

Verifiable Credentials with expressive zero knowledge query JESSE WRIGHT | FEB 2025





DPhil @ Oxford

Building neurosymbolic Web agents that

- Represent legal entities
- Support user autonomy
- Protect user data







Solid Lead via Open Data Institute

<u>Solid</u> is an open standard for managing digital identities and storing personal data for re-use across applications on the Web. The goal of Solid is for people to have more agency over their data.







inrupt



Formerly working on Solid @ Inrupt

- Enterprise software engineer and data architect
- Inrupt donated a Solid-based data wallet to the Open Wallet Foundation





W3C

Groups 14

Decentralized Identifier Working Gr... Linked Web Storage Working Group Notation 3 (N3) Community Group RDF Dataset Canonicalization and ... RDF JavaScript Libraries Commun... RDF Surfaces Community Group RDF Test Suite Curation Communi... RDF-star Working Group SHACL Community Group

Х





What is a wallet?



A way to prove **someone** said **something**



A *generalisation* is to supply **some evidence** that you can **take something to be true**

We call this provenance

The two core features (in my opinion) of (W3C) Verifiable Credentials



The two core features (in my opinion) of (W3C) Verifiable Credentials





But lets consider the use cases ...



Secondary Data Reuse

gaia-x

~ ~ ~ / / / /

~ ~ ~ / / / /

;;;;;;

- - (



Give me an anonymised dataset of the **age** and **BMI** of all men in the UK



On demand data integrity



Prove that you're eligible to hire a car in this country.



Prove that you're eligible to hire a car in this country.

What if the company didn't need to write bespoke business logic integrating age, driving, and travel (e.g. visa) credentials? Can the standards make this data integration easier?

Can the standards make this data integration more privacy preserving?



Can we support zero knowledge proof over arbitrary SPARQL query?



```
:UKDrivingAuthority :claims <<:Jesse :dob "06-00-2000"^^xsd:dateTime>> .
    :signature [...] .
:UKImmigrationAuthority :claims <<:Jesse :hasCitizenship :Australia>> .
    :signature [...] .
:UKImmigrationAuthority :claims <<:Australia a :CommonwealthCountry>> .
:signature [...] .
I want to be able to execute the following SPARQL ASK query (API to be refined). I also want to be able to execute
all other read-only SPARQL operations (SELECT and CONSTRUCT).
ASK {
    :Jesse :dob ?x .
        :hasCitizenship [ a :CommonwealthCountry ]
    FILTER (?x >= "03-01-2006"^^xsd:dateTime)
}
```

And have a zero-knowledge proof proving that the statement is true if you trust claims issued by :UKDrivingAuthority and :UKImmigrationAuthority .



RDF and Query Overview



https://eyereasoner.github.io/eye-js/example/_____

https://query.comunica.dev/



🛃 SPA	RQL	🕸 Graph C	QL-LD				
1 •	CONS	TRUCT {		-			K J
) 2	<pre>ruben:me dbpedia-owl:sharesBirthPlace f</pre>					?person ;	
3		dbpedia-0	owl:ol	lderThan ?person	•		
4 •	} WH	ERE {					
5	ru	ben:me dl	opedia	a-owl:birthPlace	?place	;	
6	-	di	opedia	a-owl:birthDate	?dateR		
7	?p	erson di	opedia	a-owl:birthPlace	?place;		
8			opedia	a-owl:birthDate	?date.		
10	1	LIER(?dai	te > 4	(datek)			
10	3						
						Press (TRI – <spacebar< td=""><td>> to autocomplete</td></spacebar<>	> to autocomplete
Exec	ute d	query					18 results in 0.4
Ouerv	res	ults					
<i>query</i>	163	uits.					
uhan'r	ne dh	nedia-owl:	charoc	BirthPlace dhoedia:	Rirgor Vo	rstraata.	
100111	ne ub	peula-owi.	31101 25	bil the lace uppedia.	Singer_ve	isciaete,	

dbpedia-owl:olderThan dbpedia:Birger_Verstraete; dbpedia-owl:sharesBirthPlace dbpedia:Brecht_Capon; dbpedia-owl:olderThan dbpedia:Brecht_Capon; dbpedia-owl:sharesBirthPlace dbpedia:Dieter_Wittesaele; dbpedia-owl:olderThan dbpedia:Dieter_Wittesaele; dbpedia-owl:sharesBirthPlace dbpedia:Divock_Origi; dbpedia-owl:olderThan dbpedia:Divock_Origi; dbpedia-owl:sharesBirthPlace dbpedia:Igor_Vetokele; dbpedia-owl:olderThan dbpedia:lgor_Vetokele; dbpedia-owl:sharesBirthPlace dbpedia:Jelle_Florizoone; dbpedia-owl:olderThan dbpedia:Jelle_Florizoone; dbpedia-owl:sharesBirthPlace dbpedia:Kimmer_Coppejans; dbpedia-owl:olderThan dbpedia:Kimmer_Coppejans; dbpedia-owl:sharesBirthPlace dbpedia:Niels_Coussement; dbpedia-owl:olderThan dbpedia:Niels_Coussement; dhadia-owlicharacRirthPlace dhadia:Oliver Naecen



