



Brisbane, Australia 6-10 April

Presented at the FIG Working Week 2025, 6-10 April 2025 in Brisbane, Australia

Collaboration, Innovation and Resilience: Championing a Digital Generation

# Artificial Intelligence for Querying Land and Property Data from Cadastral Plans



**Hamid Hosseini**

PhD Candidate, Centre for Spatial Data Infrastructures and Land Administration, The University of Melbourne



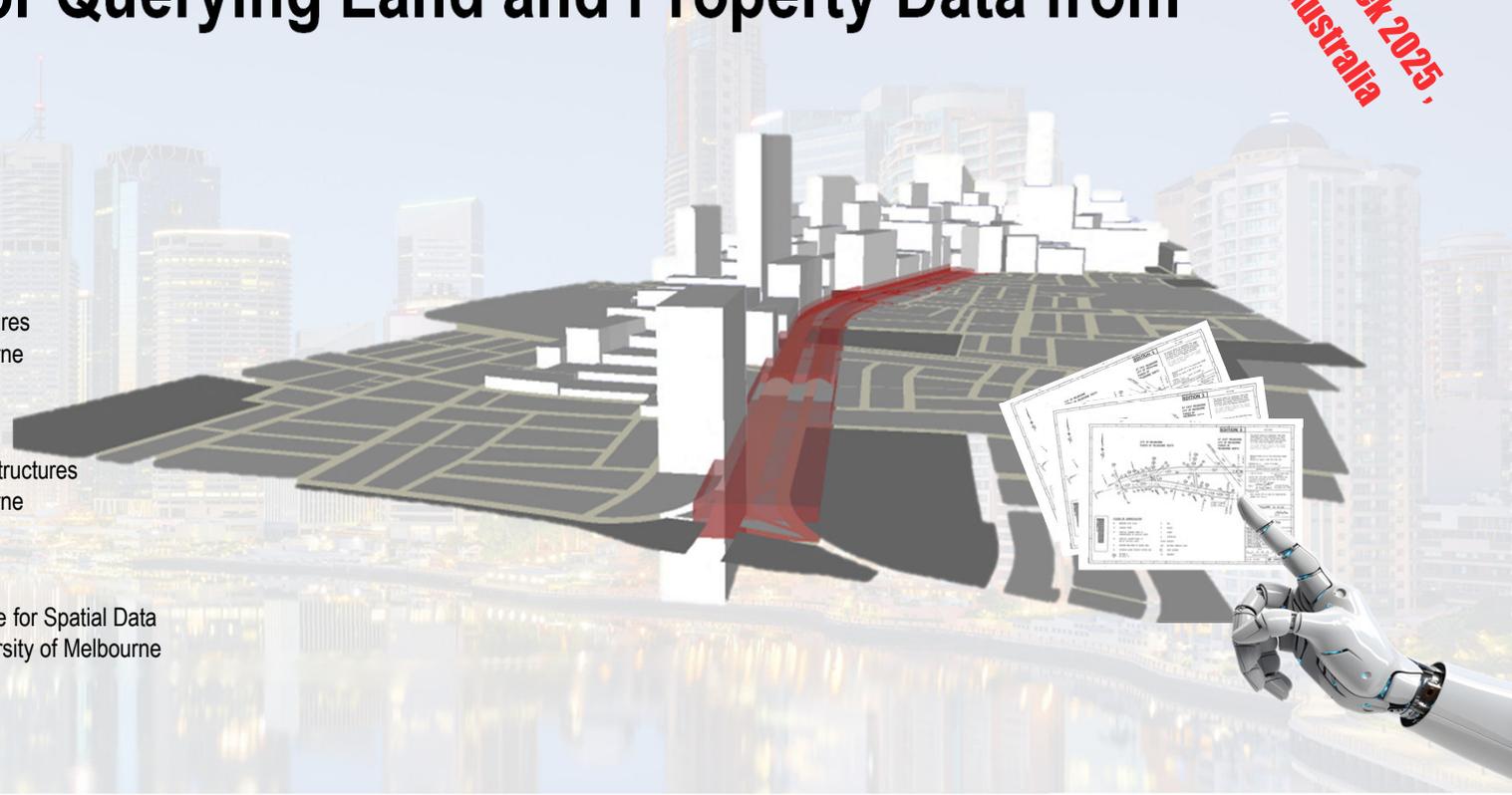
**Behnam Atazadeh**

ARC DECRA Fellow, Centre for Spatial Data Infrastructures and Land Administration, The University of Melbourne



**Abbas Rajabifard**

Professor and Discipline Leader Geomatics, Centre for Spatial Data Infrastructures and Land Administration, The University of Melbourne



PLATINUM SPONSORS

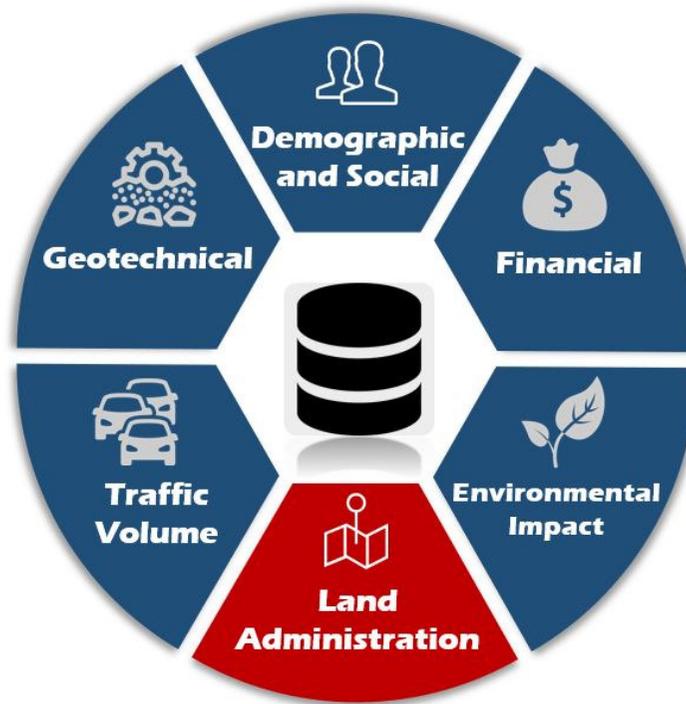


# Land and Property Data

## Semantic and Spatial Information about

- Physical reality of land parcels and properties (i.e., Physical data)
- Legal boundaries and associated rights, restrictions, and responsibilities (RRRs) (i.e., Legal data)

**Critically important** across the entire project lifecycle including design, planning, construction, and maintenance.



# Land and Property Data

**Effective management and communication** of land and property data in **Land Administration Systems (LASs)** is essential!

