



checkMate

A novel approach to Building Code
Compliance with Large Language Models

Carla Cortez
Redwan Hussain
Kam Kasravi
Edward Kirton
Ryan Wilson

OUR TEAM



Carla Cortez
Software & Web Dev, Blockchain



Redwan Hussain
Industrial Controls & Automation



Kam Kasravi
Artificial Intelligence, Cloud



Ed Kirton
Data Science, Bioinformatics



Ryan Wilson
Analytics, Education Management

MOTIVATION

Address Construction Industry Challenges

- Architects and construction firms experience **delays with plan-checkers**
- Verifying code compliance is very complex, **and still performed manually**
- Regulatory costs can raise the **cost of a home by 24%***

*National Association of Home Builders, 2021



MOTIVATION

Digitizing this process can reduce delays and associated overhead costs for firms and prospective homeowners

Target Audience

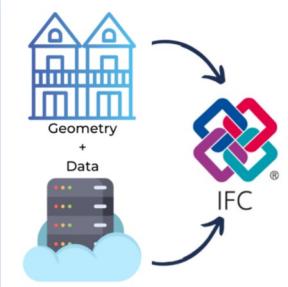
Architecture, Engineering, and Construction (AEC) firms



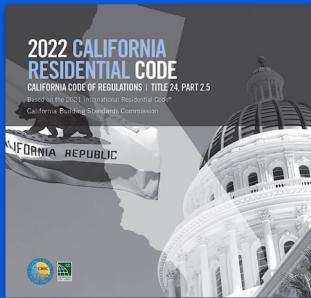
VALUE PROPOSITION

Compare project files against building codes...

Building Information Modeling (BIM) Program Files



Building Standards and Codes



Automated Compliance Reports

Element Details for 277472 (Roof)

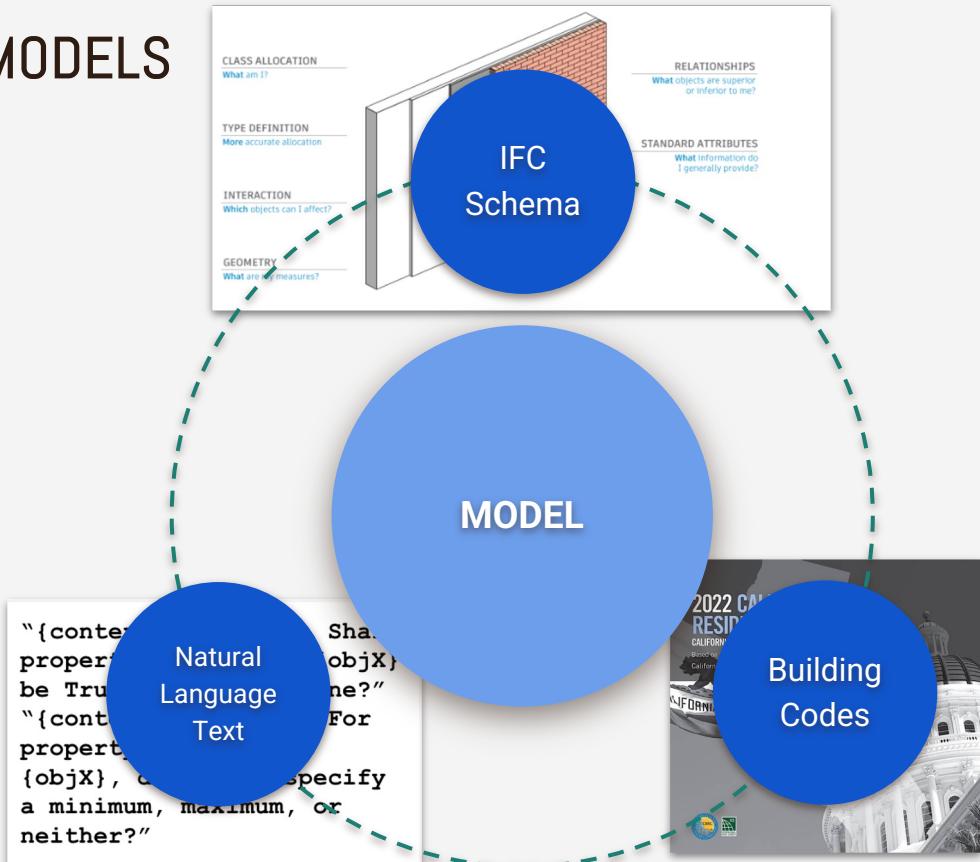
Name	SMRV 9" Unrated	
Flame Spread	C	
Fire Rating(hours)	0	
Related Building Codes		
Section	Result	Reason
602.1	Failed	The roof is required to have a minimum fire rating of (1 hours). The roof has a fire rating of (0.0 hours) which does not meet or exceed the minimum required rating

