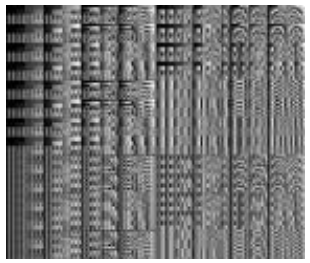
Open Science with Closed Data

Allan Hanbury

TU WIEN Informatics

“Classic” Data Science



Data

Analysis

Data Repository

Code Repository

Publication

https://creazilla.com/nodes/54725-page-facing-up-emoji-clipart

https://www.flickr.com/photos/kylemcdonald/6187343093/

https://www.needpix.com/photo/1331454/

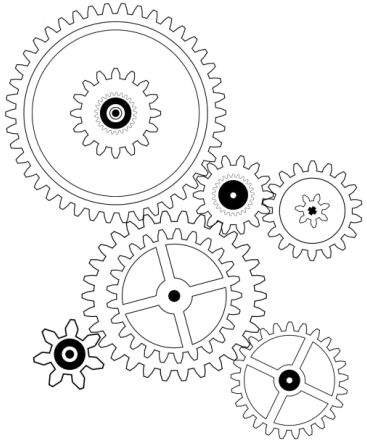
Closed Data

https://pixabay.com/vectors/padlock-silver-lock-security-147915/

Huge

Real-time

Non-distributable



Data Repository

Code Repository

Publication

https://creazilla.com/nodes/54725-page-facing-up-emoji-clipart

https://www.flickr.com/photos/kylemcdonald/6187343093/

https://www.needpix.com/photo/1331454/

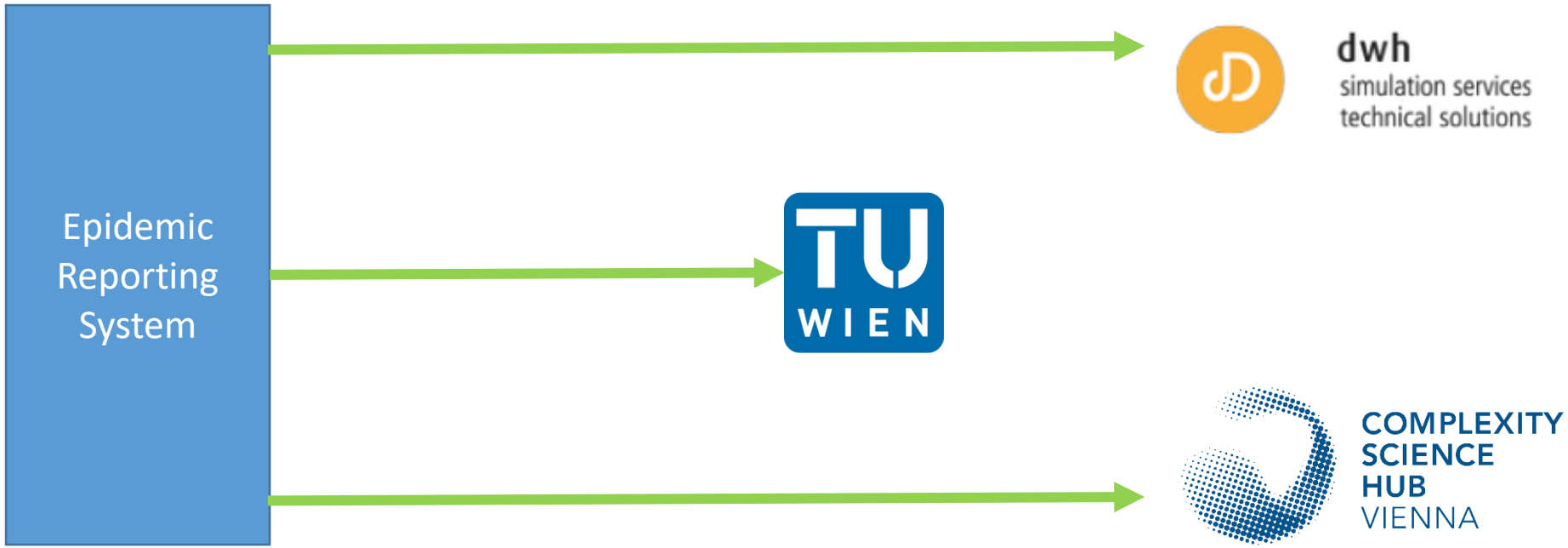
Example:
COVID-19 Analysis Pipeline

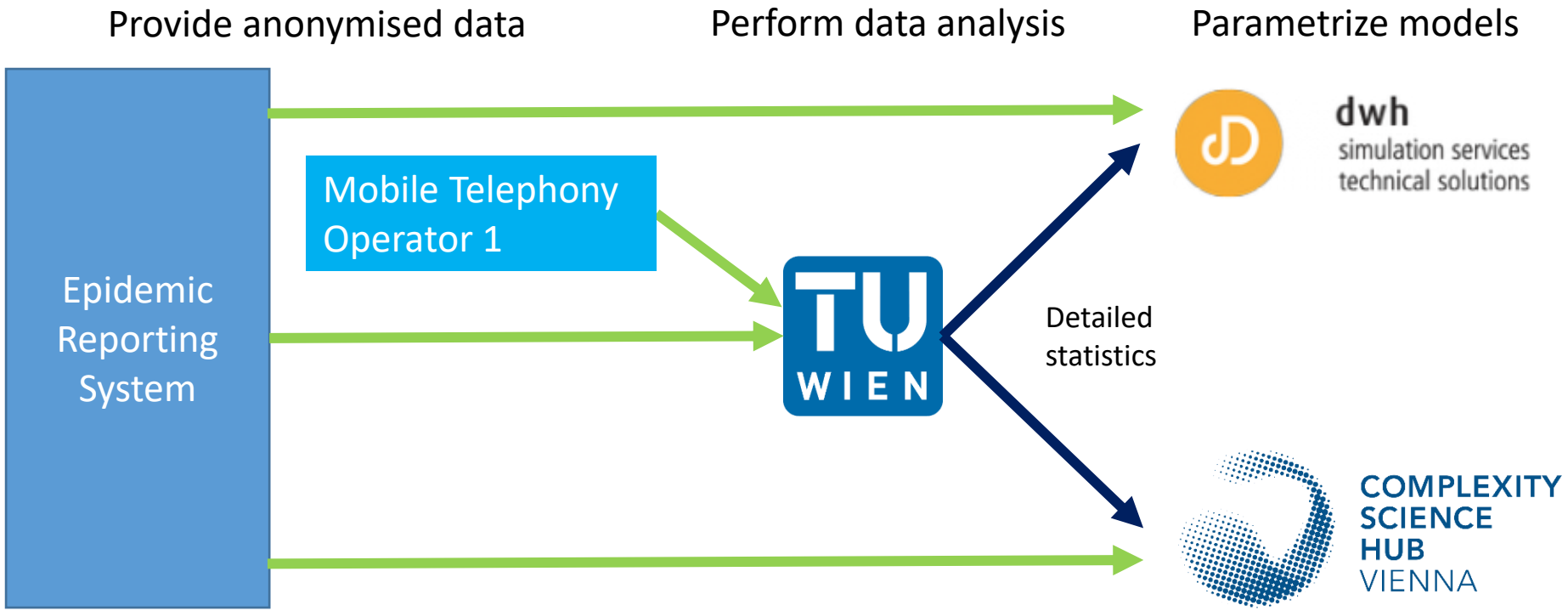
dwh
simulation services
technical solutions