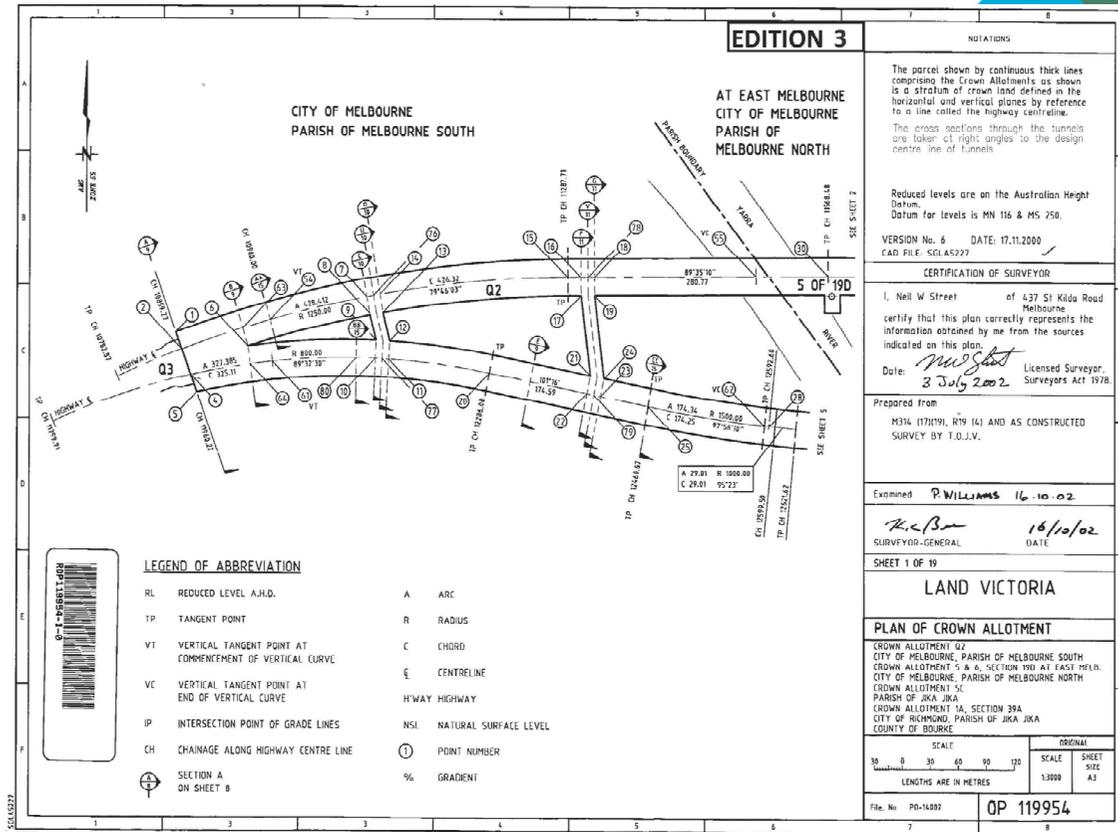
- Significant amount of **Land Surveying and Land Registry**
- Reliable and complete representation of legal spaces can **prevent unnecessary costs, delays and disruptions** in many land development projects
- **Unclear definition of legal spaces** can raise land and property disputes.



## ➤ Cadastral Plans

## ➤ Abstract of Field Records

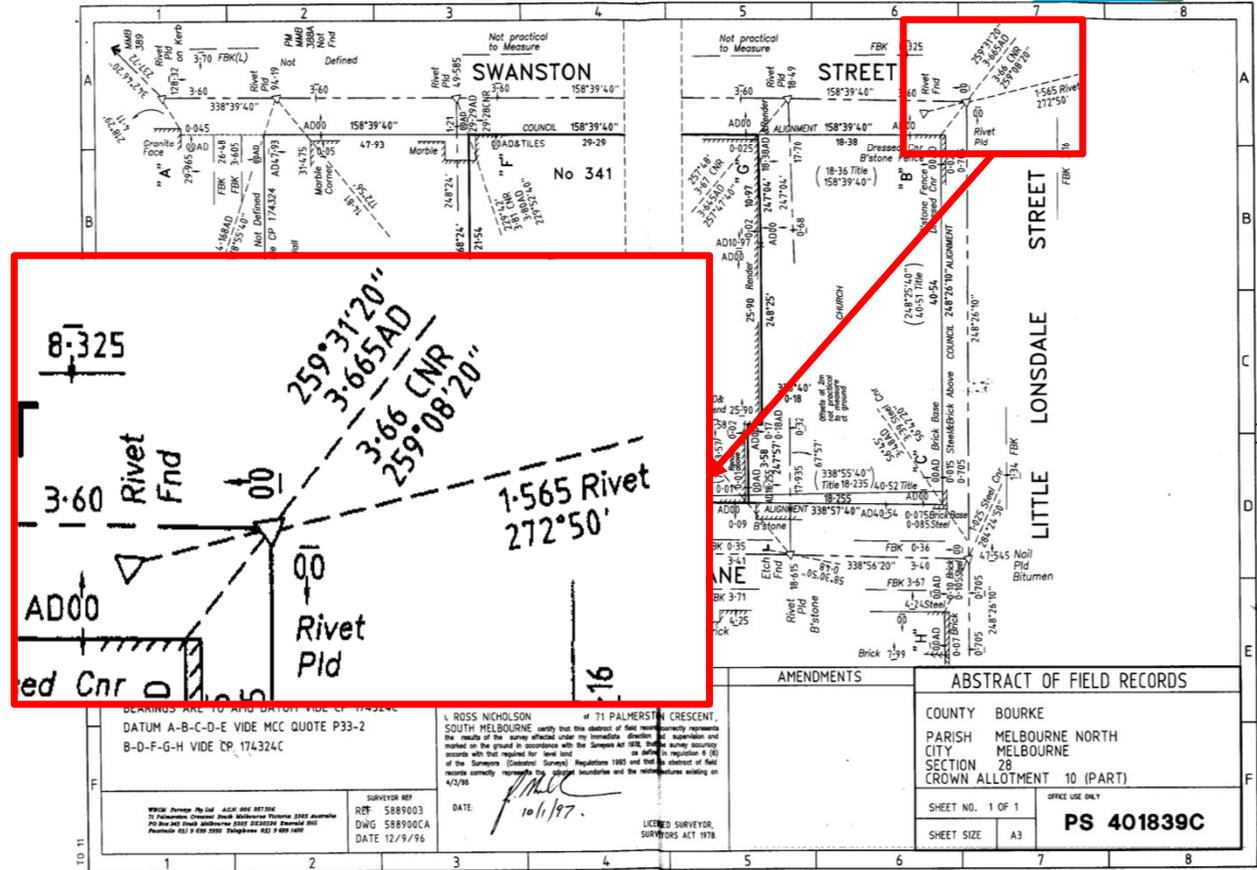
- **Administrative** information
- **Legal** information
- **Survey** information
  
- **Textual** such as notations
- **Geometric** such as boundaries



## Challenges associated with Cadastral Plans and AFRs

- ✘ 1. Dense and detailed
- 2. Fragmentation
- 3. Visualization
- 4. Manipulation
- 5. Storage
- 6. Query

Leading to difficulties to find specific information quickly.





