Knowledge Graphs and Neuro-Symbolic Al

An introduction





• A discussion of the magic predicates that allow you to create and curate knowledge graphs

• A higher level discussion about NeuroSymbolic AI with AllegroGraph.



- Ilm:response
 - Ask a question and force LLM to return list of elements or JSON structures
- Ilm:node
 - String results can be turned into IRIs, needed to create Graph Structures
- Ilm:askForTable
 - Ask for a table all at once
- Ilm:askMyDocuments & Ilm:nearestNeighbor
 - Instead of looking for answer in LLM you use our embedded vectorstore to query your private documents
- Ilm:chatState
 - Build 'stateful' bots: to mix local expertise with the history of a conversation.

Ilm:response as function



Recursive SPARQL fun

Unt	titled-1 🗙 🖬 0 - Hello, GPT 🗙	+ •
1 2 3	<pre># This query demonstrates the use of `response` SPARQL function. # It returns a single string response to the given prompt.</pre>	
4 5	<pre>PREFIX llm: <http: 8.0.0="" allegrograph="" franz.com="" llm="" ns=""></http:> SELECT ?response {</pre>	
6	BIND (llm:response("Write a SPARQL query to extract 10 cities out of dbpedia, only return the SPARQL str: include the prefixes") AS ?response)	ing,
7	}	XECUTE
1	ROW DOWNLOAD RESULTS QUERY INFORMATION	Q SEARCH
res	ponse	
"PR Pre	:EFIX dbo: <http: dbpedia.org="" ontology=""></http:> EFIX dbr: <http: dbpedia.org="" resource=""></http:>	
SEL	LECT ?city	
WH ?c	IERE { sity a dbo:City .	
} LIM	1IT 10"	

Ilm:response as magic predicate



Catalog Repository Statements AllegroGraph root IIm-playground-1 0		? C TEST
0 - Hello, GPT X 🖬 1 - Entry lists X	+ •	Saved queries 🔨
1		💄 0 - Hello, GPT 🍵
2 SELECT ?entry {		💄 1 - Entry lists 🧻
4 # Also try:		💄 2 - Resources 📋
5 # ?entry llm:response "Suggest some names for a male cat."		States
<pre>6 # ?entry llm:response "Enumerate the causes of the American Civil War." 7 # ?entry llm:response "Name the colors of the rainbow."</pre>		• 1 - Borders
8 }	EVECUTE	
	PERECOTE	5 - Untology
		💄 6 - Summari 🔳
50 ROWS DOWNLOAD RESULTS QUERY INFORMATION	Q SEARCH	Query Templates \land
entry	A	View triples
"Alabama"		View quads
"Alaska"		View glasses
"Arizona"		view classes
"Arkansas"		View predicates
"California"		>
"Colorado"	•	

Ilm:node turns strings into repeatable URIs

	AllegroGraph Catalog Repository root IIm-play	Statements ground-1 0		? 🕒	TEST
	0 - Hello, GPT X 1 - Entry lists X	2 - Resources X	+ •	Saved queries	^ ^
	<pre>2 3 PREFIX llm: <http: 2mode="" 2state="" 4="" allegrog="" franz.com="" ns="" pre="" select="" {<=""></http:></pre>		 1 - Entry lists 2 - Resources 	1	
	<pre>5 BIND (llm:response("Largest US state, n 6 BIND (llm:node(?state) AS ?node) 7 # Node can also be ontionally returned</pre>		3 - States		
*	<pre># Node can also be optionally recurred # (?state ?node) llm:response "List US 9 }</pre>	states.".	► EXECUTE	4 - Borders5 - Ontology	i
	1 ROW DOWNLOAD RESULTS QUERY INFO	RMATION	Q SEARCH	6 - SummariQuery Templates	•
	state	node		View triples	
	"Alaska"	llm:node#7a067bb6f837b5fd800a	16dbbd18c629	View quads View classes View predicates	Ŧ
>				>	