COMPLEXITY
SCIENCE
HUB
VIENNA

W|W|T|F
WIENER WISSENSCHAFTS-,
FORSCHUNGS- UND TECHNOLOGIEFONDS

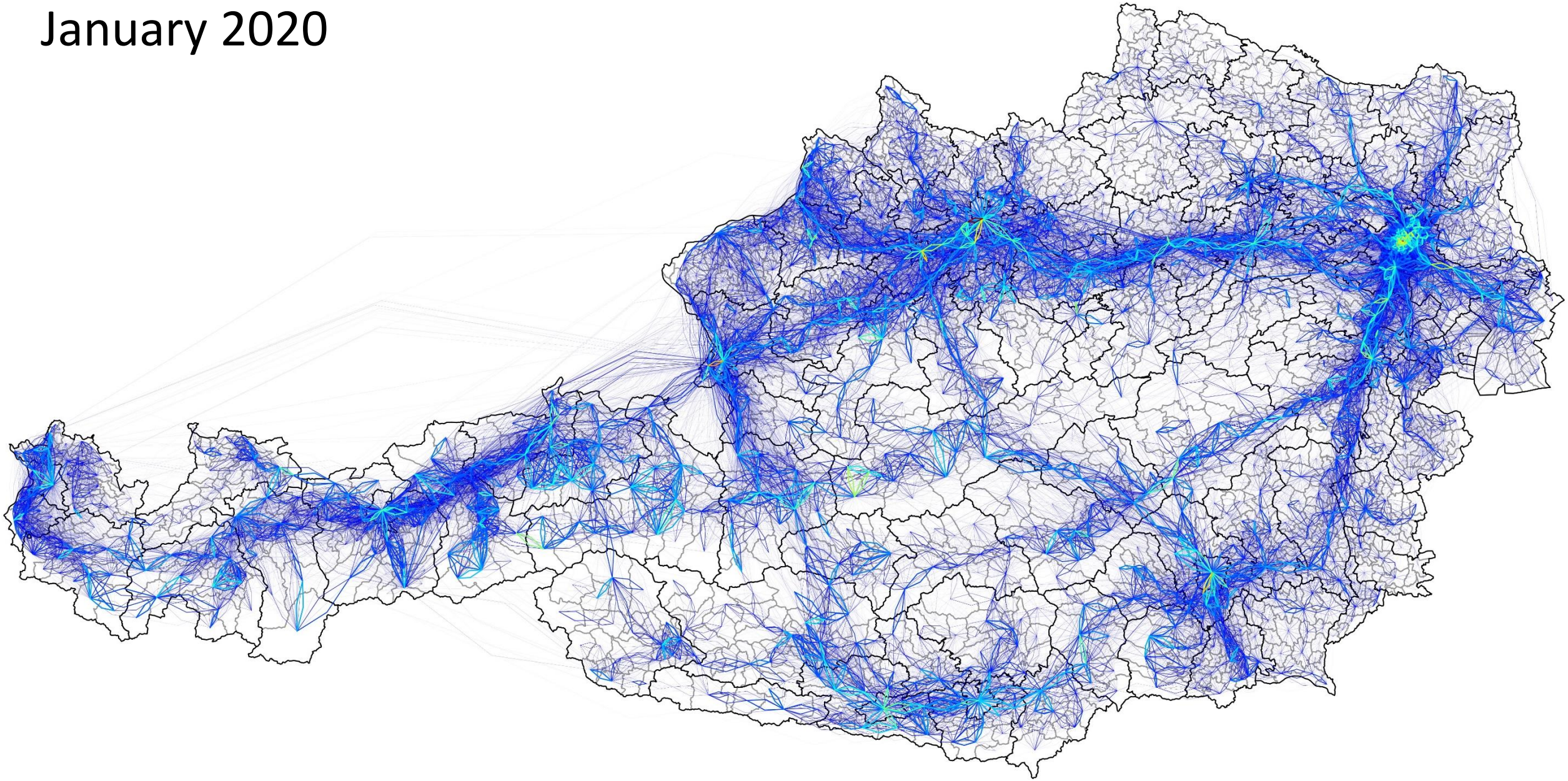
