

#### Analyze Decentralized Personal Health Data using Solid, Digital Consent, and Federated Learning

Citation for published version (APA):

Sun, C., van Soest, J., & Dumontier, M. (2023). Analyze Decentralized Personal Health Data using Solid, Digital Consent, and Federated Learning. In Semantic Web Applications and Tools for Health Care and Life Sciences 2023 (Vol. 3415, pp. 169-170) https://bbuelens.solidcommunity.net/public/html/SolidSymposium2023\_ChangSun.pdf

Document status and date: Published: 01/01/2023

**Document Version:** Publisher's PDF, also known as Version of record

**Document license:** Taverne

#### Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.

• The final author version and the galley proof are versions of the publication after peer review.

 The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these riahts.

Users may download and print one copy of any publication from the public portal for the purpose of private study or research.

You may not further distribute the material or use it for any profit-making activity or commercial gain
 You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

#### Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.





**Chang Sun**, Ph.D. Institute of Data Science, Maastricht University

Solid Symposium 2023 - Applications for Health Data Management with Solid 31 March 2023



















OJ L 119, 4.5.2016, p. 1-88 (BG, ES, CS, DA, DE, ET, EL, EN, FR, GA, HR, IT, LV, LT, HU, MT, NL, PL, PT, RO, SK, SL, FI, SV)

In force: This act has been changed. Current consolidated version: 04/05/2016

ELI: http://data.europa.eu/eli/reg/2016/679/oj

**≡** MENU

:



Institute of Data Science



A lot work focused on analysing the data controlled by organizations, but what about using data directly controlled by individuals?





#### Citizens would like to have more access and control over the data they generate.

Law and technology should EU should make major investments in enable us to make available their technologies and infrastructures that **enhance** data for the public interest data access and use, while giving us as well as without any direct reward. public and private organisations full control over the data they generate. We would be willing to make such data There are **no sufficient tools and** available, especially for health-related mechanisms for us to 'donate' research and for aspects relating to the locality our data. we live in (e.g. mobility, environment).

Maastricht University
 Summary report of the public consultation on the European strategy for data:
 Institute of Data Science
 https://ec.europa.eu/digital-single-market/en/news/summary-report-public-consultation-european-strategy-data

## A key challenge

*"Individuals value the high level of protection granted by the GDPR and ePrivacy legislation. However, they suffer from the absence of technical tools and standards that make the exercise of their rights simple and not overly burdensome."* 

-- European Committee <An European Strategy for Data> 2020





European Health Data Space: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space en

### Solutions:

provide individuals the tools to decide at a granular level what is done with their data using consent management, personal information management, decentralised networking, etc, in order to give greater oversight, transparency, control over their personal data.









**Based on Solid (SOcial Linked Data) to** store and structure personal data, and give individual control access to their fine-grained personal data

ciTlzen-cei	ntric DAta
pLatform	1 (TIDAL)
Welcome! https://chang.ini	rupt.net/profile/card#me
Github Get :	Started [WIP] →
Fetch data files from Solid Pods	WRITE DATA TO THE CREATED FILE

Create a file and write your data triples in					
Please give the location where you	want	to create/modify t	he file (pa	th+file name):	
https://user.podprovider.com/priva	t 🕇	newFile.ttl		Create	
WRITE DATA	тот	THE CREATED FI	LE		
ase give the data element URI (predi	cate) a	and the value (obje	ct), then o	click the (+) buttor	
ttps://schema.org/name		TIDAL		¢	
Add abo	ve tri	ples to my pod			
J					



**Based on Solid (SOcial LInked Data) to** store and structure personal data, and give individual control access to their fine-grained personal data



Q

G	DF	R	
-	~	×	
÷			е,
	-		×

Integrating Data Privacy Vocabulary to structure personal data requests as digital consents



@prefix : <https://exampleparticipant.solidprovider.com/profile/card#>.
@prefix SNOMEDCT: <http://purl.bioontology.org/ontology/SNOMEDCT/>.

:me a SNOMEDCT:116154003; # Patient SNOMEDCT:397669002 "27"^^xsd:int; # Age SNOMEDCT:50373000 "165"^^xsd:int; # Height SNOMEDCT:726527001 "55"^^xsd:int; # Weight SNOMEDCT:263495000 SNOMEDCT:248152002; # Gender, Female SNOMEDCT:271649006 "110"; # Systolic blood pressure SNOMEDCT:271650006 "90"; # Diastolic blood pressure SNOMEDCT:405751000 SNOMEDCT:44054006. # Type 2 diabetes

Listing 4: An example of the RDF data file in a participant's Solid pod.