Infinitely composable



2 -	Resources X	3 - States X					+ •				
9	PREFIX llm: <h< td=""><td>ttp://franz.com/ns</td><td>— s/allegrograph/8.0.0/</td><td>llm/></td><td></td><td></td><td></td></h<>	ttp://franz.com/ns	— s/allegrograph/8.0.0/	llm/>							
10	SELECT ?state	<pre>?capital ?pop ?are</pre>	ea ?admit ?gov ?celeb	{							
11	{ SELECT ?st	ate { ?state llm:r	response "List the US	<pre>states.". } ORDER E</pre>	BY RAND() LIMIT 10	}					
12	BIND (llm:response(CONCAT("Name the capital of the state of ", ?state, ". Return the city name only.")) AS ?capital).										
13	BIND (llm:response(CONCAT("State the population of the state of ", ?state, ". Respond with a number only.")) AS ?pop).										
14	BIND (llm:response(CONCAT("Tell me the square mile area of the state of ", ?state, ". Respond with a number only.")) AS ?area).										
15	BIND (llm:re	sponse(CONCAT("In	what year was ", ?st	ate, " admitted to 1	the Union?. Respond	with a year only.")) AS ?admit).				
16	BIND (llm:re	sponse(CONCAT("Who	is the governor of	the state of ", ?sta	ate, "?. Respond wi	th the governor's n	ame only.")) AS ?				
	gov).										
17	BIND (llm:re	esponse(CONCAT("Who	o is the most famous	celebrity from the s	state of ", ?state,	"?. Respond with t	he celebrity's name				
	only.")) AS ?c	eleb).									
18	}						► EXECUTE				
19	ORDER BY ?stat	e									
10	ROWS DOWN	LOAD RESULTS Q	UERY INFORMATION				Q SEARCI				
etat		capital	non	2702	admit	001	calab				
Sla						gov					
De	laware	Dover	"9/3/64"	2489	" /8/" "1000"	"Jonn Carney"	"Aubrey Plaza"				
"Ida	iho"	"Boise"	"1826913"	"83569"	"1890"	"Brad Little"	"Aaron Paul"				
"Lo	uisiana"	"Baton Rouge"	"4659978"	"51843"	"1812"	"John Bel Edwards"	"Louis Armstrong"				
"Mi	ssouri"	"Jefferson City"	"6169.91"	"69704"	"1821"	"Mike Parson"	"Brad Pitt"				
"Ne	w Mexico"	"Santa Fe"	"2117522"	"121590"	"1912"	"Michelle Lujan Grish	"Neil Patrick Harris"				
						am"					
"No	rth Carolina"	"Raleigh"	"10439571"	"53819"	"1789"	"Roy Cooper"	"Michael Jordan"				
"No	rth Dakota"	"Bismarck"	"779094"	"70698"	"1889"	"Doug Burgum"	"Josh Duhamel"				
"Oh	io"	"Columbus"	"11.799 million"	"44826"	"1803"	"Mike DeWine"	"LeBron James"				
"Pe	nnsylvania"	"Harrisburg"	"12801989"	"46054"	"1787"	"Tom Wolf"	"Taylor Swift"				
"Uta	ah"	"Salt Lake City"	"3.206 million"	"84899"	"1896"	"Spencer Cox"	"Donny Osmond"				

In many cases *askForTable* is easier ③

States Table	×					+	•			
1 PREFIX	<pre>llm: <http: franz.<="" pre=""></http:></pre>	com/ns/allegrograph/8	.0.0/llm/>							
2										
<pre>SELECT ?state ?capital ?pop ?area ?admitted ?governor ?celebrity {</pre>										
<pre>1 (?state ?capital ?pop ?area ?admitted ?governor ?celebrity) llm:askForTable "Make a table of US States, their</pre>										
capita	ls, population, area	, year admitted to th	e Union, current	t governor and the	e most famous ce	elebrity from that				
state.	}									
						P EXECUT	-			
51 ROWS	DOWNLOAD RESULTS	QUERY INFORMATION				Q SEA	RCH			
state	capital	рор	area	admitted	governor	celebrity				
"State"	"Capital"	"Population"	"Area (sq. mi)"	"Year Admitted"	"Governor"	"Famous Celebrity"				
"Alabama"	"Montgomery"	"4,903,185"	"52,419"	"1819"	"Kay Ivey"	"Hank Aaron"				
"Alaska"	"Juneau"	"731,545"	"665,384"	"1959"	"Mike Dunleavy"	"Sarah Palin"				
"Arizona"	"Phoenix"	"7,278,717"	"113,990"	"1912"	"Doug Ducey"	"John McCain"				
"Arkansas"	"Little Rock"	"3,017,804"	"53,179"	"1836"	"Asa Hutchinson"	"Bill Clinton"				
"California"	"Sacramento"	"39,512,223"	"163,696"	"1850"	"Gavin Newsom"	"Arnold Schwarzenegger"				
"Colorado"	"Denver"	"5,758,736"	"104,094"	"1876"	"Jared Polis"	"Tim Allen"				
"Connecticut	" "Hartford"	"3,565,287"	"5,543"	"1788"	"Ned Lamont"	"Meryl Streep"				
"Delaware"	"Dover"	"973,764"	"1,949"	"1787"	"John Carney"	"Joe Biden"				
"Florida"	"Tallahassee"	"21,477,737"	"65,755"	"1845"	"Ron DeSantis"	"Dwayne Johnson"				
"Georgia"	"Atlanta"	"10,617,423"	"59,425"	"1788"	"Brian Kemp"	"Julia Roberts"				
"Hawaii"	"Honolulu"	"1,415,872"	"10,931"	"1959"	"David Ige"	"Bette Midler"				
"Idaho"	"Boise"	"1,787,065"	"83,569"	"1890"	"Brad Little"	"Aaron Paul"				
"Illinois"	"Springfield"	"12,671,821"	"57,914"	"1818"	"J.B. Pritzker"	"Barack Obama"				
"Indiana"	"Indianapolis"	"6,732,219"	"36,418"	"1816"	"Eric Holcomb"	"Michael Jackson"				
"lowa"	"Des Moines"	"3.155.070"	"56.273"	"1846"	"Kim Revnolds"	"John Wavne"				

Build a wine knowledge graph..

```
INSERT {
  o:root o:hasChild ?typeId.
  ?typeId o:hasChild ?wineId.
  ?wineId o:hasChild ?vintageId.
  ?typeId rdf:type o:WineType.
  ?wineId rdf:type o:Wine.
  ?vintageId rdf:type o:Vintage.
  ?typeId rdfs:label ?type.
                                         ?wineId rdfs:label ?wine.
  ?vintageId rdfs:label ?vintage.
} WHERE {
  (?type ?wine ?vintage) llm:askForTable "Create a table of the top 4 types of wine, the 3 major wines of each
type, 3 vintages of each wine/type. The table should have 4 \times 3 \times 3 = 36 rows."
  bind(concat(?type,"-",?wine,"-",?vintage) as ?vintageName).
  bind(concat(?type,"-",?wine) as ?wineName).
  bind(llm:node(?type) as ?typeId).
  bind(llm:node(?wineName) as ?wineId).
  bind(llm:node(?vintageName) as ?vintageId)
```

