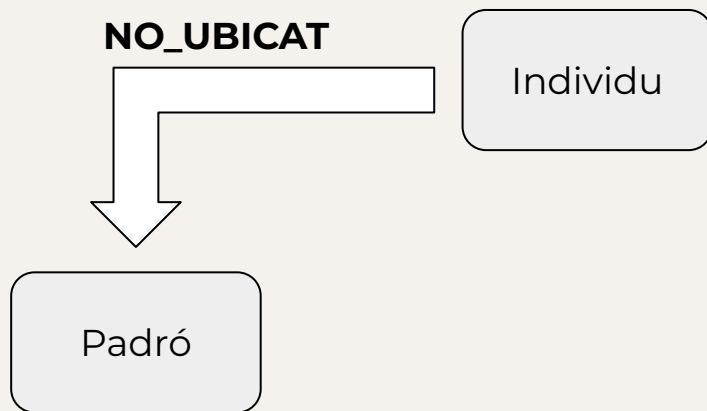
03 Canvis realitzats

Separació d'habitatge i ubicació



04 Canvis realitzats

Gestió de persones sense domicili



Avantatges

Temps com a llista encadenada



Reducció de redundància i millor normalització temporal

De cliques a graf estel·lar



Menys arestes, menys memòria i eliminació de redundàncies familiars

Separació d'habitatge i ubicació

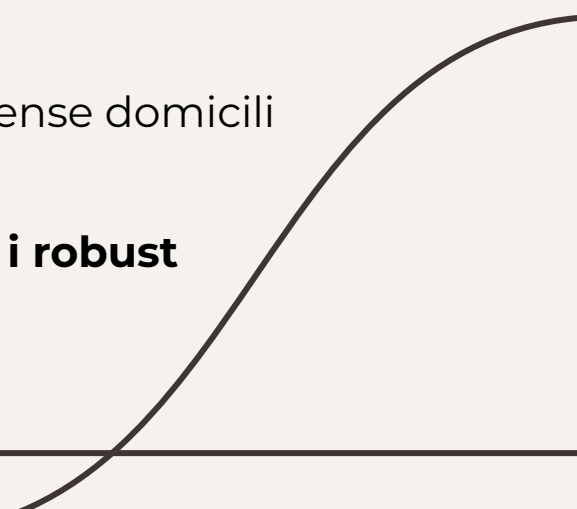


Millor organització i menys supernodes

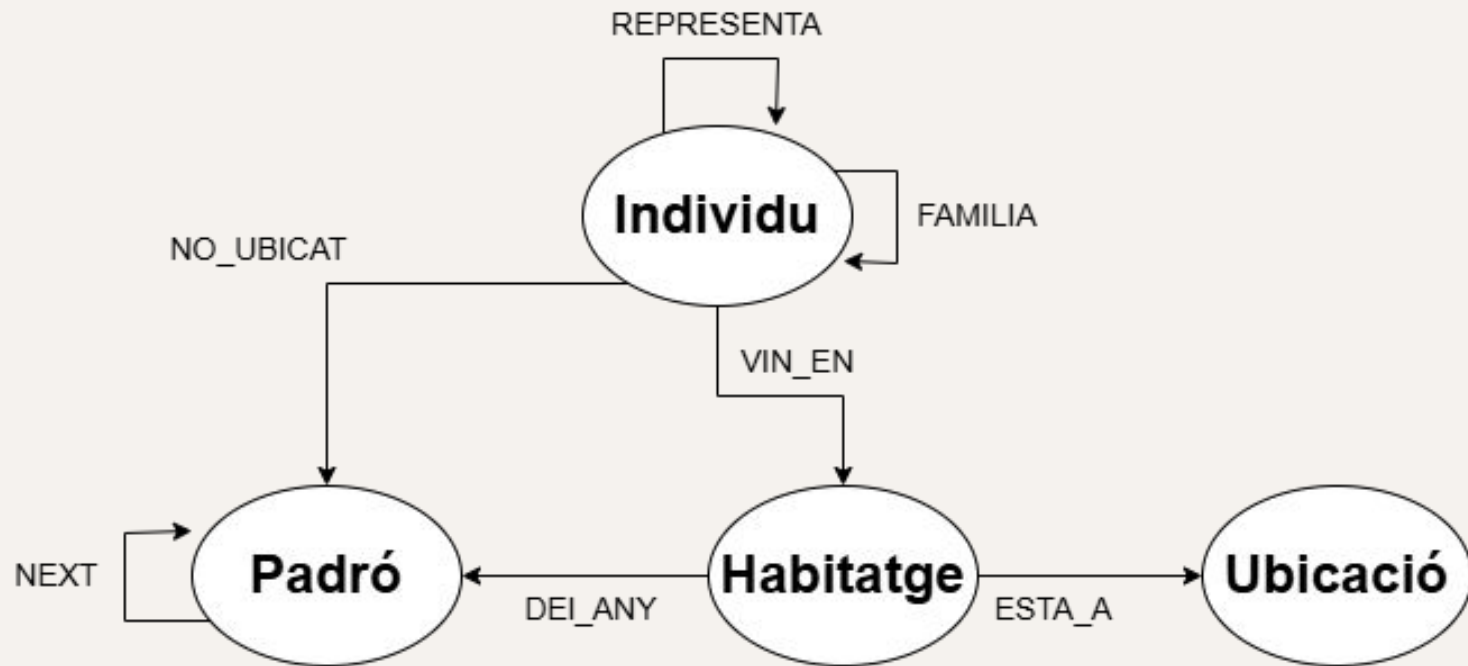
Gestió de persones sense domicili



Model més complet i robust



Diagrama



Consultes: Nou model

Persones que sempre han viscut al mateix lloc:

```
MATCH (rep:Representant)-[:VIU_EN]->(:Habitatge)-[:ESTA_A]->(u:Ubicacio)
WITH rep, size(collect(u)) AS n_ubicacions, collect(DISTINCT u) AS ubicacions
WHERE size(ubicacions) = 1 AND n_ubicacions > 1
RETURN size(collect(rep)) AS SEDENTARIS
```

SEDENTARIS	
1	268

Consultes: model original

Persones que sempre han viscut al mateix lloc:

```
MATCH (p:Person)
OPTIONAL MATCH (p)-[:SAME_AS]-(q:Person)
WITH p, collect(q) + p AS group_nodes
WITH [x IN group_nodes WHERE x IS NOT NULL | x.id] AS ids, group_nodes
WITH apoc.coll.sort(ids)[0] AS group_id, collect(DISTINCT p) AS persons
UNWIND persons AS person
MATCH (person)-[:VIU]->(h:Habitatge)
WITH group_id, collect(DISTINCT h.municipi + '|' + h.carrer + '|' + h.numero) AS
adreces
WHERE size(adreces) = 1
RETURN count(*) AS SEDENTARIS
```

	SEDENTARIS
1	6541

Consultes: Nou model

Persones que apareixen a dos padrons consecutius:

```
MATCH (p1:Padro {year: $any_padro})-[:NEXT]->(p2:Padro)
MATCH (rep:Representant)-[:VIU_EN]->(:Habitatge)-[:DEL_ANY]->(p1)
WITH p1, p2, collect(DISTINCT rep) AS reps1
MATCH (rep:Representant)-[:VIU_EN]->(:Habitatge)-[:DEL_ANY]->(p2)
WITH reps1, collect(DISTINCT rep) AS reps2
WITH reps1, reps2,
    [r IN reps1 WHERE r IN reps2] AS enAmbdos
RETURN