WORKING WEEK 2025

AND

Locate25 THE NATIONAL GEOSPATIAL CONFERENCE

Collaboration, Innovation and Resilience: Championing a Digital Generation



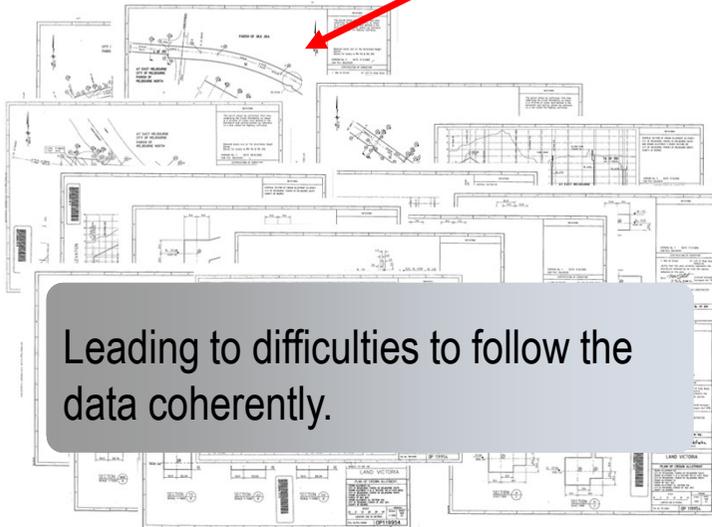
Geospatial Council of Australia

Brisbane, Australia 6-10 April

# Challenges associated with Cadastral Plans and AFRs



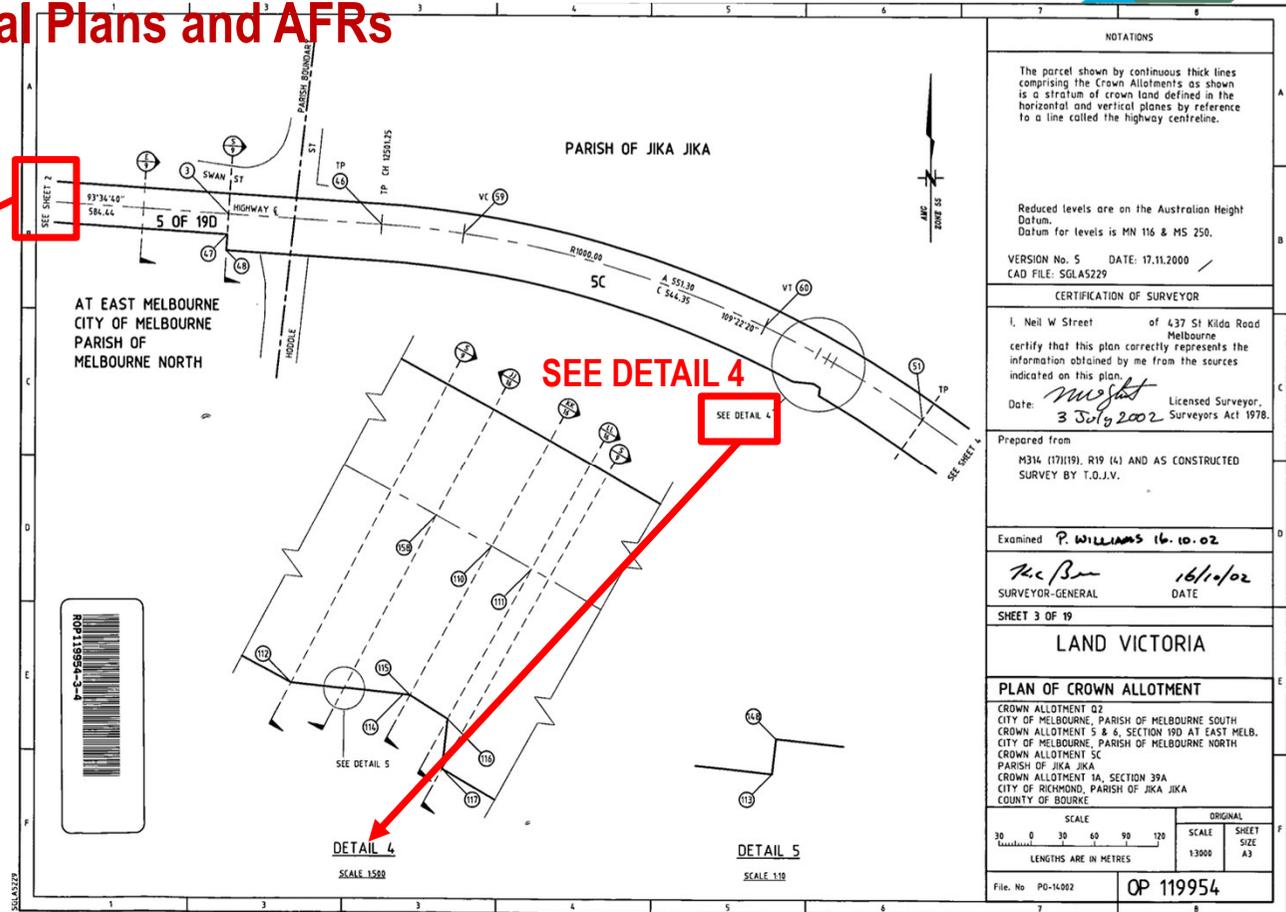
1. Dense and detailed
2. Fragmentation
3. Visualization
4. Manipulation
5. Storage
6. Query