};

And write some reviews for each wine

```
INSERT {
  ?vintageId o:hasChild ?reviewId.
  ?reviewId rdf:type o:Review.
  ?reviewId rdfs:label ?review.
} WHERE
{SELECT DISTINCT ?vintageId ?type ?wine ?vintage {
  ?vintageId rdf:type o:Vintage.
  ?wineId o:hasChild ?vintageId.
  ?typeId o:hasChild ?wineId.
  ?vintageId rdfs:label ?vintage.
  ?wineId rdfs:label ?wine.
  ?typeId rdfs:label ?type.}}
  bind(concat("Write three phantom reviews of ",?type," ",?wine," ",?vintage) as ?prompt)
  (?review ?reviewId) llm:response ?prompt.
```

And a fragment...





Extract the graph of connecting US states



The border graph



Gruff	9.0.0 on	AllegroGraph	8.0.0 lim-play	ground-1 rea	id / write 650 t	riples server 12	7.0.0.1:10035			×
<u>F</u> ile	<u>V</u> iew	Te <u>x</u> t Search	<u>D</u> isplay <u>E</u> dit	Global <u>O</u> ptions	Query Options	Ta <u>b</u> le Options	<u>H</u> elp			
		● SPARQL	Dup Quert	Reindent	Name	Query	isit 🔶 🖨	<u>G</u> raph View	Tab <u>l</u> e View	Graphical Qu <u>e</u> ry View
Q <u>u</u> ery	,	○ <u>P</u> rolog	Run <u>Q</u> uery	Select <u>A</u> ll						
sele ?s } limi	ct ?s fr:b t 100	?o where orders ?o O	{ ·							

216 <u>R</u> esults	<u>C</u> reate Visual Graph	Add to Visual Graph	Write Text Report	Save as CSV					
?s			?0						
Wyoming		Idaho		A					
Wyoming		Utah							
Wyoming		Colorado							
Wyoming		Nebraska							
Wyoming		South Dakota							
Wyoming		Montana							
Wisconsin		Michigan	Michigan						
Wisconsin		Illinois							
Wisconsin		Iowa							
Wisconsin		Minnesota							
West Virginia		Virginia							
West Virginia		Pennsylvania							
West Virginia		Ohio							
West Virginia		Maryland							
West Virginia		Kentucky							
Washington		Idaho							
Washington		Oregon							
Virginia		West Virginia							
Virginia		Tennessee							
Visainia		North Carolina							

The border graph visualized





The border graph visualized





The border graph visualized





LLM not good at shortest path ©





\$

Imagine a graph where each state in the US is a node, and for every state that borders that node there is a connection between the two nodes. List the nodes in the shortest path from Maine to California

Content

To find the shortest path from Maine to California on a graph where each state is a node and each border between states is a connection, you would typically use a pathfinding algorithm like Dijkstra's algorithm or A* search. Unfortunately, without the actual graph data or running such an algorithm, I can't compute the exact shortest path.

However, a general shortest path might resemble something like this based on geography:

- 1. Maine
- 2. New Hampshire
- Vermont
- 4. New York
- Pennsylvania
- 6. Ohio
- Indiana
- 8. Illinois
- 9. Iowa
- 10. Nebraska
- 11. Wyoming
- 12. Utah
- 13. Nevada
- California

This list assumes a fairly direct westward path across the United States, avoiding detours south or north. The actual shortest path could vary based on the specific network topology of the state borders. To get an exact path, you would typically input the graph into a GIS

Fight hallucinations with other magic predicates



(?car ?price) askMyTable: list the ten most expensive cars and their prices as integers. Format integers without commas. Omit table header

```
`# The following (four-step) query shows how we can deal with
# potential hallucinations. The task is to create a knowledge
# graph with the 10 most expensive cars and their price according to
# LLM and Google (via the Serp Api)
# Query 1 simply cleans up from a previous run
# Query 2 find ten most expensive cars and their price according to LLM and insert in KG
# Query 3 finds the price according to Google (Serp) and insert in KG
# Query 4 queries the KG to find if the price between LLM and Google is more > 20 % . If yes, it signals a 'Hallucination'
PREFIX llm: <http://franz.com/ns/allegrograph/8.0.0/llm/>
PREFIX franzOption serpApiKey: <franz:cb61b7cb2077c507583ab621e57d3d6b04361b7638c76b7959408e17038e161c>
# Query 1: Clean up from a previous run
delete { ?x ?y ?z } where { ?x a rdf:Car ; ?y ?z . }
# Query 2: Insert 10 cars with their prices according to LLM
# Note how we use the llm predicate llm:askForTable (see documentation)
# to create a table in one go.
insert {
  ?resource a rdf:Car ; rdfs:label ?car ; llm:llmPrice ?price . }
 where
  (?car ?price) llm:askForTable "List the ten most expensive cars and their prices as integers. Format integers without commas. Omit table header"
 bind(llm:node(?car) as ?resource)
```



?answer Ilm:askSerp "find the price of Bugatti Veyron?"

?price IIm:response "Extract the price of the car as an integer from the following text ?answer, respond with an integer only and no text at all. Format something like \$80,000,000 as 8000000.