Provide anonymised data



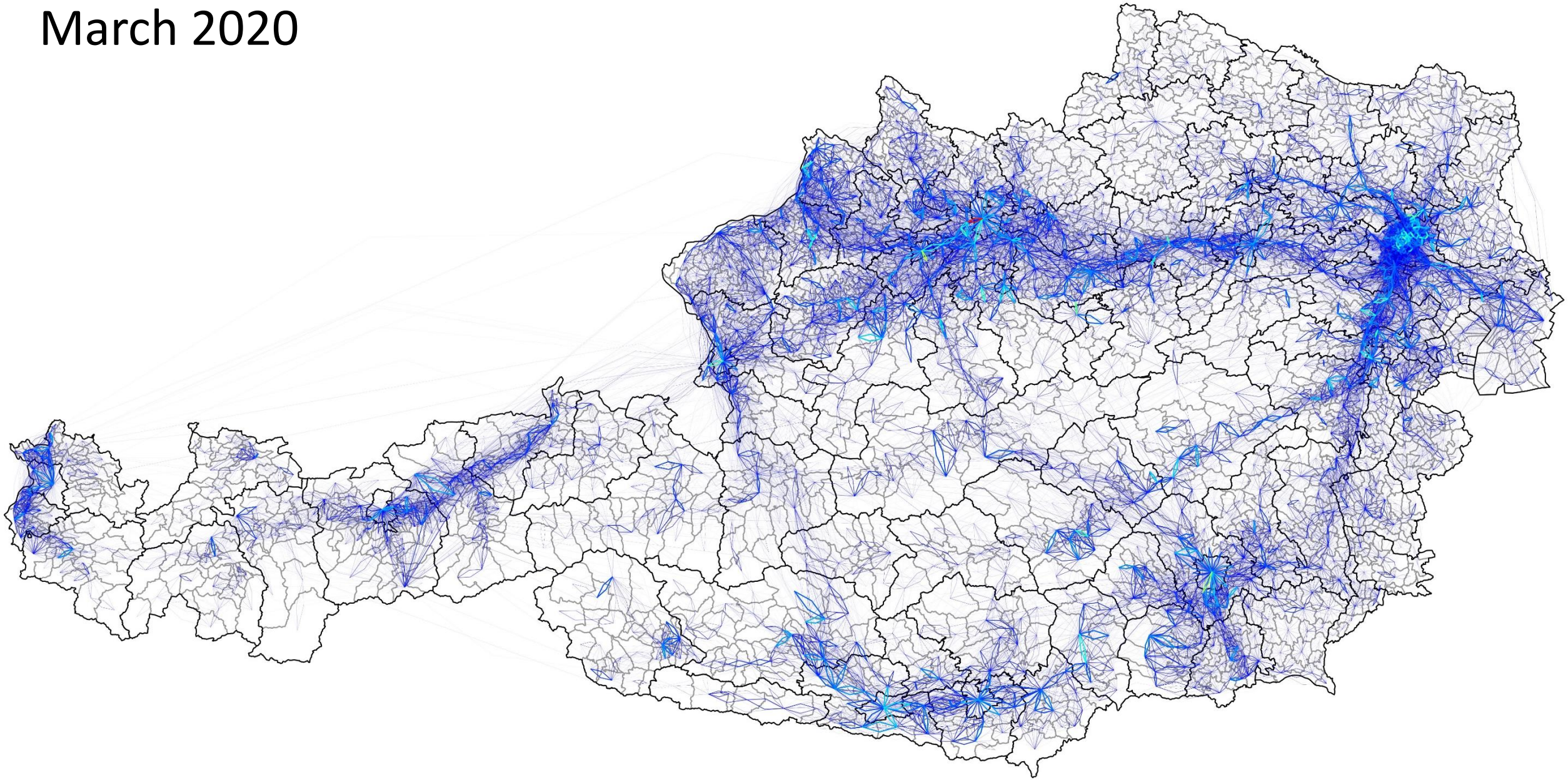


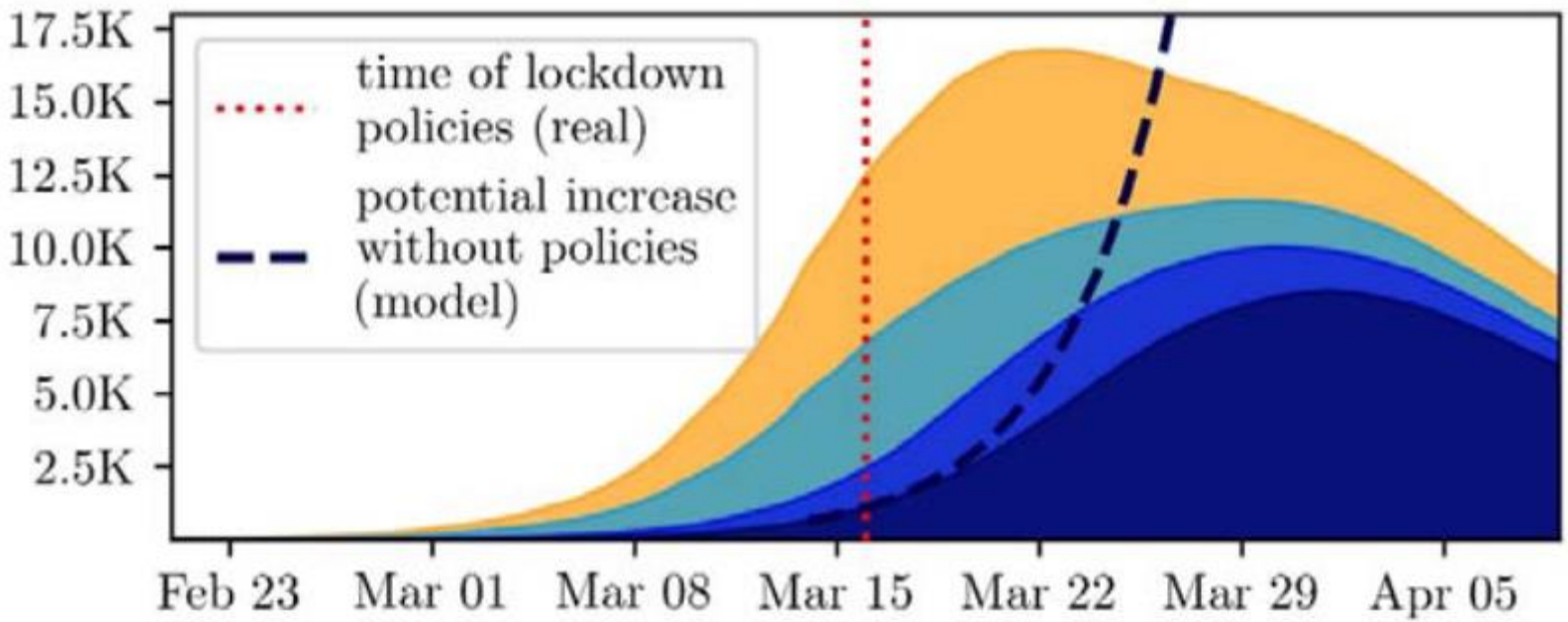
WWTF