Red - Building Code
Orange - Ifc input

... and auto-generate compliance reports with targeted feedback for **reduced errors, faster approvals, and saved costs.**

OPPORTUNITY: LARGE LANGUAGE MODELS

- Leverage IFC standard across all BIM applications to ensure uniformity
- Natural language processing of inputs/outputs
- LLM integrated with dedicated building code repository for information extraction

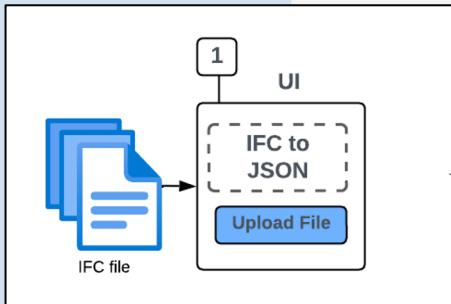


FUNCTIONAL FLOW



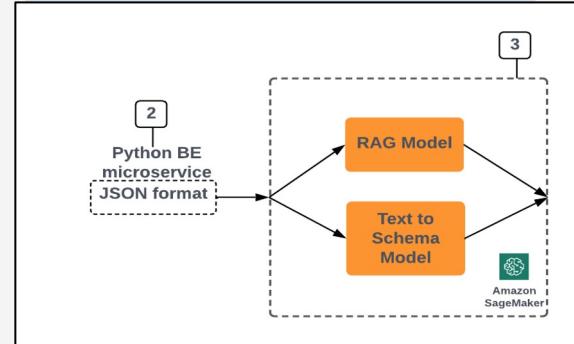
Intuitive UI

UI allows easy selection of building elements that can be submitted for code compliance



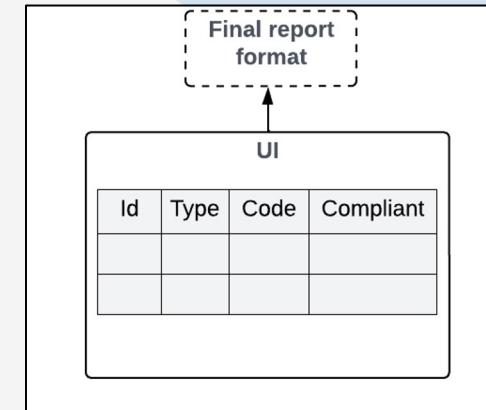
Backend Processing

Building elements are matched with relevant codes. Each element is evaluated for code compliance



Reports

A compliance report is returned to the user





checkMate

APPROACH 1: RAG MODEL

Prompt Example

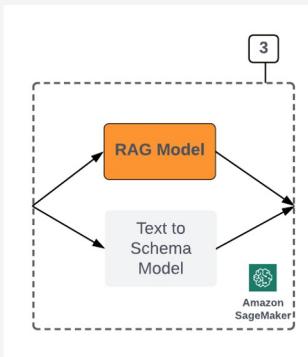
Document (code='R302.2.1Each townhouse unit shall be separated from other townhouse units by two 1-hour ...

+

```
{  
  "type": "Door Metal, Operable"  
  "IsExternal": true,  
  "FireRating": "1.0 hr",  
  "Infiltration": "0.30 m"  
}
```