RDF 1.2 Overview









```
PREFIX : <file:///Users/jesght/Documents/GitHub/jeswr/rust-test/bar/example.spargl#>
PREFIX dbpedia-owl: <http://dbpedia.org/owl/>
PREFIX ruben: <https://ruben.verborgh.org/>
CONSTRUCT {
  [] :asserts <<ruben:me dbpedia-owl:sharesBirthPlace ?person; dbpedia-owl:olderThan ?person>> .
     :proof ?derivedProof .
  BIND_PROOF(?derivedProof)
} WHERE {
  ?e1 :asserts <<ruben:me dbpedia-owl:birthPlace ?place; dbpedia-owl:birthDate ?dateR>> .
     :proof ?proof1 .
  ?e2 :asserts <<?person dbpedia-owl:birthPlace ?place; dbpedia-owl:birthDate ?dateR>> .
     :proof ?proof2 .
  FILTER(?date > ?dateR)
 FILTER(?e1 a :EUGov)
 FILTER(?e2 a :EUGov)
```



The derived proof should reveal **at most*** which entities **stated** the facts used to **establish** the **derived facts**.



Now let's talk Zero Knowledge Proof



Naïve approach

- Take a Zero Knowledge Virtual Machine (ZKVM)

- Add a <u>SPARQL Query</u> Engine

- Done (?)

The RISC Zero zeroknowledge virtual machine zkVM (zkVM) lets you prove correct execution of arbitrary Rust code.

Anyone with a copy of the receipt can verify the guest program's execution and read its publicly shared outputs.





Naïve approach

https://github.com/jeswr/risc0-sparql-poc/blob/main/core/src/lib.rs



RISC	
7 F R O	
1771mm: ^27366B8@85	71
·7PG.1^~·YP.15.17.1B@@@@@@#·	· VRD7 ·
	. TOP7.
	51&@@@@R7
	\$.&@@@@@@}7. &&@@@@@@@@#7
	7.000000000000000000000000000000000000
	000000000000000000000000000000000000000
	000000000000000000000000000000000000000
	&@@@@GJY&&7#
Y7 G@@@@@@@@@@@@@@@@@@&J!!5Y: !BP#@P?!?	YPG@#Y??5@B~
7Y ^@@@@@@@@@@@@@B!^7:. 7B^^7!	~!:P!Y###@@B
.P. ?@@@@@@@@@@@@@P!:	P7::!:::J@@@
	@@&&@@&&PB@@
P~ 75@@@5J5J#P .Y&@@@@@@@@@@@@@@	ବୂଜ୍ଭ୍ଭ୍ଭ୍ଭ୍ଭ୍ଭ୍ର୍୍୍୍୍୍୍୍୍୍୍୍୍୍୍୍୍୍ରେ କୁଦ୍ର୍
.G. ^@@P ^? 7@@@@@@@@@@@@@@@@@@	00000000000075
.G. Y@#~?P^^~J?^!:. J@@@@@@@@@@@@@@@@@	000000000000000000000000000000000000000
.G. ^JG@G!:~J7^ 7#@@@@@@@@@@@@@@@@	000000000000000000000000000000000000000
5~ .^Y@? ^~^ :5&@@@@&B@@@	000000000000000000000000000000000000000
7J ~YY7?@@@&&B! :!~:~. ~7E	B@@@@@@@@@@@
.G 7@@@@@@@##B? ()	G@@@@@@@@@#!
.B@@@@@@@@@@@@05Y7^^.	: B@@@@@@@@?
G@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@	G@@@@@@@@!
5!	P@@@@@@B7~G
.Y7 ~#@@@@@@@@@@@@@@@@@@@@	5@@@@@@P .P7
?J. :^J&@@@@@@@@@@@@@@@	Y@@@@&7
~Y! ?@@@@@@@@Y!:	7@&G?: ~
.7J~^&@@@@@@B:	:^: ^?7
.7J~. J@@@@&5^	^?7.
.!?7^ .P@@G~	:!7~.
:!?7^. J@B.	.^!7!:
	~





Gotcha's

- Proving time
- Proof is code dependent
- Hash and proof description is not part of SPARQL result