WIENER WISSENSCHAFTS-, FORSCHUNGS- UND TECHNOLOGIEFONDS

January 2020



March 2020



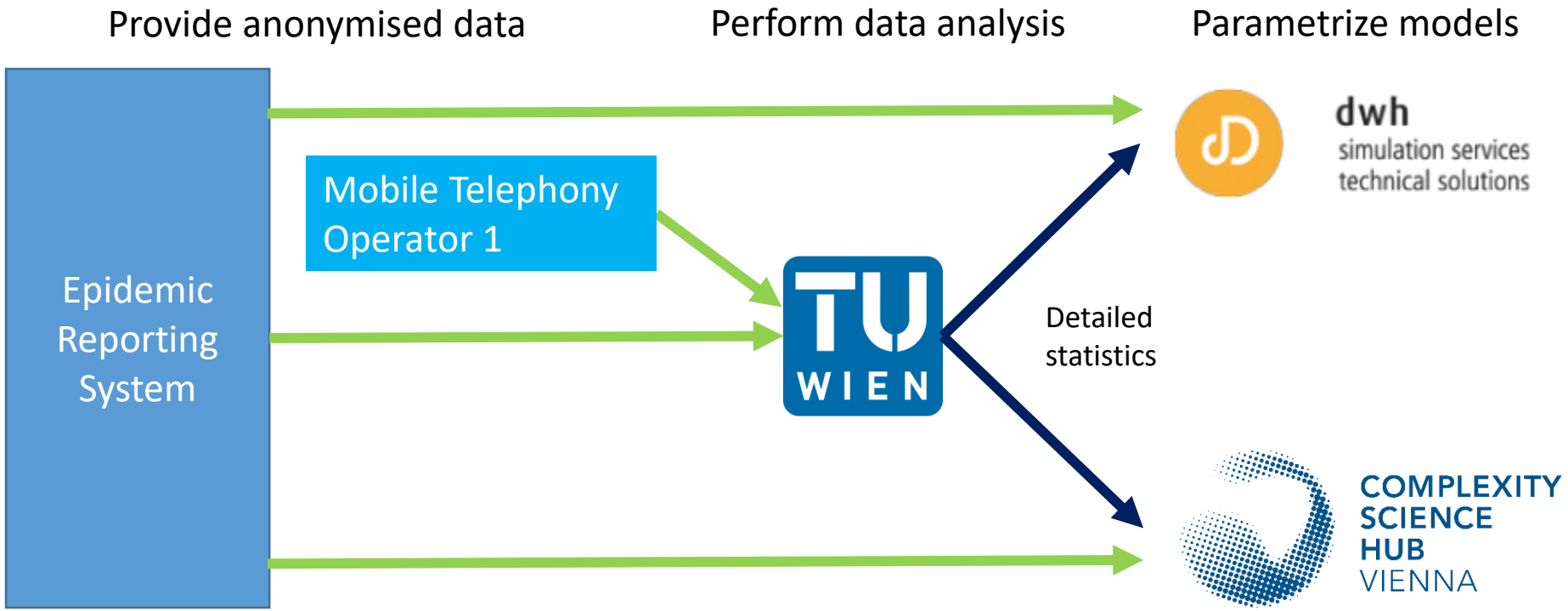


confirmed cases (model)