SEE SHEET 2

SEE DETAIL 4

SEE DETAIL 4



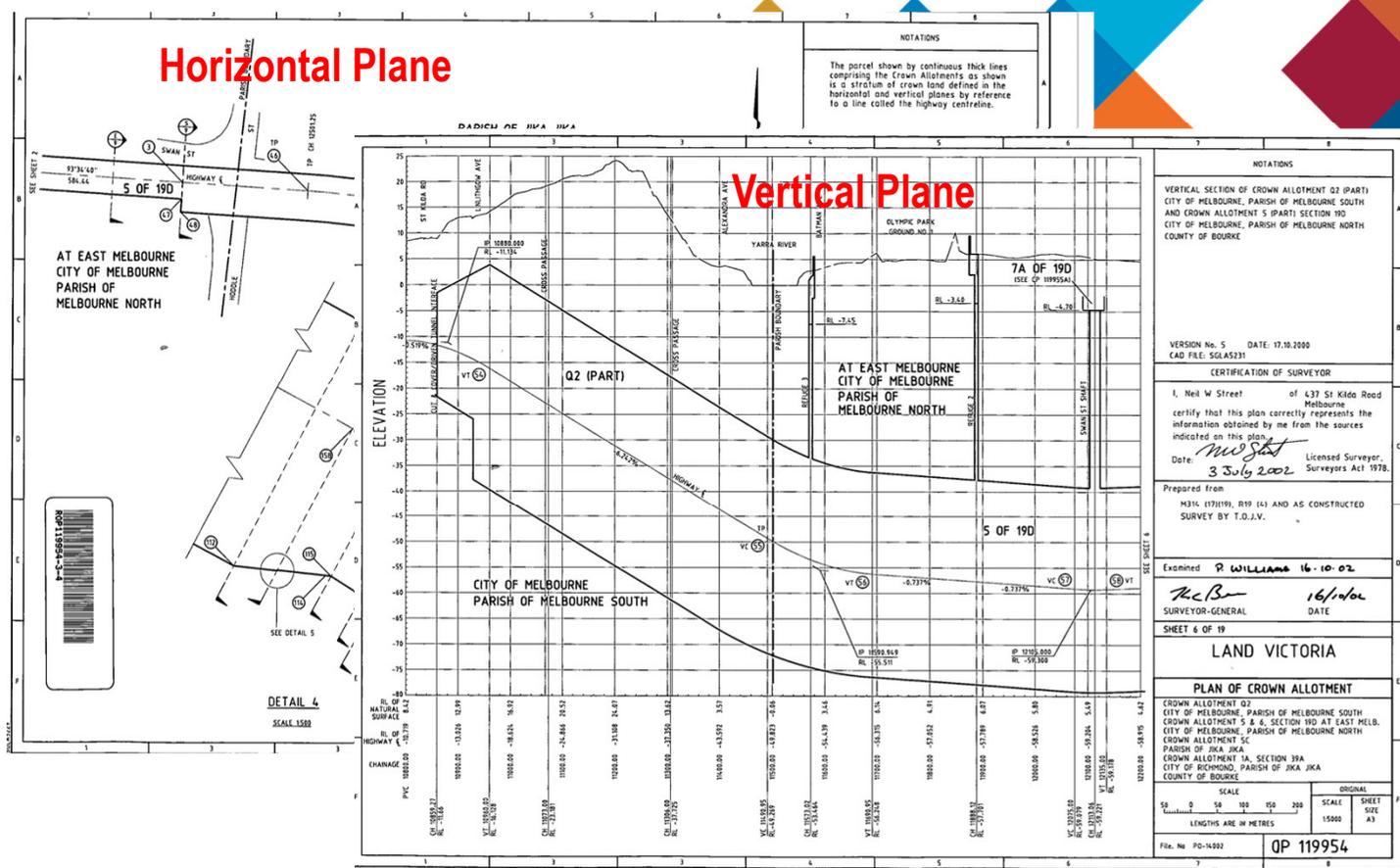
|   |   |
|---|---|
| NOTATIONS   |   |
| The parcel shown by continuous thick lines comprising the Crown Allotments as shown is a stratum of crown land defined in the horizontal and vertical planes by reference to a line called the highway centreline.  |   |
| Reduced levels are on the Australian Height Datum.<br>Datum for levels is MN 116 & MS 250.  |   |
| VERSION No. 5    DATE: 17.11.2000<br>CAD FILE: SGLA5229   |   |
| CERTIFICATION OF SURVEYOR   |   |
| I, Neil W Street    of 437 St Kilda Road<br>Melbourne<br>certify that this plan correctly represents the information obtained by me from the sources indicated on this plan.  |   |
| Date: <i>Neil Street</i><br>3 July 2002    Licensed Surveyor,<br>Surveyors Act 1978.  |   |
| Prepared from<br>M314 (17/11/19), R19 (4) AND AS CONSTRUCTED<br>SURVEY BY T.O.J.V.  |   |
| Examined <i>P. Williams</i> 16-10-02  |   |
| <i>Neil Street</i> 16/11/02<br>SURVEYOR-GENERAL    DATE   |   |
| SHEET 3 OF 19   |   |
| LAND VICTORIA   |   |
| PLAN OF CROWN ALLOTMENT   |   |
| CROWN ALLOTMENT 02<br>CITY OF MELBOURNE, PARISH OF MELBOURNE SOUTH<br>CROWN ALLOTMENT 5 & 6, SECTION 190 AT EAST MELB.<br>CITY OF MELBOURNE, PARISH OF MELBOURNE NORTH<br>CROWN ALLOTMENT 5C<br>PARISH OF JIKA JIKA<br>CROWN ALLOTMENT 1A, SECTION 39A<br>CITY OF RICHMOND, PARISH OF JIKA JIKA<br>COUNTY OF BOURKE |   |
| SCALE<br>30 0 30 60 90 120<br>LENGTHS ARE IN METRES   | ORIGINAL<br>SCALE<br>13000<br>SHEET<br>SIZE<br>A3 |
| File No. PD-14002   | OP 119954   |

## Challenges associated with Cadastral Plans and AFRs

1. Dense and detailed
2. Fragmentation
3. Visualization
4. Manipulation
5. Storage
6. Query



Data representation is limited to 2D floor-plans and cross-sectional diagrams which may lead to ambiguity.



## Challenges associated with Cadastral Plans and AFRs

1. Dense and detailed
2. Fragmentation
3. Visualization
- ×** 4. Manipulation
- ×** 5. Storage
- ×** 6. Query



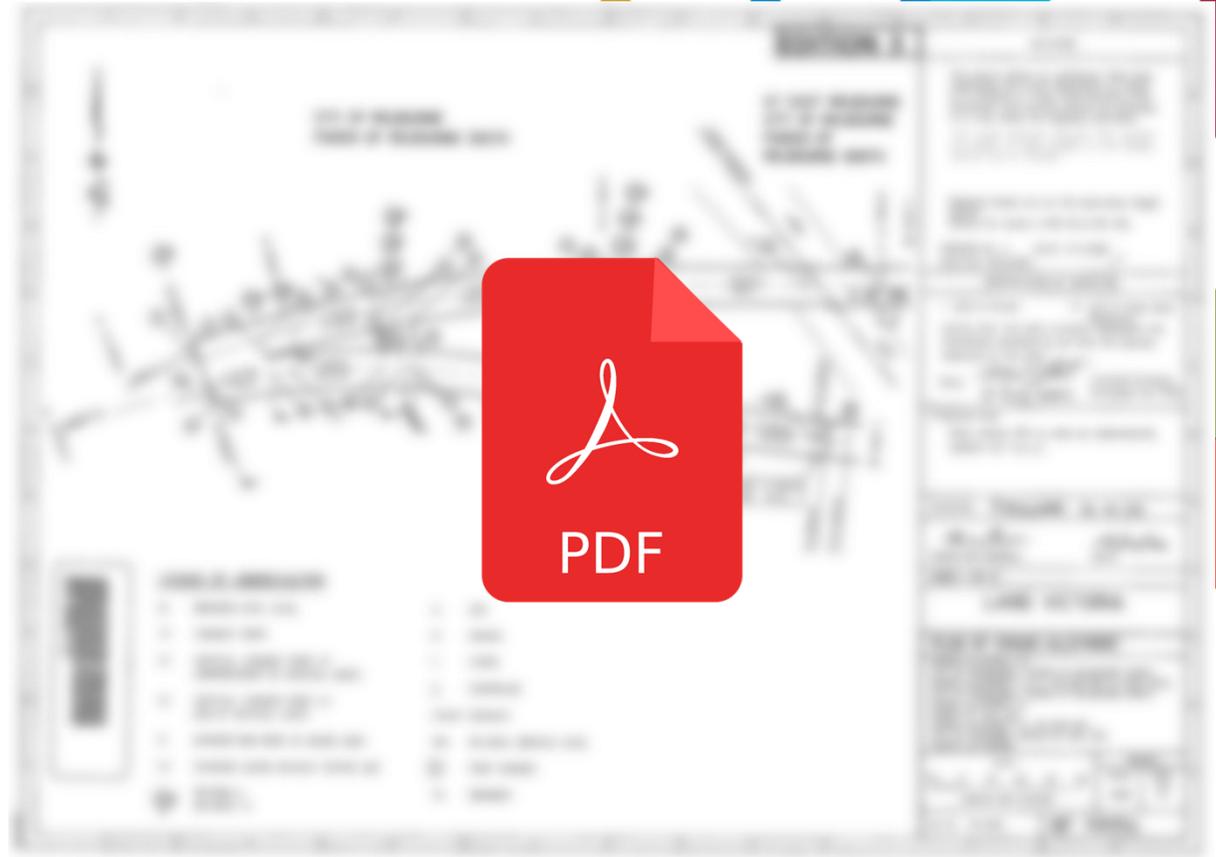
Data is static and recorded once and cannot be updated.