RISC		Contraction of the local division of the loc
ZERO	************	~~^^:
	:^!777!~~~^^!^^::	:^^~~!!7!~^:
	:!77!~~::^?7JGGB&@&5	.:^!77!.
	:7PGJ^~:YPJ5J7JB@@@@@#:	:YBP7:
	.!P&@@5JP?5GGP5. 7@@@@@@B^	:?J~^: J@@&P!
	.!Y@@@@@@#5PG5JB@Y .P@@@&GGJ!~	~B@&B@5!&@@@@@B7.
	.7JJP@@@@@@#@P!!YBGP 7@@@5^ .!?	7. ?@@G:#&&@@@@@@@@#7
· · · ·	Y~7@@@@@@G?.:!GP?P~ .YB5	!! 7P5@5.7@@@@@@@@@@
?3	J.7@@@@@@&~ G@@YG^ .	!GBP^ 7&GJ#@@@@@@@@@@
	Y@@@@@@@@W~G@@@@@G!	~7J&GP@@@@@@@@@@@@@@@@@
5!^B	@@@@@@@@@@&&@@@&BBG#^	~G@@@@@&@@@@GJY&&7#
Y7. G@	@@@@@@@@@@@@@&&J!!5Y:	!BP#@P?!?YPG@#Y??5@B~
7Y ^@@	@@@@@@@@@@B!^7:.	7B^^7!~!:P!Y###@@B
.P. ?@@	00000000000000000000000000000000000000	.?#B#@@@@P7::!:::J@@@
7J Y@@	@@@@@@@BY.	.5000000000000&&00&&PB00
P~ 75@	@@5J5J#P	.Y&@@@@@@@@@@@@@@@@@@&YG@
.G. ^@		7@@@@@@@@@@@@@@@@@@@@@?5
.G. Y	@#~?P^^~J?^!:.	J@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
.G.	^JG@G!:~J7^	7#@@@@@@@@@@@@@@@@@@@@@@@@@@@@
5~	.^Y@? ^~^	:5&@@@@&B@@@@@@@@@@@@@@@@
V J	~YY'?'@@@&&B!	:!~:~. ~7B@@@@@@@@@@@@@
.G.	. 7@@@@@@##B?	
	. B@@@@@@@@@@6777	: B@@@@@@@@@
1/ E1		
· 0! V7		
27	• ^.1&@@@@@@@@@@@@@@@@@@@@@@	00000000000000000000000000000000000000
		7@&6? · · ·
	7.1~ ^&@@@@@@@B`	· · · · · · · · · · · · · · · · · · ·
	7.1~	^?7
		:!7~.
	:!?7^. J@B.	.^!7!:
	:~77!~:. ^??^:	.:^!7!~:
	:^!!7!!!~~^:::::::	:^^~~!!!!~^:



Optimization!



• Circuitree [code]: A Datalog Reasoner in Zero-Knowledge

- Pros:
 - Close to SPARQL technology (existing N3 reasoning and SPARQL querying engines are built on top of)
- Cons:
 - Not tailored for proof of derived facts and likely has large proofs size; instead is defined for full ZKP of arbitrary prolog execution.
- Lean ZKP [code]: A zero knowledge theorem prover compatible with lean4:
 - Pros:
 - Specialised for proving that theorems are true in zero knowledge including SAT and other formal logic problems. This is a problem that can easily be translated into "is the following query true given this set of facts".
 - Has a relatively small proof size according to the paper
 - Cons:
 - Codebase appears to have been unmaintained for last 8 months
 - Unclear how easy it would be to translate this to a production system would a theorem prover be needed to make every implementation work?
 - Would likely need to collaborate directly with the original author to make this work.
- ZKSMT [code]: Similar to Lean ZKP, but specifically targeted towards SAT problems



100-1000x



Standardization!



Standardization

- Bind the proof to the SPARQL 1.2 algebra rather than to the Rust Risc-v implementation
- This could become a new proof method in <u>W3C</u> verifiable credentials, but ...





Abstraction!



User abstractions != Standards abstractions

• The and wallet for *work*, *personal* etc.; and the credential for *education*, *transport*, *health* should be user **abstractions** not what we are tied to in the specifications.



Shapes as a method for shaping credentials

• <u>Data shapes</u> languages such as <u>SHACL</u> enable the ability to query, frame and validate data; and thus generate a credential that conforms to an expected schema.



Call for collaboration

- Performance optimisation (rust)
- ZKP: Custom algo design with specialised circuits
- Modelling (rdf-star)
- Deployment in particular use-cases, including:
 - "the better versions" of examples from the Gamma trust framework
 - Classes of use cases which require the "integrate and derive" pattern
 - Perhaps too academic for some "declarative OIDC"





Call for use-cases





Future Work

- Emergent multi-party computation
 - User data stores declare "I permit my salary to be used to publicly declare the average salary of my workplace, but no one can know my individual salary"
- Policy aware query
 - See ODRL: <u>https://www.w3.org/community/reports/odrl/CG-FINAL-</u> profile-bp-20240808.html





My Recommended Reading List

- <u>https://www.w3.org/TR/sparql12-query/</u>
- <u>https://www.w3.org/community/reports/odrl/CG-FINAL-profile-bp-20240808.html</u>
- <u>https://solidproject.org</u>
- <u>https://www.w3.org/TR/prov-o/</u>

Questions

Email: jesse@jeswr.org Mastodon: jeswr@sfba.social