**Formulating personal data and data request into RDF format** with integration of vocabulary services and standards





**Based on Solid (SOcial Linked Data) to** store and structure personal data, and give individual control access to their fine-grained personal data



**Integrating Data Privacy Vocabulary** to structure personal data requests as digital consents to meet the requirements of GDPR





Integrating federated learning (Personal Health Train) using the parameters promised in the data request and only the results are sent to the researchers



000000

### TIDAL framework - Data req



Research and Development	×			•	
escription of your purpose:					
Learn association between diabetes status and dietary pattern Rec					
ersonal data categories 🚯 🍍					
Medical Health [Special] (hys	sical Health, Mer	ntal Health, DNA	Code, Disability, Health Hi	story) 🗙 🚽	
Demographic (Physical Trait	, Income Bracket	, Geographic) 🗙	)		
ata elements (URI) 🕕 🍍					
Q diagnosis					
consent Duration (Days) *		ີ່ Searchi	ng terms from BioPo	ortal ontologie	
90	NCIT	Diagnosis The investigation, and condition, or injury fro of any kind; the concis	http://ncicb.nci.nih.gov/xml/ow alysis and recognition of the presence om expressed signs and symptoms; a se results of such an investigation.	I/EVS/Thesaurus.owI#C: e and nature of disease, lso, the scientific determin	
100	PREMEDONTO	Diagnosis The investigation, ana condition, or injury fro of any kind; the concis	http://purl.obo alysis and recognition of the presence om expressed signs and symptoms; a se results of such an investigation.	library.org/obo/NCIT_C e and nature of disease, lso, the scientific determin	
<ul> <li>Analyse ×</li> </ul>	CRISP	diagnosis general term for deter	http://purl.bioontology cting and classifying diseases.	.org/ontology/CSP/4000	
nalysis Model *	IOBC	Diagnosis	http://purl.jp/bio	o/4/id/20090600161154	
Linear Regression				•	
onsequences of data processing	and impact of you	ir research:			

#### TIDAL framework - Data request and consent

```
@prefix : <http://exampleresearcher.solidprovider.com/public/request.ttl#>.
@prefix schema: <https://schema.org/>.
@prefix exre: <http://exampleresearcher.solidprovider.com/profile/card#>.
@prefix dpv: <http://w3id.org/dpv#>.
@prefix dpvpd: <http://w3id.org/dpv/dpv-pd#>.
@prefix dpvtech: <http://w3id.org/dpv/dpv-tech#>.
@prefix SNOMEDCT: <http://purl.bioontology.org/ontology/SNOMEDCT/>.
```

```
:161964062096710764675982245664
```

```
a schema:AskAction, dpv:PersonalDataHandling;
dpv:hasLegalBasis dpv:Consent;
rdfs:label "Learn association between diabetes status and dietary pattern";
schema:collectionSize 32;
schema:creator exre:me;
schema:DataFeedItem SNOMEDCT:10396001, SNOMEDCT:230125005, SNOMEDCT:56718006, SNOMEDCT:73211009;
schema:dateCreated "2021-01-18T00:00:00Z"^^XML:dateTime;
dpvtech:isImplementedUsingTechnology dpvtech:LinearRegression;
dpv:hasImpact "Help diabetes patients understand the impact of their diet pattern";
dpv:hasContext SNOMEDCT;
dpv:hasDataController exre:me;
dpv:hasDataController exre:me;
dpv:hasPersonalData dpvpd:Health;
dpv:hasPersonalData dpvpd:Health;
dpv:hasPurpose dpv:ResearchAndDevelopment.
```

## TIDAL framework - Data request and consent



## TIDAL framework - Data request and consent



15

Listing 2: An example of generated participation statements in a RDF format in the participant's Solid pod.

#### **TIDAL framework - Data Analysis**





#### **TIDAL framework - Data Analysis**



@prefix : <http://examplertrustedparty.solidprovider.net/inbox/triggermessage#>. @prefix req: <http://exampleresearcher.solidprovider.com/public/request.ttl#>. @prefix exre: <http://exampleresearcher.solidprovider.com/profile/card#>.

```
:160622932739325095672093710975
schema:actionStatus schema:ActivateAction;
schema:creator exre:me;
schema:dateCreated "2021-04-20T09:37:57.4992"^^XML:dateTime;
schema:target req:161964062096710764675982245664.
```

Listing 3: An example of generated trigger message (Activate Action) sent by the researcher.

#### **TIDAL framework - Data Analysis**

1.Send a

Request

JS

Researcher

Trigger

Analysis



#### Discussion

- Individuals data will be read only when the analysis gets executed.
- The analysis can only be triggered by the data requester under conditions (have enough participants, in the valid period, for the indicated purpose, etc)
- Individuals can revoke the permission anytime.
- Individuals can change the data values at any moment (even after giving researchers permission, when the experiments are being conducted)
- (Future work) Individuals will be able to indicate their preference to which data requests they want to receive.
- and more...

Paper: ciTlzen-centric DAta pLatform (TIDAL): Sharing Distributed Personal Data in a

Privacy-Preserving Manner for Health Research

**Open Source tool and code:** <u>https://github.com/sunchang0124/TIDAL</u>

Web application is live  $\rightarrow$ 

Demo Youtube Video:



Contact:

- chang.sun@maastrichtuniversity.nl
- www.linkedin.com/in/chang-sun-maastricht/



IN SCIENCE & TECHNOLOG