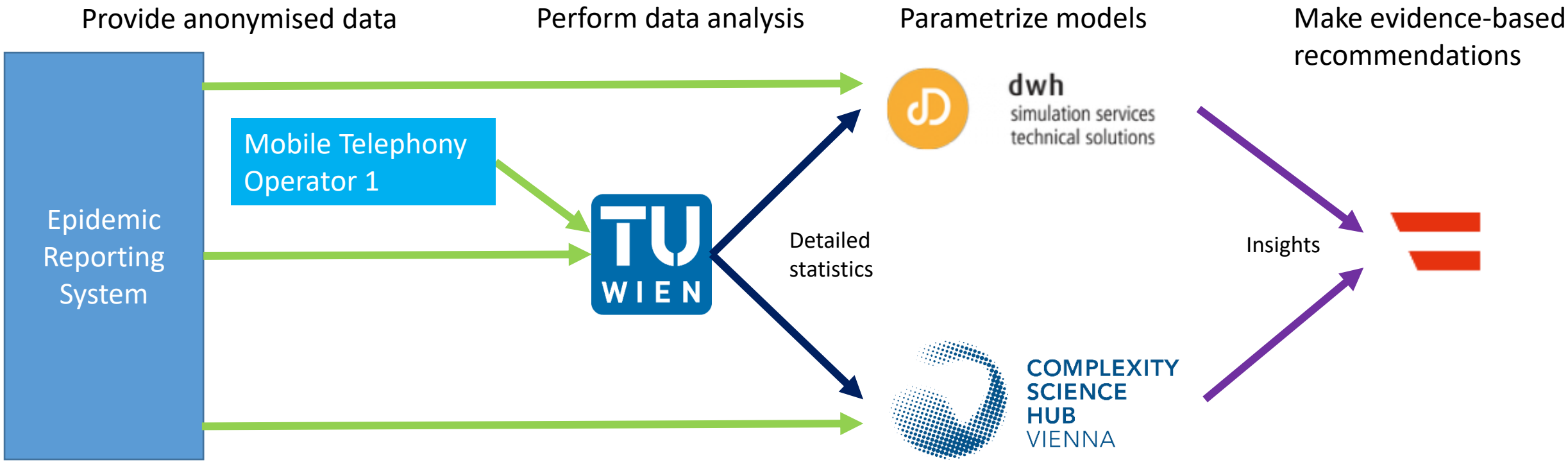
unconfirmed symptomatic cases (model)

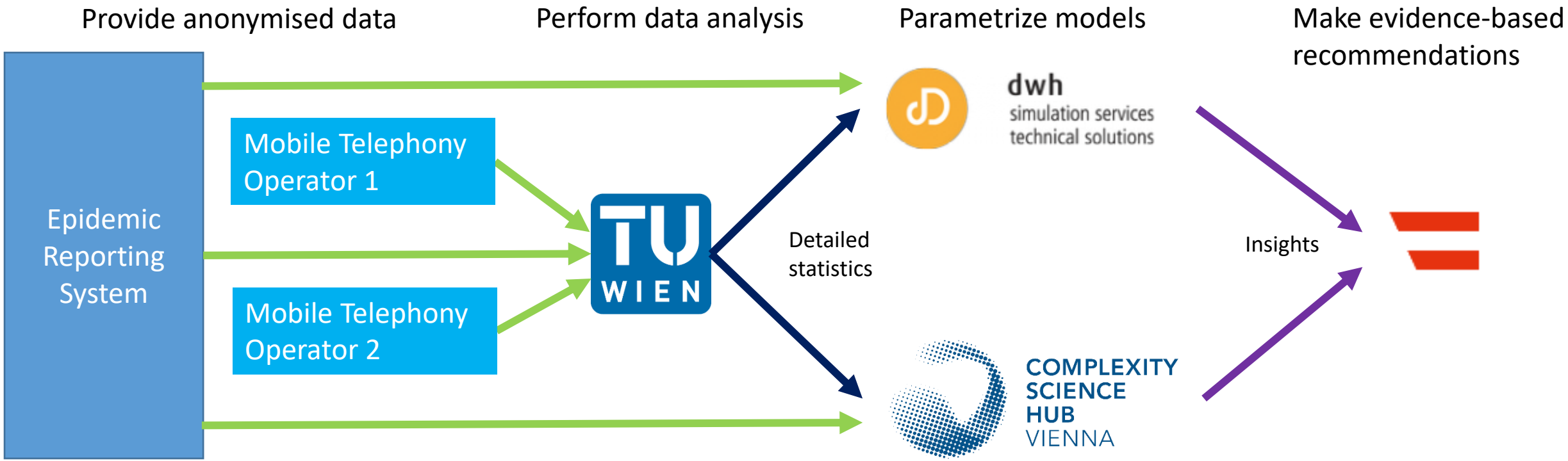
pre-symptomatic cases (model)

asymptomatic cases (model)