Query 3: Search for cars prices in Google # So this query first finds all the car in our Knowledge Graph # Then for each car we ask for the price in google using the predicate llm:askSerp. # Unfortunated the returned answer is NOT a number but free-text, # So just for the fun of it we use the llm:response function to retrieve # the actual number as an integer from the text and insert it as llm:serpPrice # we insert the answer also in the graph so you can see how well llm extracted the price insert { ?x llm:serpPrice ?price ; rdf:serpAnswer ?answer } where { ?x a rdf:Car . ?x rdfs:label ?name . bind(concat("What is the price of ", ?name) as ?prompt) . (?answer ?citation ?path) llm:askSerp (?prompt 1) bind(llm:response(concat("Extract the price of the car as an integer from the following text: ", ?answer, "Respond with an integer only and no text at all. Format something like \$8 million as 800000")) as ?price

Find difference between LLM and Google



 But I don't know who to trust, so let's just signal hallucination if difference > 20 %

```
# Query 4: detect hallucinations
# This query finds the serp and llm price for each car and inserts a
# 'Hallucination' object in the graph if the price is > 20 %
# So now another process or human can decide which one to belief.
insert { ?x rdf:hallucination [a rdf:Hallucination; rdfs:label "YES"] } where {
    ?x a rdf:Car ; llm:serpPrice ?serp ; llm:llmPrice ?llm .
    bind(xsd:integer(?serp) as ?number1) .
    bind(xsd:integer(?llm) as ?number2) .
    filter(ABS(?number1 - ?number2) / ((?number1 + ?number2) / 2) > 0.20 )
    }
```

Now look in Gruff, do a select * { ?x a rdf:Car ; ?p ?o } and visualize all triples.



A vector store for search and RAG









Magic predicate: askMyDocuments



NN Te	est 🗙 🖬 Bot3c 🗙			+ •					
3 S	ELECT ?response ?citation ?score	<pre>?content {</pre>							
4 5 b	<pre>4 5 bind("What policy would you suggest to reduce inequality in America?" as ?query)</pre>								
6	Prochance Score Scitation Score	nt)]]m+askMyDocuments ()query	"chomsku	47" 10 0 0)					
8		inc) iim.asknybocuments (iquery	CHOIISKY	47 10 0.0).					
9				► EXECUTE					
10 }									
6 RO	DOWNLOAD RESULTS QUE	RY INFORMATION		Q SEARCH					
respor	nse	citation	score	content					
"To rec include e, stron poor to c awar ould al	duce inequality in America, policies could e a wealth tax, increased minimum wag nger labor rights, and block grants to the o start in commerce or agriculture. Publi reness about regressive state policies sh lso be raised to shift political pressures."	Global_Discontents_Question_15	"8.5315 454E-1"	"What will it take to change U.S. policy?"					
"To rec include e, stron poor to c awar ould al	duce inequality in America, policies could e a wealth tax, increased minimum wag nger labor rights, and block grants to the o start in commerce or agriculture. Publi reness about regressive state policies sh lso be raised to shift political pressures."	Requiem_for_the_American_Drea m_Paragraph_179	"8.3274 47E-1"	"Economic Research: How Increasing Income Inequ Is Dampening U.S. Economic Growth, and Possible s to Change the Tide,"					
"To rec include e, stroi poor to c awar ould al	duce inequality in America, policies could e a wealth tax, increased minimum wag nger labor rights, and block grants to the o start in commerce or agriculture. Publi reness about regressive state policies sh Iso be raised to shift political pressures."	Who_Rules_the_World_Paragraph _307	"8.3267 456E-1"	"More generally, the U.S. public should not learn tha \\\"state policies are overwhelmingly regressive, th nforcing and expanding social inequality,\\\" thoug igned in ways that lead \\\"people to think that the nment helps only the undeserving poor, allowing pc ans to mobilize and exploit anti-government rhetori values even as they continue to funnel support to t etter-off constituents\\\"I'm quoting here from the establishment journal, Foreign Affairs, not from son dical rag.17"					
"To rec include e, stron poor to c awar ould al	duce inequality in America, policies could e a wealth tax, increased minimum wag nger labor rights, and block grants to the o start in commerce or agriculture. Publi reness about regressive state policies sh Iso be raised to shift political pressures."	Requiem_for_the_American_Drea m_Paragraph_180	"8.3266 145E-1"	"The topic of income inequality and its effects has the subject of countless analyses stretching back (ations and crossing geopolitical boundaries. Despi tendency to speak about this issue in moral terms, entral questions are economic ones: Would the U.S nomy be better off with a narrower income gap? Ar					

27

Although sometimes you just want the answer



<	AskMyDocuments	×	shortAMD	×	Query Document-1	×	Query Docum	>	+ NEW QUERY DOCUMENTS -
Vec	tor database to query								
cho	omsky47vecs		•	_					
Que	ry document								
Wh	at are the main cau	ses	of poverty						
Her	e goes a question for you	ır doc	uments						
	EXECUTE QUERY								

Query: What are the main causes of poverty

Answer: The main causes of poverty can be attributed to a combination of factors, including economic inequality, lack of education, political instability, and limited access to resources and opportunities.

Citations:

84

Poverty is the cause of the defects of democracy. That is the reason why measures should be taken to ensure a permanent level of prosperity. This is in the interest of all classes, including the prosperous themselves; and therefore the proper policy is to accumulate any surplus revenue in a fund, and then to distribute this fund in block grants to the poor. The ideal method of distribution, if a sufficient fund can be accumulated, is to make such grants sufficient for the purchase of a plot of land: failing that, they should be large enough to start men in commerce or agriculture.

83

But that development hasn't happened. The main reason is internal. The countries have typically been dominated by small, Europeanized, mostly white elites who are enormously wealthy and linked to the West culturally and economically. These elites do not assume responsibility for their own countries, which leads to horrible poverty and oppression. There have been efforts to break out of that pattern, but they have been crushed.