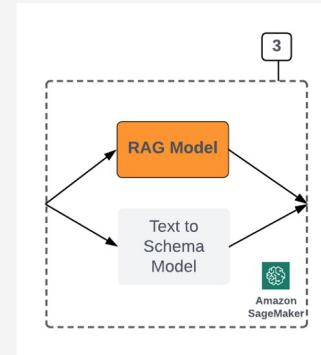
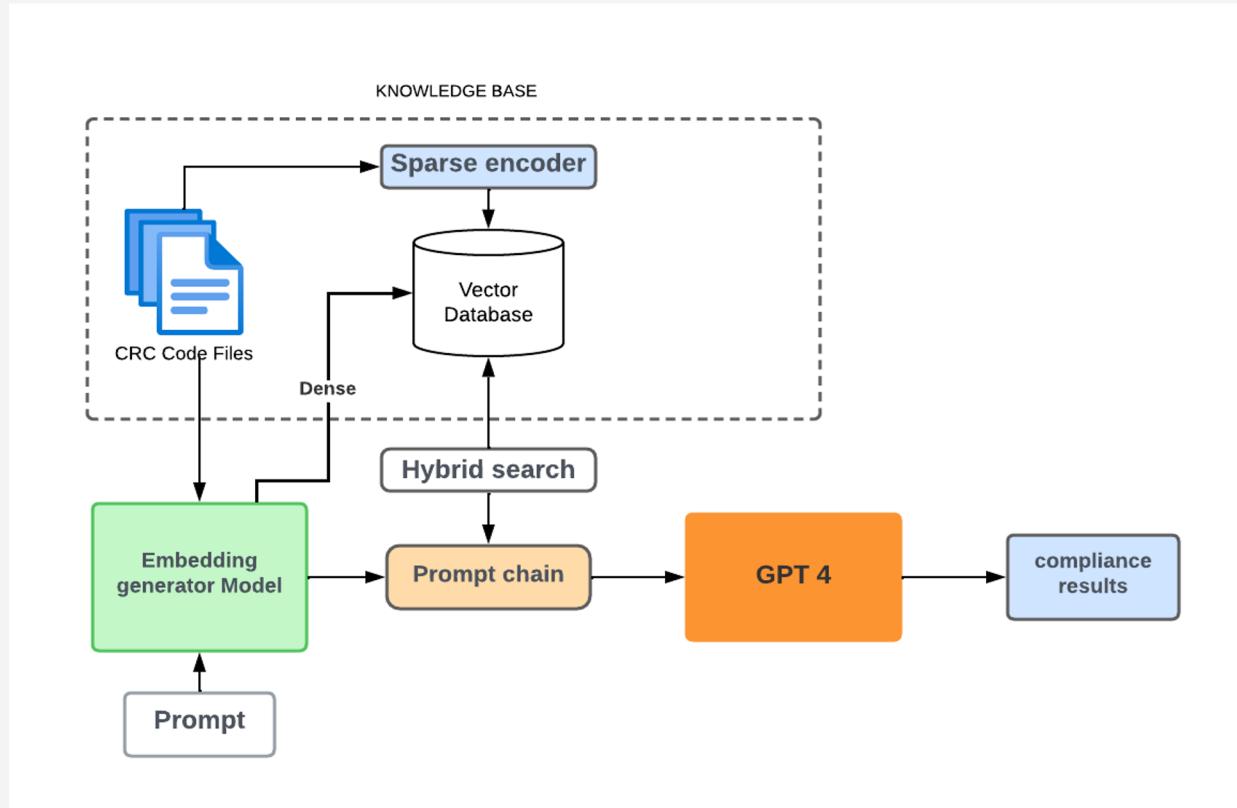
Response

```
{ R337.8.3: {  
    "compliant": True,  
    "summary": The door is  
    external and it has the correct  
    fire rating  
  }  
}
```



APPROACH 1: RAG MODEL

Architecture Diagram



APPROACH 1 - MODEL EVALUATION

Results:

Accuracy  63%

Precision  0.78%

Misclassification analysis:

```
"compliant": False,  
"summary": "The entity does not comply with the building code.  
Although the door has a high fire rating of 4.0 hours,  
it does not meet the specific fire-resistance rating of not less than  
20 minutes when tested according to NFPA 252. The door also does not  
meet the performance requirements of Section R337.7.3.1 and SFM  
Standard 12-7A-1."
```

DISCUSSION, PRO'S/CON'S OF RETRIEVAL-AUGMENTED-GENERATION

RAG

PROs

- Versatility
- Better understanding of context
- Rich Content Generation
- Based on up to date data

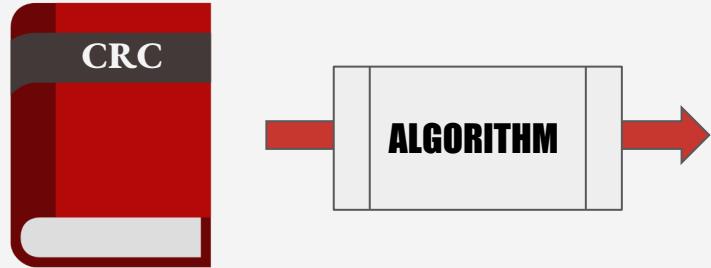
CONS

- Challenging to present the answer in a structured format.
- High Resource Requirements