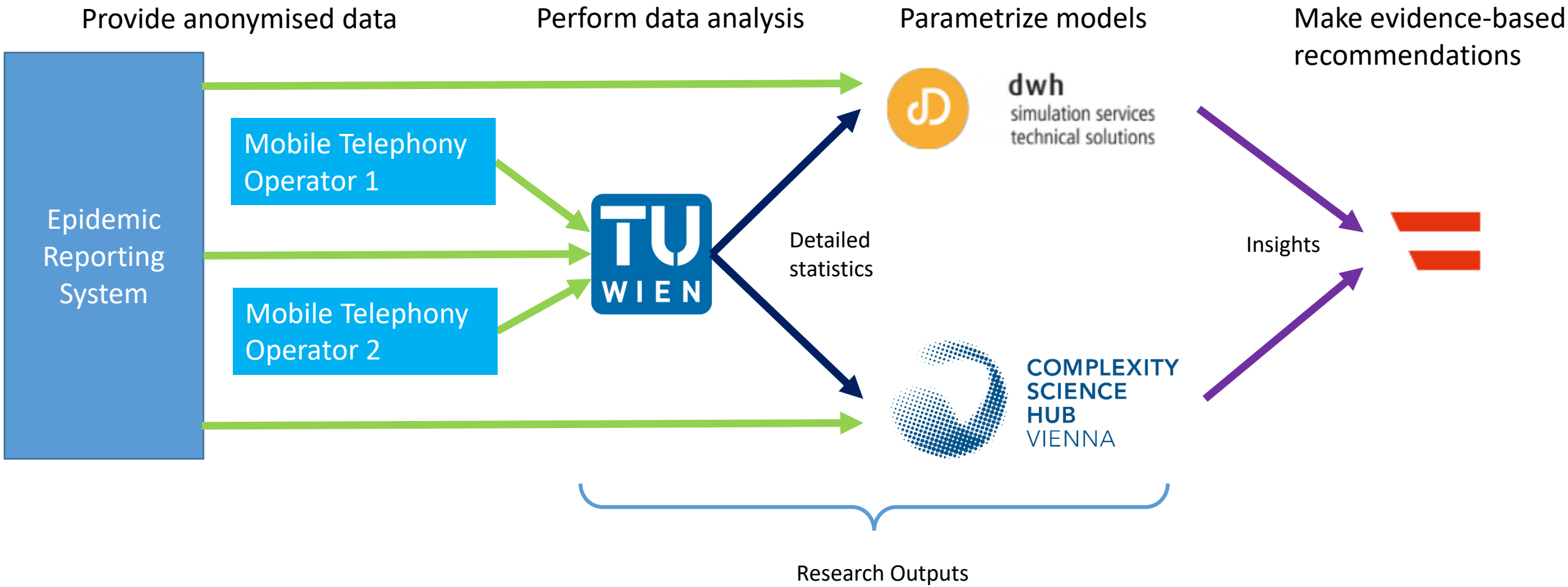


WWTF
WIENER WISSENSCHAFTS-, FORSCHUNGS- UND TECHNOLOGIEFONDS


Provide anonymised data
Perform data analysis
Parametrize models
Make evidence-based recommendations
Epidemic Reporting System
Mobile Telephony Operator 1
TU WIEN
Detailed statistics
dwh
simulation services
technical solutions
COMPLEXITY SCIENCE HUB VIENNA
Insights



WWTF
WIENER WISSENSCHAFTS-, FORSCHUNGS- UND TECHNOLOGIEFONDS


Provide anonymised data
Perform data analysis
Parametrize models
Make evidence-based recommendations
Epidemic Reporting System
Mobile Telephony Operator 1
Mobile Telephony Operator 2
TU WIEN
dwh
simulation services
technical solutions
Detailed statistics
Insights
COMPLEXITY SCIENCE HUB VIENNA
Research Outputs

Provide

dence-based
ndations

Epidemic Reporting System

The impact of COVID-19 on relative changes in aggregated mobility using mobile-phone data

Georg Heiler[a,b], Allan Hanbury[a,b] and Peter Filzmoser[c]

[a]Institute of Information Systems Engineering, TU Wien, Favoritenstr. 9-11, 1040 Vienna, Austria; [b]Complexity Science Hub, Josefstdter Str. 39, 1080 Vienna, Austria; [c] Computational Statistics Institute of Statistics and Mathematical Methods in Economics, TU Wien, Wiedner Hauptstrasse 8-10 1040 Vienna

ARTICLE HISTORY
Compiled September 9, 2020

ABSTRACT
Evaluating relative changes leads to additional insights that would remain hidden when only evaluating absolute changes. We analyze a dataset describing the mobility of mobile phones in Austria before, during COVID-19 lock-down measures until recently.

By applying compositional data analysis we show that formerly hidden information becomes available: we see that the elderly population groups increase relative mobility and that the younger groups, especially on weekends, also do not decrease their mobility as much as the others.

KEYWORDS
compositional-data-analysis, mobility, pandemic, big-data, geospatial-data

Provide a

e evidence-based

mmendations

Epidemic Reporting System

Supporting Austria through the COVID-19 Epidemics with a Forecast-Based Early Warning System

Martin Bicher[1,2,†], Martin Zuba[3,†], Lukas Rainer[3], Florian Bachner[3], Claire Rippinger[2], Herwig Ostermann[3,4], Nikolas Popper[1,2,5], Stefan Thurner[6,7,8,9], Peter Klimek[6,7*]

[1] *Institute of Information Systems Engineering, TU Wien, Favoritenstraße 8-11, A-1040 Vienna, Austria*

[2] *dwh simulation services, dwh GmbH, Neustiftgasse 57-59, A-1070 Vienna, Austria*

[3] *Austrian National Public Health Institute, Stubenring 6, A-1010 Vienna, Austria*

[4] *Private University for Health Sciences, Medical Informatics and Technology GmbH, UMIT, Eduard-Wallnöfer-Zentrum 1, A-6060 Hall in Tirol, Austria*

[5] *Society for Decision Support Policy and Planning, DEXHELPP, Neustiftgasse 57-59, A-1070 Vienna, Austria*

[6] *Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria*

[7] *Complexity Science Hub Vienna, Josefstädterstraße 39, A-1080 Vienna, Austria*

[8] *International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria*

[9] *Santa Fe Institute, 1399 Hyde Park road, Santa Fe, NM 87501, USA*

[†] *equal contributions*

(Dated: July 2020)

Background: The corona crisis hit Austria at the end of February 2020 with one of the first European superspreading events. In response, the governmental crisis unit commissioned a forecast consortium with regularly projections of case numbers and demand for hospital beds.

Methods: We consolidated the output of three independent epidemiological models (ranging from agent-based micro simulation to parsimonious compartmental models) and published weekly short-term forecasts for the number of confirmed cases as well as estimates and upper bounds for

mobility and that the younger groups, especially on weekends, also do not decrease their mobility as much as the others.