Or build chatbot applications



Ps	+ NEW QUERY ▼										
1 2 3 4 5 6	<pre>PREFIX lim: <http: 8.0.0="" allegrograph="" franz.com="" llm="" ns=""></http:> SELECT distinct ?response { bind("localhost:10810/aizelvecs" as ?expertise) bind("localhost:10810/multi" as ?history) ?response llm:chatState ("There's nothing you can do for me" ?expertise 10 0.75 ?history 4 0.8) }</pre>										
7	7 EXECUTE										
1	I ROW DOWNLOAD RESULTS QUERY INFORMATION	Q SE	ARCH								
res	ponse										
"I u ur r wh jus	"I understand that you're feeling this way, and I truly empathize with the sense of overwhelm you're experiencing. You might remember in o ur previous sessions, we've talked about how any complex problem can be addressed by dividing it into manageable parts. This helped us when discussing your work stress and family issues. It may not seem like it right now, but I believe we can find a way through this together, just as we have in the past. Whenever you're ready, I'm here to listen and support you."										



- Many hidden variables visible in the magic predicate call.
- Contact us for help if you want to experiment with chatState

В	ot-1 X psychiatrist-full X
1	PREFIX llm: <http: 8.0.0="" allegrograph="" franz.com="" llm="" ns=""></http:>
2	SELECT * {
3	Dind(localnost:10810/alzeivecs as respertise)
5	Chesning Solution S
6	PexpertiseNatchString PexpertiseNatches PexpertiseNints
7	?historyMatchString ?historyMatches ?historyHints
8	<pre>?prompt) llm:chatState ("There's nothing you can do for me" ?expertise 10 0.75 ?history)</pre>
9	
10	
	λ.

10 RC	0 ROWS DOWNLOAD RESULTS QUERY INFORMATION												
cite	expertise	expertiseHints	expertiseMatchString	expertiseMatches	feed	history	historyHints	historyMatchString	historyMatches	prompt	response	score	source
" <htt< td=""><td>"localhost:1</td><td>"Here is a list of citation IDs a</td><td>"Patient: There's nothing you ca</td><td>"citation_id: <http: exp<="" td=""><td>"Patient: There's nothing you c</td><td>"localh</td><td>"Here is a list of citation IDs</td><td>"Patient: There's nothing you c</td><td>"citation_id: <ht< td=""><td>"We are writing a fictional transcrip</td><td>"I understand you're feeling overwhelm</td><td>"8.908792E-</td><td>"40. General Topic:# I</td></ht<></td></http:></td></htt<>	"localhost:1	"Here is a list of citation IDs a	"Patient: There's nothing you ca	"citation_id: <http: exp<="" td=""><td>"Patient: There's nothing you c</td><td>"localh</td><td>"Here is a list of citation IDs</td><td>"Patient: There's nothing you c</td><td>"citation_id: <ht< td=""><td>"We are writing a fictional transcrip</td><td>"I understand you're feeling overwhelm</td><td>"8.908792E-</td><td>"40. General Topic:# I</td></ht<></td></http:>	"Patient: There's nothing you c	"localh	"Here is a list of citation IDs	"Patient: There's nothing you c	"citation_id: <ht< td=""><td>"We are writing a fictional transcrip</td><td>"I understand you're feeling overwhelm</td><td>"8.908792E-</td><td>"40. General Topic:# I</td></ht<>	"We are writing a fictional transcrip	"I understand you're feeling overwhelm	"8.908792E-	"40. General Topic:# I
p://e	0810/aizelv	nd hints about how Doctor sh	n do for me	ertise.be/vdb/id/499>	an do for me	ost:108	and hints about how Doctor	an do for me	tp://history.be/	t of a conversation between Patien	ed right now and it may seem like ther	1"	m stressed
xpert	ecs"	ould respond, based on his k	Doctor: I understand that you're	hint:40. General Topic:	Doctor: I understand that you'r	10/mul	should respond, based on hi	Doctor: I understand that you'r	vdb/id/58> hin	t and Doctor. Here is a list of citati	e's no solution. But remember, we've na		Doctor: If you dont tak
ise.b		nowledge ane experience. B	feeling like this situation is bey	# Im stressed	e feeling like this situation is b	ti"	s previous conversations wit	e feeling like this situation is b	t:43.	on IDs and hints about how Doctor	vigated through challenging times befo		e any action at all, it is
e/vd		ase your response on the styl	ond help, but remember, we've f	Doctor: If you dont tak	eyond help, but remember, w		h this Patient: citation_id: <h< td=""><td>eyond help, but remember, w</td><td></td><td>should respond, based on his kno</td><td>re, like your work stress and family issu</td><td></td><td>exquisitely unlikely the</td></h<>	eyond help, but remember, w		should respond, based on his kno	re, like your work stress and family issu		exquisitely unlikely the
b/id/		e of the Doctor's responses i	aced challenges before and ma	e any action at all, it is	e've faced challenges before a		ttp://history.be/vdb/id/58> h	e've faced challenges before a		wledge ane experience. Base your	es. We did so by breaking down these c		situation will change.
499		n these hints but do not discl	de it through. We've managed y	exquisitely unlikely the	nd made it through. We've ma		int:43.	nd made it through. We've ma	Patient: You go	response on the style of the Docto	omplex problems into smaller, more ma		Patient: Well I thought
>"		ose any patient's personal he	our work stress by breaking it d	situation will change.	naged your work stress by bre			naged your work stress by bre	t me there. Oh	r's responses in these hints but do	nageable parts. This same approach ca		my wife could handle i
		alth information: citation_id:	own into manageable pieces. T	Patient: Well I thought	aking it down into manageabl			aking it down into manageabl	my God.	not disclose any patient's personal	n be applied to your current situation. It		t, Im at work all day.
		<http: 4<="" expertise.be="" id="" td="" vdb=""><td>he same approach can be appli</td><td>my wife could handle i</td><td>e pieces. The same approach</td><td></td><td>Patient: You got me there. O</td><td>e pieces. The same approach</td><td></td><td>health information: citation_id: <htt< td=""><td>may not seem like it right now, but toge</td><td></td><td>Doctor: Thats understa</td></htt<></td></http:>	he same approach can be appli	my wife could handle i	e pieces. The same approach		Patient: You got me there. O	e pieces. The same approach		health information: citation_id: <htt< td=""><td>may not seem like it right now, but toge</td><td></td><td>Doctor: Thats understa</td></htt<>	may not seem like it right now, but toge		Doctor: Thats understa
		99> hint:40. General Topic:# I	ed here. It may not seem like it r	t, Im at work all day.	can be applied here. It may no		h my God.	can be applied here. It may no		p://expertise.be/vdb/id/499> hint:	ther we can find a way forward. Could y		ndable. But now that y
		m stressed	ight now, but we can find soluti	Doctor: Thats understa	t seem like it right now, but we			t seem like it right now, but we	Doctor: You fee	40. General Topic:# Im stressed	ou tell me more about what's making y		oure feeling the stress
		Doctor: If you dont take any a	ons together. Can you share mo	ndable. But now that y	can find solutions together. C			can find solutions together. C	I something? D	Doctor: If you dont take any action	ou feel this way?"		from this situation dor
		ction at all, it is exquisitely un	re about what's causing you to f	oure feeling the stress	an you share more about wha		Doctor: You feel something?	an you share more about wha	on't hold back.	at all, it is exquisitely unlikely the si			t you think your input
		and a second second second second	and the second sec		and the second sec		the second second second second	and the second sec					