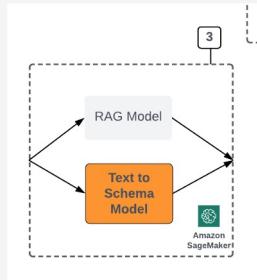
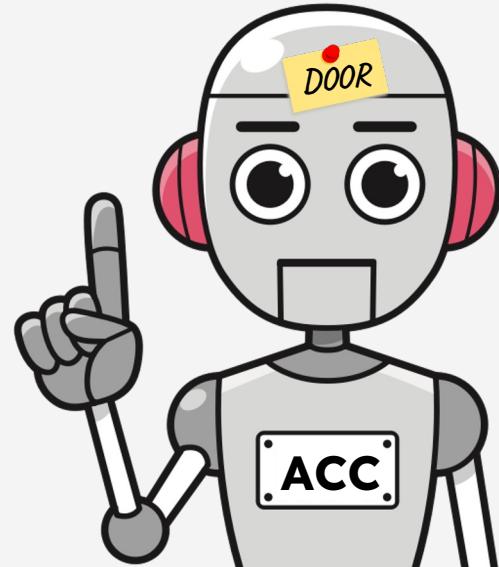
APPROACH 2: TEXT-TO-SQL

The RAG approach asks a LLM to **reason**: “Does <wall-1> satisfy <code-X>?”

A LLM can be used to **extract information** from natural language.
The extracted values may be used to construct SQL queries programmatically.



While You Were Gallivanting		
For	Date	Time
M		
of		
Phone	Fax	
Email		
<input type="checkbox"/> Called <input type="checkbox"/> Returned Call		
<input type="checkbox"/> Called to See You <input type="checkbox"/> Will Call Again		
<input type="checkbox"/> Stopped By <input type="checkbox"/> Will Return		
<input type="checkbox"/> Wants You to Call <input type="checkbox"/> Wants You to Email		
<input type="checkbox"/> Other		
Message:		<input type="radio"/> General <input type="radio"/> Urgent
Signed		



ZERO-SHOT TEXT-TO-SQL

Section R308 Glazing

R308.6.2: Materials

Glazing materials shall be limited to the following:

- Laminated glass with not less than a 0.015-inch (0.38 mm) polyvinyl butyral interlayer for glass panes 16 square feet (1.5 m²) or less in area located such that the highest point of the glass is not more than 12 feet (3658 mm) above a walking surface; for higher or larger sizes, the interlayer thickness shall be not less than 0.030 inch (0.76 mm).
- Fully tempered glass.
- Heat-strengthened glass.
- Wired glass.
- Approved rigid plastics.

```
SELECT GlobalId FROM IfcWindow
```

```
WHERE ( ... OR ... )
```

- Named entity recognition
- List operator: [and | or]

ZERO-SHOT TEXT-TO-SQL

Section R308 Glazing

R308.6.2: Materials

Glazing materials shall be limited to the following:

- Laminated glass with not less than a 0.015-inch (0.38 mm) polyvinyl butyral interlayer for glass panes 16 square feet (1.5 m²) or less in area located such that the highest point of the glass is not more than 12 feet (3658 mm) above a walking surface; for higher or larger sizes, the interlayer thickness shall be not less than 0.030 inch (0.76 mm).
- Fully tempered glass.
- Heat-strengthened glass.
- Wired glass.
- Approved rigid plastics.

```
SELECT GlobalId FROM IfcWindow  
WHERE ( IsLaminated = true OR ... )
```

- Property
- Comparison operator
- Value

ZERO-SHOT TEXT-TO-SQL

Section R308 Glazing

R308.6.2: Materials

Glazing materials shall be limited to the following:

- Laminated glass with not less than a 0.015-inch (0.38 mm) polyvinyl butyral interlayer for glass panes 16 square feet (1.5 m²) or less in area located such that the highest point of the glass is not more than 12 feet (3658 mm) above a walking surface; for higher or larger sizes, the interlayer thickness shall be not less than 0.030 inch (0.76 mm).
- **Fully tempered glass.**
- **Heat-strengthened glass.**
- **Wired glass.**
- **Approved rigid plastics.**

```
SELECT GlobalId FROM IfcWindow
```

```
WHERE ( IsLaminated = true OR IsTempered = true OR IsWired = true );
```

- Score based on number of parsed vs unparsed lines
- Provide unparsed lines to user in report for interpretation