KEYWORDS

WIENER WISSENSCHAFTS-, FORSCHUNGS- UND TECHNOLOGIEFONDS

Provide anonymised data

Perform data analysis

Parametrize models

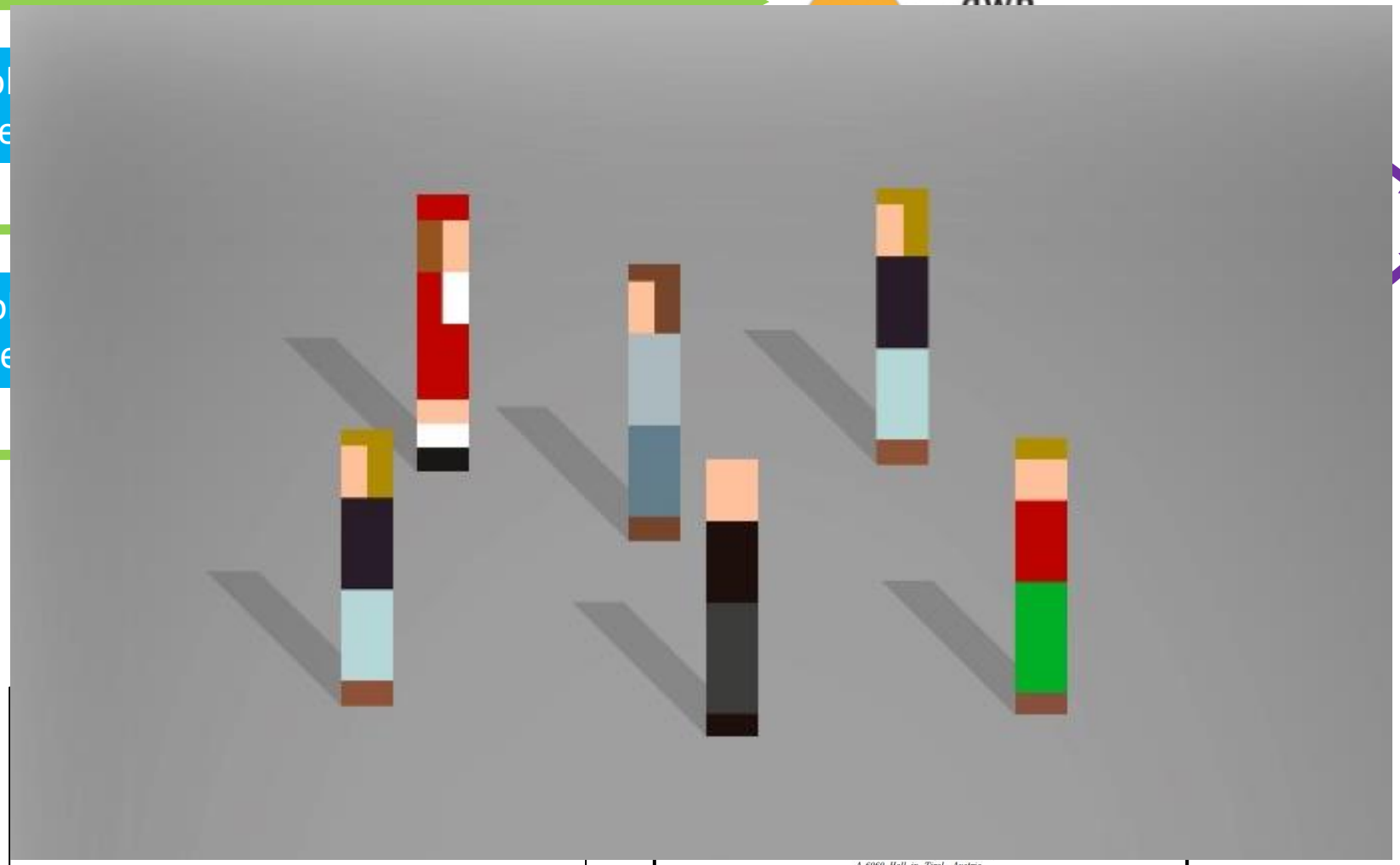
Make evidence-based recommendations

Epidemic Reporting System

Mo
Ope

Mo
Ope

dwh



=

ARTICLE HISTORY
Compiled September 9, 2020

ABSTRACT
Evaluating relative changes leads to additional insights that would remain hidden when only evaluating absolute changes. We analyze a dataset describing the mobility of mobile phones in Austria before, during COVID-19 lock-down measures until recently.

By applying compositional data analysis we show that formerly hidden information becomes available: we see that the elderly population groups increase relative mobility and that the younger groups, especially on weekends, also do not decrease their mobility as much as the others.

KEYWORDS

A-6060 Hall in Tirol, Austria
[5]Society for Decision Support Policy and Planning, DEXHELPP, Neustiftgasse 57-59, A-1070 Vienna, Austria
[6]Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria
[7]Complexity Science Hub Vienna, Josefstädterstraße 39, A-1080 Vienna, Austria
[8]International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria
[9]Santa Fe Institute, 1399 Hyde Park road, Santa Fe, NM 87501, USA
[1]equal contributions

(Dated: July 2020)

Background: The corona crisis hit Austria at the end of February 2020 with one of the first European superspreading events. In response, the governmental crisis unit commissioned a forecast consortium with regularly projections of case numbers and demand for hospital beds.

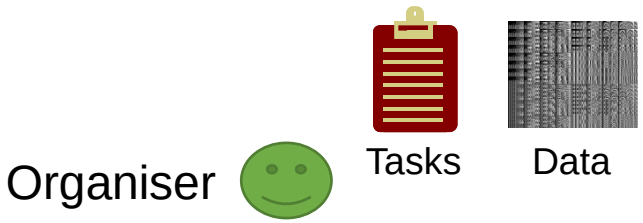
Methods: We consolidated the output of three independent epidemiological models (ranging