Neuro-Symbolic Al The Next Step for Knowledge Graphs





2023: Neuro-Symbolic AI is on the Rise





Source: Gartner © 2023 Gartner, Inc. and/or its affiliates. All rights reserved. 2079794

Enterprise Drivers for Neuro-Symbolic Al



Trust	Limitations of AI models that rely purely on machine learning techniques that focus on correlation over understanding and reasoning. LLMs sometimes give factually incorrect answers or produce unexpected results (hallucinations)
Explainability	The need for explanation and interpretability of AI outputs that are especially important in the regulated industry use cases and in systems that use private data.
Semantics	The need to move toward semantics over syntax in systems that deal with real-world entities in order to ground meaning to words and terms in specific domains.
Future-proof Systems	The integration of multiple reasoning mechanisms necessary to provide agile AI systems eventually leading to adaptive AI systems.
Composite AI Tooling	The set of tools available to combine different types of AI models is increasing and becoming easier to use for developers, data scientists and end users.









What-if Scenarios













Symbolic – Logic, Reasoning, Rules



StrengthsWeaknesses• Precision & Consistency
• Diagnoses, classifications,
predictions are repeatable• Complexity & Maintenance
• Need continuous updates

- Transparency & Explainability
 - Clear insight into how conclusions are reached
- Reliable in well-understood domains

- Limited Flexibility
 - Need more rules if novel cases come up

Dependent on Experts

 Need domain experts that understand their own rules and knowledge and can explain to rule-builder

Neural - Machine Learning, Statistics





LLM – Vector, RAG



Strengths

Natural Language Understanding

• At a level we have never seen before. Reasoning is impressive

All the knowledge in the world

 LLMs integrate knowledge from many different sources and domains

• Up-to-date

 With the latest from OpenAl and Google we now have up-to-date info

Reliability and Accuracy

•

Weaknesses

Never trust any LLM conclusion, prediction, or recommendation.

Contextual understanding Lacks human embodiment for true understanding of the physical world

Ethical and Privacy Concerns

Will this LLM train on my data? Are the results acceptable from an ethical point-of-view

Knowledge Graphs at the Core





Source: https://www.sciencedirect.com/science/article/pii/S266732582100159X



Healthcare Needs Neuro-Symbolic Al



Entity Event Knowledge Graph for Powerful Health Informatics

Ravi Bajracharya Franz Inc. Kathmandu, Nepal ravi@datum.md Richard Wallace Franz Inc. Portland, ME, USA rsw@franz.com

Abstract—This paper introduces the Entity-Event Knowledge Graph (EEKG) model for clinical data stored in graph databases. We describe how the EEKG model dramatically simplifies the representation of patient data, facilitates temporal queries, enables a 360 view of patients and promotes scalability by partitioning patient data into shards. We solved the practical problem that not all clinical data and life science knowledge can be sharded. The solution is to federate each individual shard with common shared data in a knowledge graph. One such shared data source is the UMLS (Unified Medical Language System) knowledge base, which contains genetic, drug clinical trials and Metathesaurus data that we link to individual patient recroix. We report on several use cases including EMR patient retrieval, matching patients with clinical trials, patient control group selection, and care quality measures.