Data is isolated and stored separated from others.

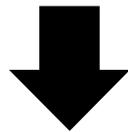


Query is not directly possible.



## Challenges associated with Cadastral Plans and AFRs

Data inside cadastral plans and AFRs is **not easily accessible and readily usable** and significant cognitive load is required to understand and interpret the data/content



Resulting in **slow and inefficient administrative processes** within land development projects that intersect with many land parcels and properties.



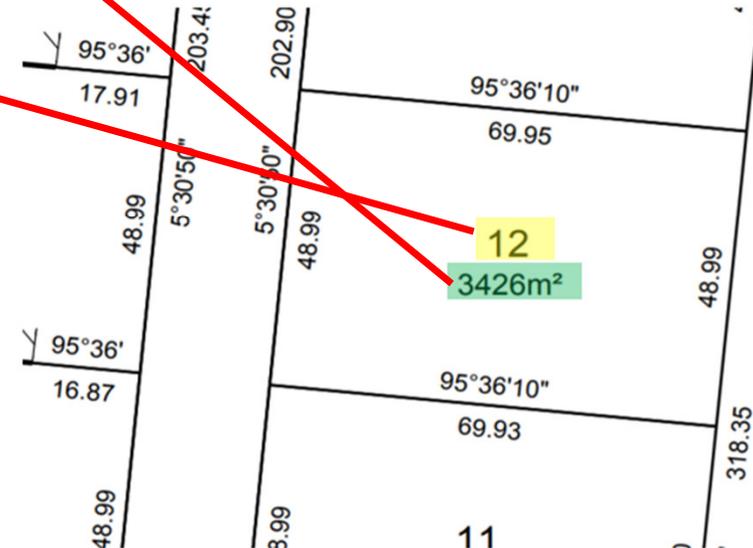
# Current practices

## Mapping into new Data Models



```

<Parcel name="12\PS123456" class="Lot" state="created" area="3426.003" parcelFormat="Standard" parcelType="Single">
  <Center pntRef="CGPNT-603"/>
  <CoordGeom name="CG-15">
    <Line>
      <Start pntRef="CGPNT-450"/>
      <End pntRef="CGPNT-449"/>
    </Line>
    <Line>
      <Start pntRef="CGPNT-449"/>
      <End pntRef="CGPNT-452"/>
    </Line>
    <Line>
      <Start pntRef="CGPNT-452"/>
      <End pntRef="CGPNT-453"/>
    </Line>
    <Line>
      <Start pntRef="CGPNT-453"/>
      <End pntRef="CGPNT-450"/>
    </Line>
  </CoordGeom>
</Parcel>
  
```



# Challenges associated with Data Models

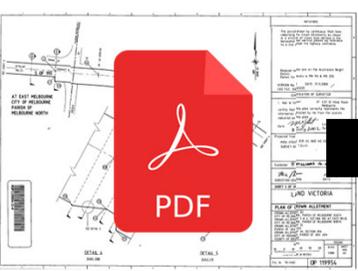


- **Manual Effort** for conversion
  - Most existing cadastral plans have yet to be mapped into these data models.
  - These data models are only capable for knowledgeable stakeholders.
- **Static Interaction** with the system
  - Users need to learn how to use software tools and understand the specific digital model formats.
  - Digital models may not fully integrate automated querying capabilities, requiring users to manually input queries.

```

<xs:schema xmlns="urn:xml-gov-au:vic:icsm:eplan:cif:enumerated-types:1.14" xmlns:xs="http://www.w3.org/2001/XMLSchema"
"urn:xml-gov-au:vic:icsm:eplan:cif:enumerated-types:1.14" elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:simpleType name="addressTypeType">
    <xs:annotation>
      <xs:documentation>Corresponds to IS_PRIMARY field in Vicmap Address</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
      <xs:enumeration value="Primary"/>
      <xs:enumeration value="Secondary"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="adminAreaTypeType">
    <xs:annotation>
      <xs:documentation>Allowable Admin Area types in VIC</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
      <xs:enumeration value="LGA"/>
      <xs:enumeration value="Locality"/>
      <xs:enumeration value="County"/>
      <xs:enumeration value="Parish"/>
      <xs:enumeration value="Township"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="addressPointTypeType">
    <xs:annotation>
      <xs:documentation>Corresponds to GEOCODE_FEATURE in Vicmap Address</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
      <xs:enumeration value="Entrance"/>
      <xs:enumeration value="Virtual"/>
    </xs:restriction>
  </xs:simpleType>
  <xs:simpleType name="adminDateTypeType">
    <xs:annotation>
      <xs:documentation>Allowable Admin Date types in VIC</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
      <xs:enumeration value="Date of Survey"/>
    </xs:restriction>
  </xs:simpleType>

```



Enabling **machines** to mimic **cognition aspects** of **human-intelligence** (i.e., **Vision and Speech**) to **understand** land and property data and **generate** appropriate responses!