AUTOMATIC CODE COMPLIANCE VIA SQL

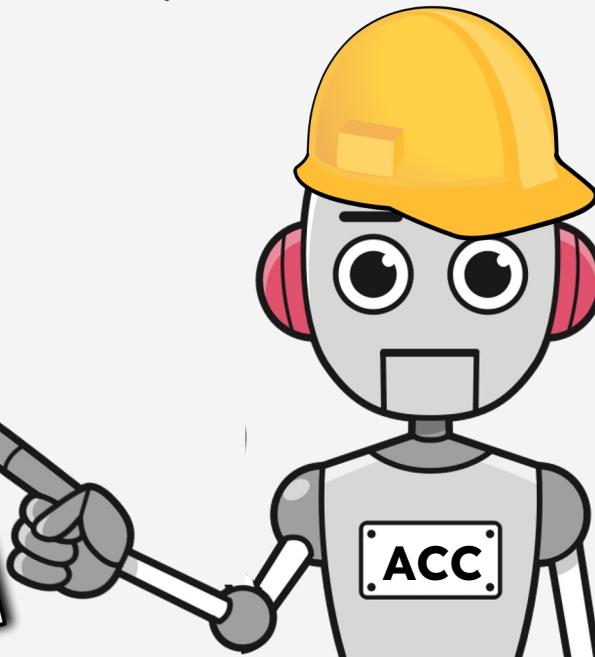
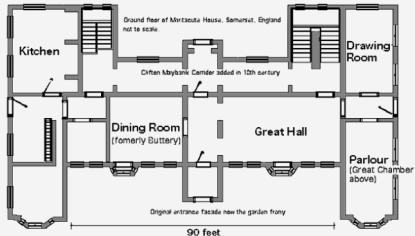
Passing records:

```
SELECT FROM <named_entity>
WHERE <defining_attribute>
AND ( <attribute1> <AND|OR> ... )
OR ( <exception1> OR ... )
```

SQL is very fast!

Failing records:

```
SELECT FROM <named_entity>
WHERE <defining_attribute>
AND NOT ( <attribute1> <AND|OR> ... )
AND NOT ( <exception1> OR ... )
```



Text-to-SQL - EVALUATION

CRC	ENTITY	HUMANS	TEXT-TO-SQL	COMPARISON
R337. 8.3	IfcDoor	(isExternal == TRUE) AND (FireRating >= 20) AND (isTempered == TRUE) OR isLaminated == TRUE OR isCoated == TRUE OR isWired == TRUE)	SELECT GlobalId from IfcDoor WHERE IsExternal = 1 AND (FireRating >= 20)	TTS BETTER
R302. 9.1	IfcDoor	SurfaceSpreadofFlame<200	SELECT GlobalId from IfcWallStandardCase WHERE (SurfaceSpreadOfFlame <= 200)	SAME Ignoring human typo for entity
R302. 1	IfcWall	FireRating >= 60	NONE (tables NYI)	HUMANS BETTER
R305. 1	IfcWall	NominalHeight >=84	SELECT GlobalId from IfcWallStandardCase WHERE (NominalHeight >= 84.0)	SAME
R337. 8.2.1	IfcWindow	FireRating >= 20	SELECT GlobalId from IfcWindow WHERE IsExternal = 1 AND (FireRating >= 20.0 OR IsTempered = 1 OR SmokeStop = 1)	TTS BETTER
R308. 6.2	IfcWindow	IsWired == 1 OR IsLaminated == 1OR isTempered == 1	SELECT GlobalId from IfcWindow WHERE (IsLaminated = 1 OR IsTempered = 1 OR IsWired = 1)	SAME

DISCUSSION, PRO'S/CON'S OF TEXT-TO-SQL

TEXT-TO-SQL

PROs

- Cheap and efficient
- **Rules are explicit, may be corrected or extended by humans**

CONS

- Additional templates+parsers need to be created for other patterns

Useful for driving the co-development of law and file standards.

LESSONS LEARNED

- **Cooperation between IFC and ICC required**
 - Flexible structure of IFC allows for multiple ways data may be encoded
 - Add additional entities and properties needed
 - Some language in ICC is ambiguous or difficult for NLP to parse
- **Extensive Data Engineering effort required**
 - IFC files don't contain all information required for code compliance
 - Building codes are not in computer-friendly format
 - E.g. tables not always square
- **GPT-4 does much better reasoning using natural language rather than structured input**
 - Compliance analysis improved after converting IFC json to natural language
- **GPT-4 challenges related to the Compliance Report**
 - Need to stream ChatCompletion API
 - Need parallel requests

RECAP: INDUSTRY CHALLENGES

- X Architects and construction firms experience delays with plan-checkers
- X Verifying code compliance is very complex, and still performed manually

checkMate VALUE PROPOSITION AND OFFERINGS



SAVE TIME AND COSTS

Automatically produce detailed compliance reports



MINIMIZE ERRORS

Leverage NLP for intuitive prompts and clear feedback responses



FUTURE-PROOF

Refresh model and constraints as regulations evolve over time



THANK YOU

Carla Cortez
Redwan Hussain
Kam Kasravi
Edward Kirton
Ryan Wilson