    size(enAmbdos) AS persistents,
    size(reps1) AS primer_any,
    toFloat(size(enAmbdos)) / size(reps1) * 100 AS percentatge
```

Consultes: Resultat Nou Model:

Persones que apareixen a dos padrons consecutius:

persistents	primer_any	percentatge
1973	2744	71.90233236151603

Consultes: Nou model

Persones que apareixen a dos padrons consecutius:

```
MATCH (p1:Padro {year: $any_padro})-[:NEXT]->(p2:Padro)
MATCH (rep:Representant)-[:VIU_EN]->(:Habitatge)-[:DEL_ANY]->(p1)
WITH p1, p2, collect(DISTINCT rep) AS reps1
MATCH (rep:Representant)-[:VIU_EN]->(:Habitatge)-[:DEL_ANY]->(p2)
WITH reps1, collect(DISTINCT rep) AS reps2
WITH reps1, reps2,
    [r IN reps1 WHERE r IN reps2] AS enAmbdos
RETURN
    size(enAmbdos) AS persistents,
    size(reps1) AS primer_any,
    toFloat(size(enAmbdos)) / size(reps1) * 100 AS percentatge
```

```
MATCH (:Habitatge)
WHERE any_padro > year_actual
WITH year_actual, min(any_padro) AS year_seguent
MATCH (p:Person)-[:VIU]->(h1:Habitatge)
WHERE h1.any_padro = year_actual
WITH DISTINCT p, year_actual, year_seguent
OPTIONAL MATCH (p)-[:SAME_AS]-(q:Person)
WITH p, year_actual, year_seguent, collect(q.id) + p.id AS ids
WITH reduce(min_id = p.id, x IN ids | CASE WHEN x < min_id THEN x ELSE min_id END) AS group_id, ids, year_actual, year_seguent
WITH group_id, ids, year_actual, year_seguent
WITH collect(DISTINCT group_id) AS grups_originals, year_actual, year_seguent
UNWIND grups_originals AS gid
MATCH (p:Person)-[:SAME_AS*0..]-(p_equiv:Person)
WHERE p.id = gid
WITH DISTINCT gid, collect(DISTINCT p_equiv.id) AS all_ids
UNWIND all_ids AS pid
MATCH (p2:Person {id: pid})-[:VIU]->(h2:Habitatge)
WHERE h2.any_padro = year_seguent
WITH DISTINCT gid AS grups_en_ambdós, count(DISTINCT gid) AS n_ambdós
WITH n_ambdós, count(*) + 0.0 AS total_grups
RETURN
  total_grups AS total_cens_inicial, n_ambdós AS persones_amb_cens_repetit,
  round(n_ambdós * 100.0 / total_grups, 2) AS percentatge;
```