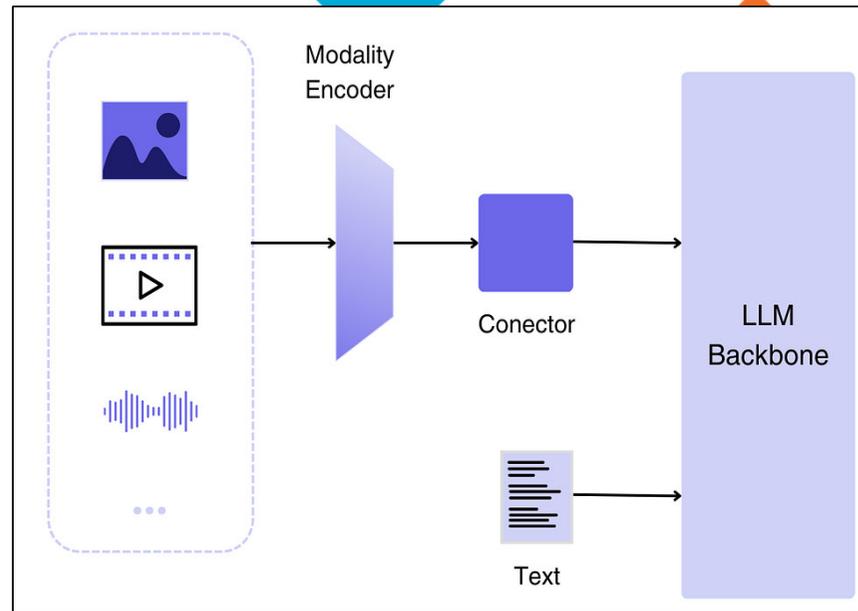


# Attention Mechanism and Transformers for Language and Vision

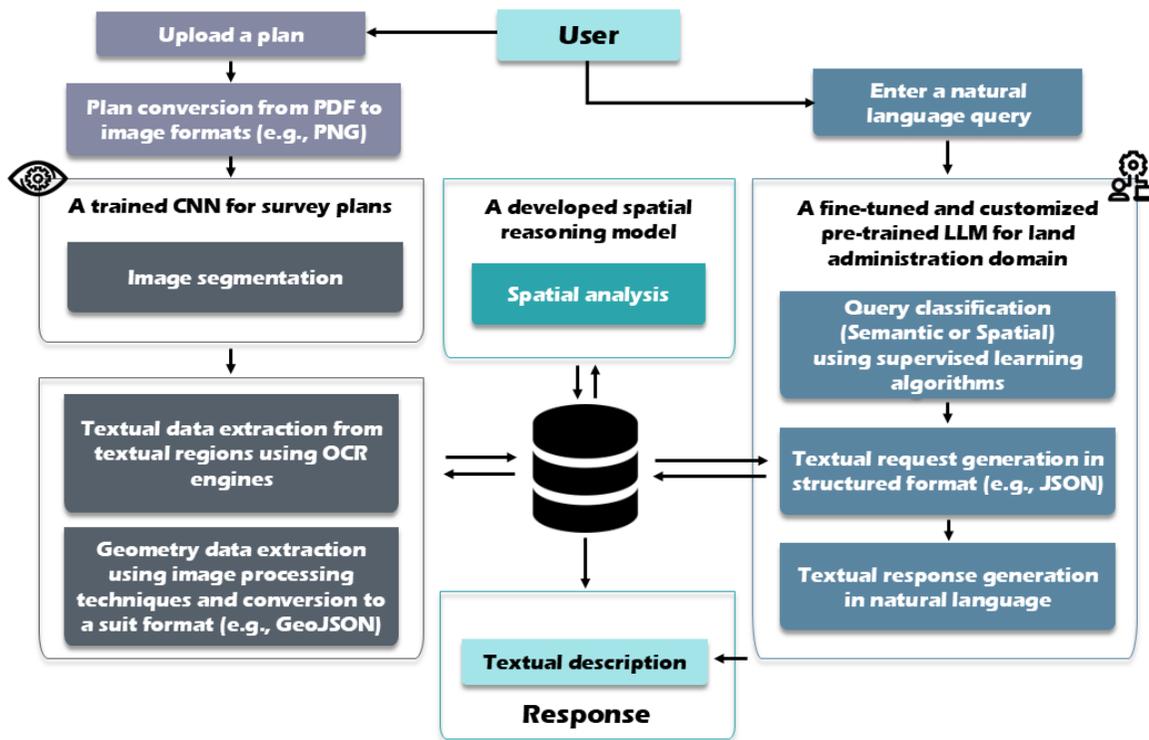
**Large Language Model:** A LM with a high number of parameters that has been trained on a vast textual data and can process sentences, paragraphs, and documents.

**Multimodal Large Language Model:** A LLM that is designed to process and understand multiple types of data, such as text, images, and audio.

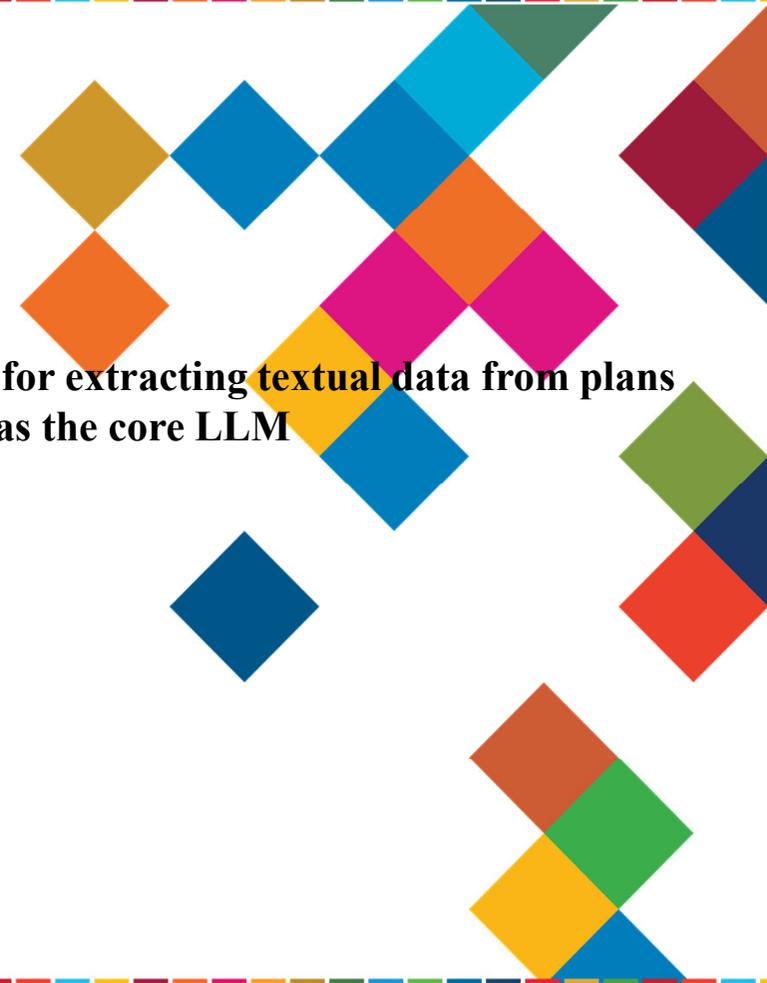
- ✓ Cross-Modal **Understanding**
- ✓ Multimodal **Generation**
- ✓ Enhanced **Interaction**



Architecture of a Typical MLLM



- OCR for extracting textual data from plans
- GPT as the core LLM



# Drawbacks of current MLLMs - examples

Lack of land administration domain knowledge in the core LLM.

what is RRR?

H

how many standard layers are there in SCFF (Single CAD File Format)?