Keywords—entity-event model, knowledge graph, distributed graph database, umls skos knowledge graph, clinical trials knowledge base

I. INTRODUCTION

We describe an EMR and Analytics data system based on the Entity-Event Knowledge Graph (EEKG) model where

patient data is sharded into a distributed graph database and linked to knowledge bases that include facility, provider, payer, coding, as well as medical and scientific knowledge. The approach is primarily based on the earlier work of P. Mirhaji at Einstein Medical College and Montefiore Health System [1]. To provide as realistic a demonstration as possible without relying on confidential patient information, we utilized an open-source synthetic patient generator called Synthea to generate demographic, clinical and claims records for 1 million patients [2]. Using the automated sharding feature of the Allegrograph1 Resource Description Framework (RDF) triple store, we sharded the patient data across a number of servers. We link the sharded patient data to a common knowledge base that also includes UMLS (Unified Medical Language System) and an NIH/NLM project that integrates Jans Aasman Franz Inc. Lafayette, CA, USA jans.aasman@franz.com Parsa Mirhaji Montefiore Medical Center Bronx, NY, USA pmirhaji@montefiore.org

multiple sources of biomedical knowledge, vocabulary and standards [3]. We report on several use cases including EMR patient retrieval, matching patients with clinical trials, patient control group selection, and care quality measures.

II. METHOD

In this section, we will describe our approach to building an EEKG model based on patient health data and linking it to a biomedical knowledge graph constructed from multiple sources of biomedical knowledge bases such as UMLS and the clinicaltrials gov dataset [6]. For the purpose of this demonstration we will use synthetic patient data generated using an open-source project called Synthea [2]. Patient health data concepts will be linked to a knowledge graph by normalising concepts to standard vocabulary which in this case will be the UMLS Metathesaurus.

A. Entity-Event Model

The Entity-Event (EE) model is a method of data organization for information stored in a graph database. By "Entity" we mean an element of the most common or, subjectively, the primary or core class of objects in the graph.



Figure 1 Patient health data and events represented as EE model linked to shared knowledge base

Submitted to the 10th IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS Rochester, Minnesota, USA June 11th -14th. 2022

An event is anything associated with the entity that has a starting time and optionally, an ending time. The authors have used this approach in call centers, where the entities are customers and the events include calls and outcomes; in

IEEE Publication

Entity Event Knowledge Graph for Powerful Health Informatics

Learn & Predict

Diagnosis, Prescribe & Report



Improve Care & Lower Costs

¹ AllegroGraph is a Horizontally Distributed, Multi-model (Document and Graph), Entity-Event Knowledge Graph technology solution from Franz Inc. : https://allegrograph.com/products/allegrograph/

Healthcare Needs Neuro-Symbolic Al



Entity Event Knowledge Graph for Powerful Health Informatics

Ravi Bajracharya Franz Inc. Kathmandu, Nepal ravi@datum.md Richard Wallace Franz Inc. Portland, ME, USA rsw@franz.com

Jans Aasman Franz Inc. Lafayette, CA, USA jans.aasman@franz.com Parsa Mirhaji Montefiore Medical Center Bronx, NY, USA pmirhaji@montefiore.org

Abstract—This paper introduces the Entity-Event Knowledge Graph (EEKG) model for clinical data stored in graph databases. We describe how the EEKG model dramatically simplifies the representation of patient data, facilitates temporal queries, enables a 360 view of patients and promotes scalability by partitioning patient data into shards. We solved the practical problem that not all clinical data and life science knowledge can be sharded. The solution is to federate each individual shard with common shared data in a knowledge graph. One such shared data source is the UMLS (Unified Medical Language System) knowledge base, which contains genetic, drug clinical trials and Metathesaurus data that we link to individual patient records. We report on several use cases including EMR patient retrieval, matching patients with clinical trials, patient control group selection, and care quality measures.

Keywords—entity-event model, knowledge graph, distributed graph database, umls skos knowledge graph, clinical trials knowledge base

I. INTRODUCTION

We describe an EMR and Analytics data system based on the Entity-Event Knowledge Graph (EEKG) model where

patient data is sharded into a distributed graph database and linked to knowledge bases that include facility, provider, payer, coding, as well as medical and scientific knowledge. The approach is primarily based on the earlier work of P. Mirhaii at Einstein Medical College and Montefiore Health System [1]. To provide as realistic a demonstration as possible without relying on confidential patient information, we utilized an open-source synthetic patient generator called Synthea to generate demographic, clinical and claims records for 1 million patients [2]. Using the automated sharding feature of the Allegrograph1 Resource Description Framework (RDF) triple store, we sharded the patient data across a number of servers. We link the sharded patient data to a common knowledge base that also includes UMLS (Unified Medical Language System) and an NIH/NLM project that integrates will be the UMLS Metathesaurus. A. Entity-Event Model

The Entity-Event (EE) model is a method of data organization for information stored in a graph database. By "Entity" we mean an element of the most common or, subjectively, the primary or core class of objects in the graph.

multiple sources of biomedical knowledge, vocabulary and

standards [3]. We report on several use cases including EMR

patient retrieval, matching patients with clinical trials, patient

II. METHOD

EEKG model based on patient health data and linking it to a

biomedical knowledge graph constructed from multiple

sources of biomedical knowledge bases such as UMLS and the

clinicaltrials.gov dataset [6]. For the purpose of this

demonstration we will use synthetic patient data generated using an open-source project called Synthea [2]. Patient health

data concepts will be linked to a knowledge graph by normalising concepts to standard vocabulary which in this case

In this section, we will describe our approach to building an

control group selection, and care quality measures.