Consultes: model original

Persones que apareixen a dos padrons consecutius:

Consultes: Resultat Model Original:

total_cens_inicial	persones_amb_cens_repetit	percentatge
2744	1973	71.90233236151603

Importar:

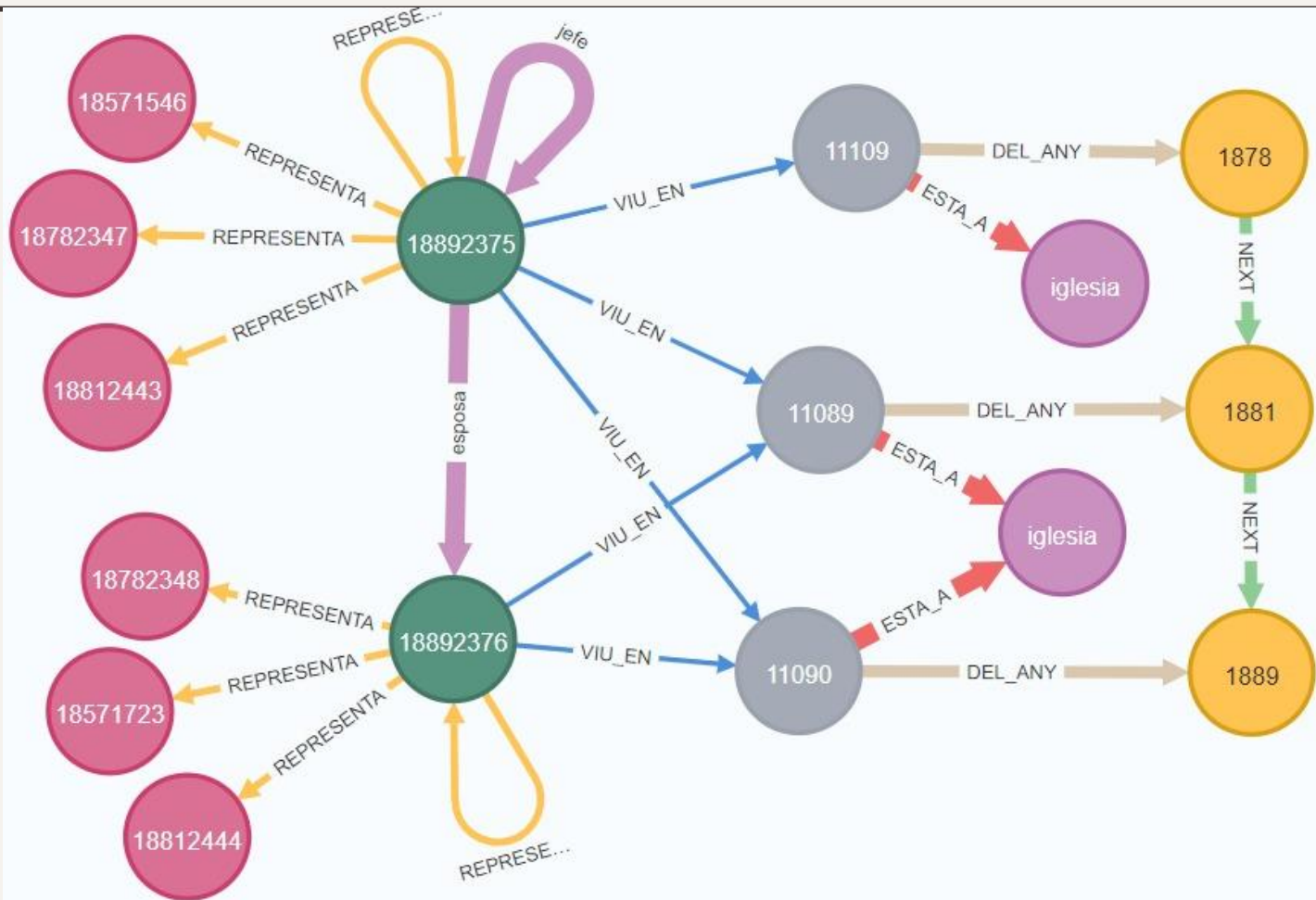
Padró llista encadenada:

```
MATCH (p:Padro)
WITH p ORDER BY p.year
WITH collect(p) AS years
UNWIND range(0, size(years)-2) AS i
WITH years[i] AS from, years[i+1] AS to
MERGE (from)-[:NEXT]->(to);
```

Importar:

Trobar cliques:

```
MATCH (n:Individu)
WHERE n.componentId IS NULL
CALL apoc.path.subgraphNodes(n, {relationshipFilter: 'SAME_AS'}) YIELD
node
WITH n, collect(node) AS component
WITH component, id(n) AS cid
UNWIND component AS x
SET x.componentId = cid
```



Gràcias