H

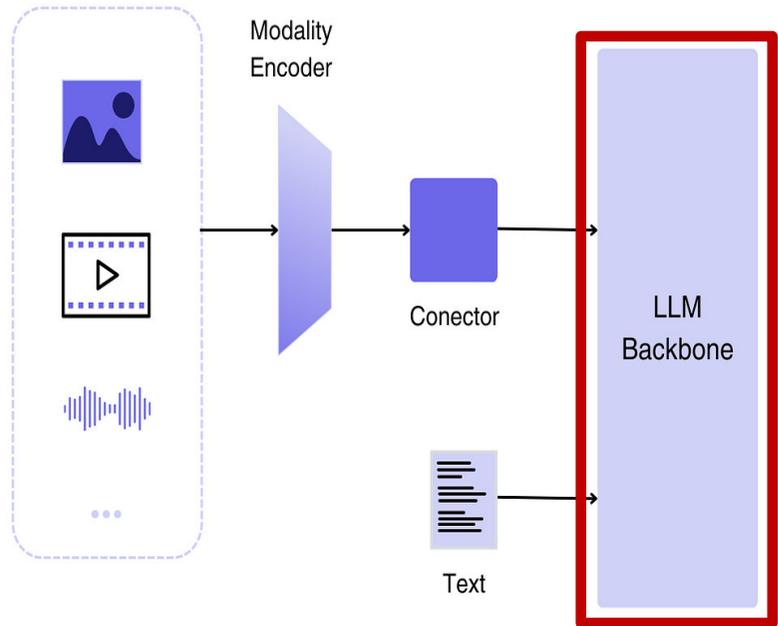
- ⚠ The Single CAD File Format (SCFF) includes a standard of 10 layers. These layers are designed to facilitate the standardized exchange of CAD data between different systems. Each layer in an SCFF is typically assigned to specific types of information or design elements, which helps organize and manage the data more effectively. If you need information on each specific layer and its function, please let me know!



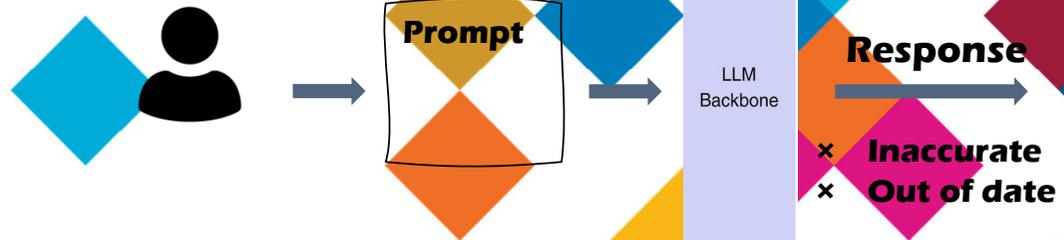
If you have a specific context in mind, please let me know!



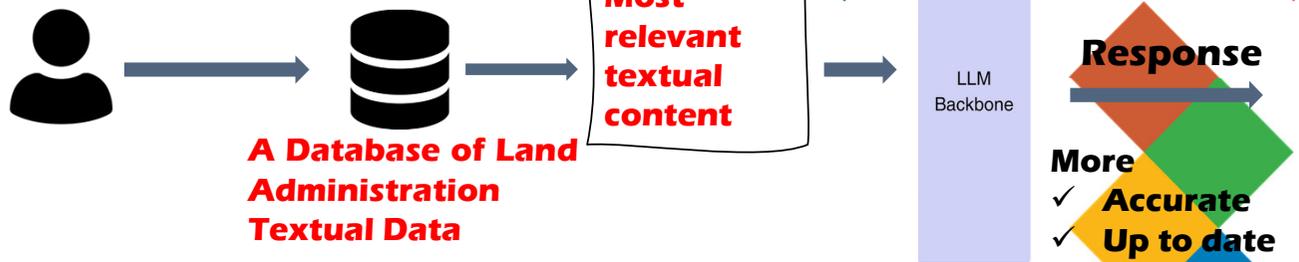
### Enhancement of the **core LLM**



#### 1. Standard



#### 2. Retrieval-Augmented Generation (RAG)



Enhancement of the **core LLM**

### How most relevant contents are identified?

**Embedding** : A numeric representation of each textual data in the core LLM



**Prompt:**

What is RRR?

- Content 1**
- Content 2**
- Content 3**
- ...
- Content m**

**Corresponding embedding:**

[ 0.564,-0.192 ... 0.023]<sub>1Xn</sub>

- Corresponding embedding 1**
- Corresponding embedding 2**
- Corresponding embedding 3**
- ...
- Corresponding embedding m**

Calculate the **similarity** between the vector representation of the prompt and the vector representations of all content items. Then, sort the results based on similarity scores and retrieve the top *k* items.

## Enhancement of the **core LLM**

**Database used:** Oracle

**Similarity method:** Cosine

**Core LLM:** GPT-3.5-Turbo

**Size of embedding vector:** 1536

**Data:** 1901 chunks extracted from

- Scholarly publications such as *Land Use Policy* journal
- Organizational documents such as *CSDILA* and *FIG*
- Governmental documents such as *Land Use Victoria* and *Victorian Legislation*





**WORKING WEEK 2025**

AND

# Locate25

THE NATIONAL GEOSPATIAL CONFERENCE

Collaboration, Innovation and Resilience:  
Championing a Digital Generation



Brisbane, Australia 6-10 April

Chatbot App - AI Chatbot

RAG\_Oracle

localhost:8501/#geoai-powered-chatbot-for-land-surveyors-and-land-registries

Deploy

**GeoAI-Powered ChatBot for Land Surveyors and Land Registries**

Ask me anything in the Land Administration Context

Submit

Please enter a message.

**THE UNIVERSITY OF MELBOURNE**

**CSDILA**  
CENTRE FOR SPATIAL DATA INFRASTRUCTURES AND LAND ADMINISTRATION

This app is developed by the Centre for Spatial Data Infrastructures and Land Administration at The University of Melbourne

ORGANISED BY



PLATINUM SPONSORS



# Initial Prototyping - User Interface

localhost:8501

**THE UNIVERSITY OF MELBOURNE**

**CSDILA**  
CENTRE FOR SPATIAL DATA INFRASTRUCTURES AND LAND ADMINISTRATION

This app is developed by the Centre for Spatial Data Infrastructures and Land Administration at The University of Melbourne

## AI-Based ChatBot for Land Surveyors and Land Registries

Ask me anything in the Land Administration Context

Upload your Plan

Drag and drop file here  
Limit 200MB per file • PDF

Browse files

tunnel-plan.pdf 65.4KB

What's your question about the Plan?