Figure 1 Patient health data and events represented as EE model linked to shared knowledge base

¹ AllegroGraph is a Horizontally Distributed, Multi-model (Document and Graph), Entity-Event Knowledge Graph technology solution from Franz Inc. : https://allegrograph.com/products/allegrograph/

Submitted to the 10th IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS Rochester, Minnesota, USA June 11th-14th 2022

An event is anything associated with the entity that has a starting time and optionally, an ending time. The authors have used this approach in call centers, where the entities are customers and the events include calls and outcomes; in (•) 0

Healthcare Needs Neuro-Symbolic Al



Entity Event Knowledge Graph for Powerful Health Informatics

Ravi Bajracharya Franz Inc. Kathmandu, Nepal ravi@datum.md Richard Wallace Franz Inc. Portland, ME, USA rsw@franz.com

Jans Aasman Franz Inc. Lafayette, CA, USA jans.aasman@franz.com Parsa Mirhaji Montefiore Medical Center Bronx, NY, USA pmirhaji@montefiore.org

Abstract—This paper introduces the Entity-Event Knowledge Graph (EEKG) model for clinical data stored in graph databases. We describe how the EEKG model dramatically simplifies the representation of patient data, facilitates temporal queries, enables a 360 view of patients and promotes scalability by partitioning patient data into shards. We solved the practical problem that not all clinical data and life science knowledge can be sharded. The solution is to federate each individual shard with common shared data in a knowledge graph. One such shared data source is the UMLS (Unified Medical Language System) knowledge base, which contains genetic, drug clinical trials and Metathesaurus data that we link to individual patient records. We report on several use cases including EMR patient retrieval, matching patients with clinical trials, patient control group selection, and care quality measures.

Keywords—entity-event model, knowledge graph, distributed graph database, umls skos knowledge graph, clinical trials knowledge base

I. INTRODUCTION

We describe an EMR and Analytics data system based on the Entity-Event Knowledge Graph (EEKG) model where

patient data is sharded into a distributed graph database and linked to knowledge bases that include facility, provider, payer, coding, as well as medical and scientific knowledge. The approach is primarily based on the earlier work of P. Mirhaii at Einstein Medical College and Montefiore Health System [1]. To provide as realistic a demonstration as possible without relying on confidential patient information, we utilized an open-source synthetic patient generator called Synthea to generate demographic, clinical and claims records for 1 million patients [2]. Using the automated sharding feature of the Allegrograph¹ Resource Description Framework (RDF) triple store, we sharded the patient data across a number of servers. We link the sharded patient data to a common knowledge base that also includes UMLS (Unified Medical Language System) and an NIH/NLM project that integrates

A. Entity-Event Model

The Entity-Event (EE) model is a method of data organization for information stored in a graph database. By "Entity" we mean an element of the most common or, subjectively, the primary or core class of objects in the graph.

multiple sources of biomedical knowledge, vocabulary and

standards [3]. We report on several use cases including EMR

patient retrieval, matching patients with clinical trials, patient

II. METHOD

EEKG model based on patient health data and linking it to a

biomedical knowledge graph constructed from multiple

sources of biomedical knowledge bases such as UMLS and the

clinicaltrials.gov dataset [6]. For the purpose of this

demonstration we will use synthetic patient data generated using an open-source project called Synthea [2]. Patient health

data concepts will be linked to a knowledge graph by normalising concepts to standard vocabulary which in this case

In this section, we will describe our approach to building an

control group selection, and care quality measures.



will be the UMLS Metathesaurus.

Figure 1 Patient health data and events represented as EE model linked to shared knowledge base

¹ AllegroGraph is a Horizontally Distributed, Multi-model (Document and Graph), Entity-Event Knowledge Graph technology solution from Franz Inc. : https://allegrograph.com/products/allegrograph/

Submitted to the 10th IEEE INTERNATIONAL CONFERENCE ON HEALTHCARE INFORMATICS Rochester, Minnesota, USA June 11th-14th, 2022

An event is anything associated with the entity that has a starting time and optionally, an ending time. The authors have used this approach in call centers, where the entities are customers and the events include calls and outcomes; in









% Facts obese(patient). smokes(patient). frequent_heartburn(patient). regurgitation(patient). stressful life(patient). eats large meals(patient). has hiatal hernia(patient). experiences nausea(patient). has sore throat(patient). % Risk Levels risk of_gerd(Patient, high) :obese(Patient), (smokes(Patient); has_hiatal_hernia(Patient)), (frequent heartburn(Patient); regurgitation(Patient)), eats_large_meals(Patient), !. risk_of_gerd(Patient, moderate) :-(stressful life(Patient); eats large meals(Patient)), (frequent_heartburn(Patient); regurgitation(Patient); experiences_nausea(Patient)), \+ risk_of_gerd(Patient, high), !. risk_of_gerd(Patient, low) :-(has sore throat(Patient); experiences nausea(Patient)), \+ risk_of_gerd(Patient, high), \+ risk of gerd(Patient, moderate), !. risk of gerd(Patient, increased) :smokes(Patient), obese(Patient), risk of gerd(Patient, moderate), !. % Default case if none of the above conditions are met risk_of_gerd(_, unknown).

Symbolic

Diagnose, Prescribe & Report

Using queries, rules, and programs: Predicting GERD using logic









LLM / Vector

NLP + Literature

Based on medical history of patient and 36 M articles in pubmed an LLM predicts Future Events or explains output of machine learning

























Creating a PatientGraph with AllegroGraph



Example Open Source Knowledge Graph



Example Open Source Knowledge Graph



80% of Medical Data is Unstructured





* Adverse vaccine reactions

** Synthea

























A patient with future events from LLM and RNN





Thank You

Dr. Jans Aasman, CEO of Franz Inc.

ja@franz.com