Please provide the name of the licensed surveyor who signed this plan

Submit

Chatbot Response:

Neil W Street

# Semantic queries and corresponding responses

| Query  | Response  | Accuracy score |
|--|---|----------------|
| Is this a plan of subdivision or a plan of consolidation?  | This document is a plan of subdivision. The plan shows the division of Crown Allotments into smaller parcels. The legend indicates "Crown Allotment" and details of the subdivision of those allotments.  | 0              |
| What is the plan number? What is the edition of plan? What is the number of the crown allotment? | The plan number is OP 119954. The edition of the plan is 3. The crown allotment number is Q2.   | 100            |
| What is its city, parish, and country?   | The plan shows the City of Melbourne and the City of Richmond, Parishes of Melbourne South, Melbourne North, and Jika Jika, all located in the County of Bourke. The country is not explicitly stated on the plan, but given the other details, it's implied to be Australia. | 100            |
| Which datum is used for reduced level?   | The Australian Height Datum is used for the reduced levels.   | 100            |
| Who is the Licensed Surveyor?  | Neil W Street of 437 St Kilda Road, Melbourne is the Licensed Surveyor.   | 100            |
| Who is the Surveyor-General?   | P. Williams is the Surveyor-General.  | 0              |
| Who is the examiner?   | P. Williams examined the plan.  | 100            |
| On what date was it signed, and by whom?   | The plan was signed on July 3, 2002, by Neil W. Street.   | 50             |
| What is the scale of the map?  | The main plan scale is 1:3000. There are also details at scales of 1:500, 1:1000, and 1:750 and 1:10. One vertical section is at 1:5000 and another is at 1:5500.   | 100            |
| What does CH mean in the legend abbreviation?  | CH stands for Chainage along the Highway Centre Line.   | 100            |
| How is the parcel shown?   | The parcel is shown by continuous thick lines.  | 100            |
| What is the CAD file number of sheet 2?  | The CAD file number for Sheet 2 is SGLA5228.  | 100            |
| How many sheets does it have?  | The plan consists of 19 sheets. There is also a modification table as a 20th sheet.   | 100            |
| How many abbreviations are there in the legend?  | There are eight abbreviations in the legend.  | 0              |

# Spatial queries and corresponding responses

| Query  | Response   | Accuracy score  |
|--|--|---|
| What are the coordinates of point numbers 17 and 19?               | Point 17: E 321761.289, N 5811548.400, RL -16.601<br>Point 19: E 321775.778, N 5811548.505, RL -38.372   | 50<br>(Do not have RL)  |
| What is the distance between point numbers 17 and 19?              | I cannot calculate the precise distance between points 17 and 19. While the Easting and Northing coordinates are provided, I lack the information necessary to perform a coordinate transformation to accurately compute the distance in meters. This would require knowing the map projection used (e.g., UTM, State Plane) and potentially applying a more complex calculation than a simple Euclidean distance. To obtain the distance, you would need to use a Geographic Information System (GIS) or specialized software that incorporates the necessary projection information. | Can not   |
| Which points have the vertical dimension within their coordinates? | The points listed in the coordinate table on Sheet 18 and Sheet 19 include a "RL" value alongside their Easting and Northing coordinates. "RL" stands for Reduced Level, which represents the vertical dimension (elevation). Therefore, all the points listed in those tables (points 1 through 172) have a vertical dimension included in their coordinate information.  | 0   |
| What is the RL of point number 1?                                  | The RL of point number 1 is -5.398.  | 0   |
| Tell me which point have RL?                                       | I cannot definitively identify which points have RL values without a more thorough and systematic review of the tables on sheets 18 and 19. The OCR results are not perfectly clean and may have missed some RL values, or they may be inconsistently formatted. A manual inspection of the original scanned document is required to confirm which points include RL data.   | Can not   |
| Is point number 59 a VC or TP?                                     | Point number 59 is a VC (Vertical Curve Tangent Point).  | 100   |
| Which type of map scale is used? Linear or ratio?                  | The map uses a ratio scale. A ratio scale (also called a representative fraction) expresses the scale as a ratio, such as 1:3000, meaning one unit on the map represents 3000 units on the ground. This is different from a linear scale which uses a visual bar to show the scale. While the document includes a linear representation of the scale, the fundamental scale used is expressed as a ratio.  | 50<br>(It is not correct for the scale on the right side of the sheets) |

# Spatial queries and corresponding responses

| Query  | Response  | Accuracy score         |
|--|---|------------------------|
| What does DETAIL 4 show?   | DETAIL 4 shows a cross-section view of a portion of the highway and surrounding area, focusing on the area around point numbers 113-117. It highlights the relationship between the highway, the natural surface level, and other details of the subsurface. The detail drawing provides measurements and levels relevant to this specific part of the project.   | 0                      |
| Is DETAIL 4 a horizontal map or vertical?  | DETAIL 4 is a vertical cross-section. It shows a slice through the earth, revealing the vertical relationships between the highway, the ground surface, and other features at that location. It's not a horizontal plan view which would show the spatial layout from above.  | 0                      |
| Is SECTION A/1 a horizontal map or vertical?   | SECTION A/1 is a vertical cross-section. The drawings labelled "SECTION" in this document are all vertical cross-sections showing subsurface details and elevations, not horizontal plan views.   | 100                    |
| How many points are observed in DETAIL 3?  | DETAIL 3 shows seven numbered points: 153, 154, 155, 157, 161, 163, and 165.  | 0                      |
| How many parcels are there?  | Crown Allotment Q2<br>Crown Allotments 5 & 6, Section 19D<br>Crown Allotment 5C<br>Crown Allotment 1A, Section 39A  | 0                      |
| Is there any easement within this plan?  | The provided document does not explicitly mention easements. While the plans show various features and boundaries, there is no text or notation that clearly indicates the presence of easements. To determine whether easements exist, you would need to consult additional documentation related to the land title or other legal records associated with these parcels.  | Can not                |
| Where is the boundary of parcel Q2?  | The boundary of parcel Q2 is shown as a thick continuous line on the provided maps, encompassing an area that extends along the Yarra River. The exact coordinates of the boundary are not explicitly listed, but the boundary is defined by the series of points and curves shown on sheets 1, 2, 3, 4 and 5. The boundary follows both sides of the Yarra River and is impacted by the highway (as the parcel is a stratum of crown land).                                      | 0<br>(Only in sheet 1) |
| With which points does Crown Allotment Q2 intersect? Please list the points that define its perimeter. | I cannot provide a complete list of points defining Parcel Q2's perimeter. The maps are complex, and a precise identification requires a manual inspection of the drawings themselves. The OCR data provides the coordinates of many points but does not explicitly connect those points to Parcel Q2's boundary in every case. To create the complete list, one would need to manually trace the boundary line on the maps and note the numbers of the points it passes through. | Can not                |

## Discussion and conclusion

- LLMs can be beneficial for **automating land and property document analysis**, especially for simple fact-extraction tasks.
- Current LLM can handle **semantic queries** explicitly defined within the plans. However, more extensive fine-tuning and prompt tuning is needed to enhance LLMs' capabilities in handling implicit information like RRR tracing.
- Challenges with **spatial queries** like spatial calculations or contextual understanding to provide meaningful spatial insights.
- Using **other LLMs like Llama** which are on-device, regarding **data privacy issues**.

# The most relevant SDGs related to the presentation and theme of this session

1st relevant SDG



15 LIFE ON LAND

2nd relevant SDG



11 SUSTAINABLE CITIES AND COMMUNITIES

3rd relevant SDG



9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

SUSTAINABLE DEVELOPMENT GOALS

International Federation of Surveyors supports the Sustainable Development Goals

# Thank You for Your Attention!

Hamid Hosseini

[hamid.hosseini@student.unimelb.edu.au](mailto:hamid.hosseini@student.unimelb.edu.au)

Behnam Atazadeh

[behnam.atazadeh@unimelb.edu.au](mailto:behnam.atazadeh@unimelb.edu.au)

Abbas Rajabifard

[abbas.r@unimelb.edu.au](mailto:abbas.r@unimelb.edu.au)



 **My LinkedIn!**