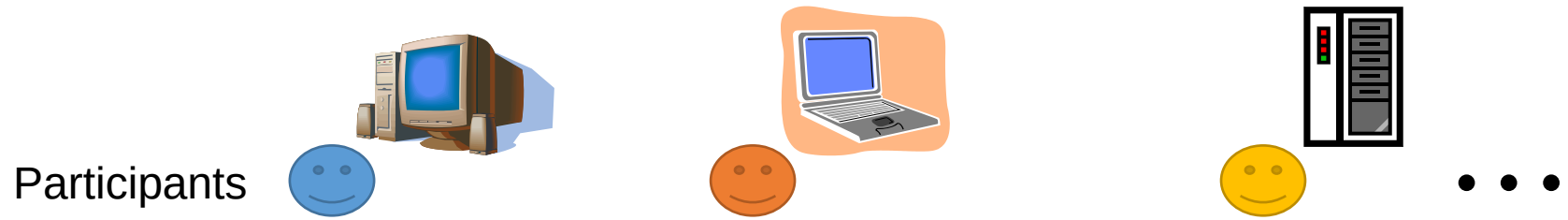
WWTF
WIENER WISSENSCHAFTS-, FORSCHUNGS- UND TECHNOLOGIEFONDS

https://www.piqsels.com/en/public-domain-photo-zbpbr

Evaluation-as-a-Service

Evaluation Campaigns / Shared Tasks / Challenges / Competitions / …



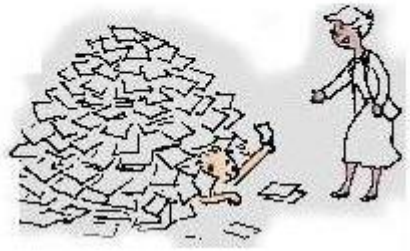


Kyle Mcdonald: http://www.flickr.com/photos/kylemcdonald/6187343093/

Academic Evaluation Campaigns

Text REtrieval Conference (TREC)

...to encourage research in information retrieval from large text collections.



NTCIR

CLEF INITIATIVE

The CLEF Initiative

Conference and Labs of the Evaluation Forum

MediaEval Benchmark

PASCAL2

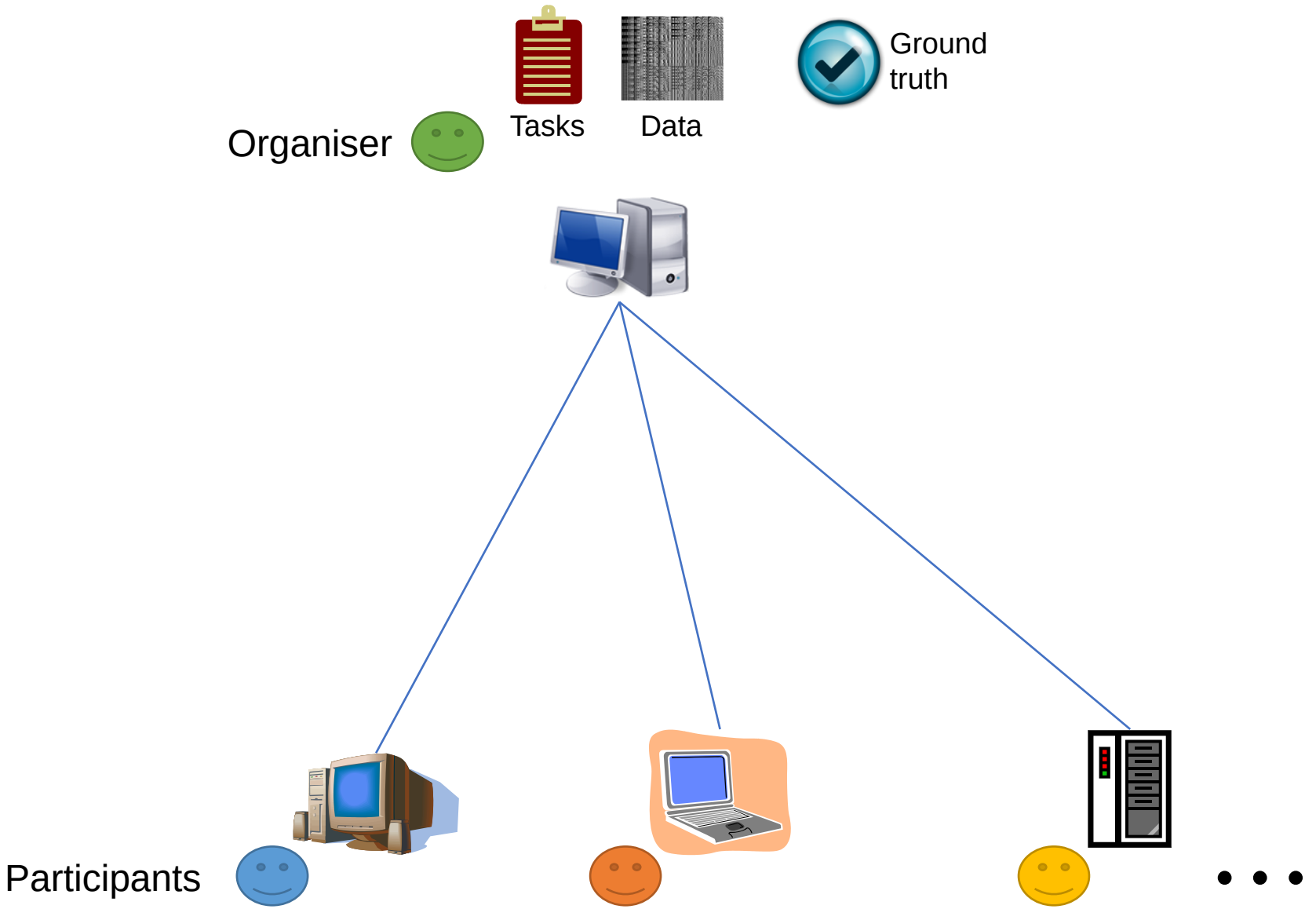
Pattern Analysis, Statistical Modelling and Computational Learning

BioASQ

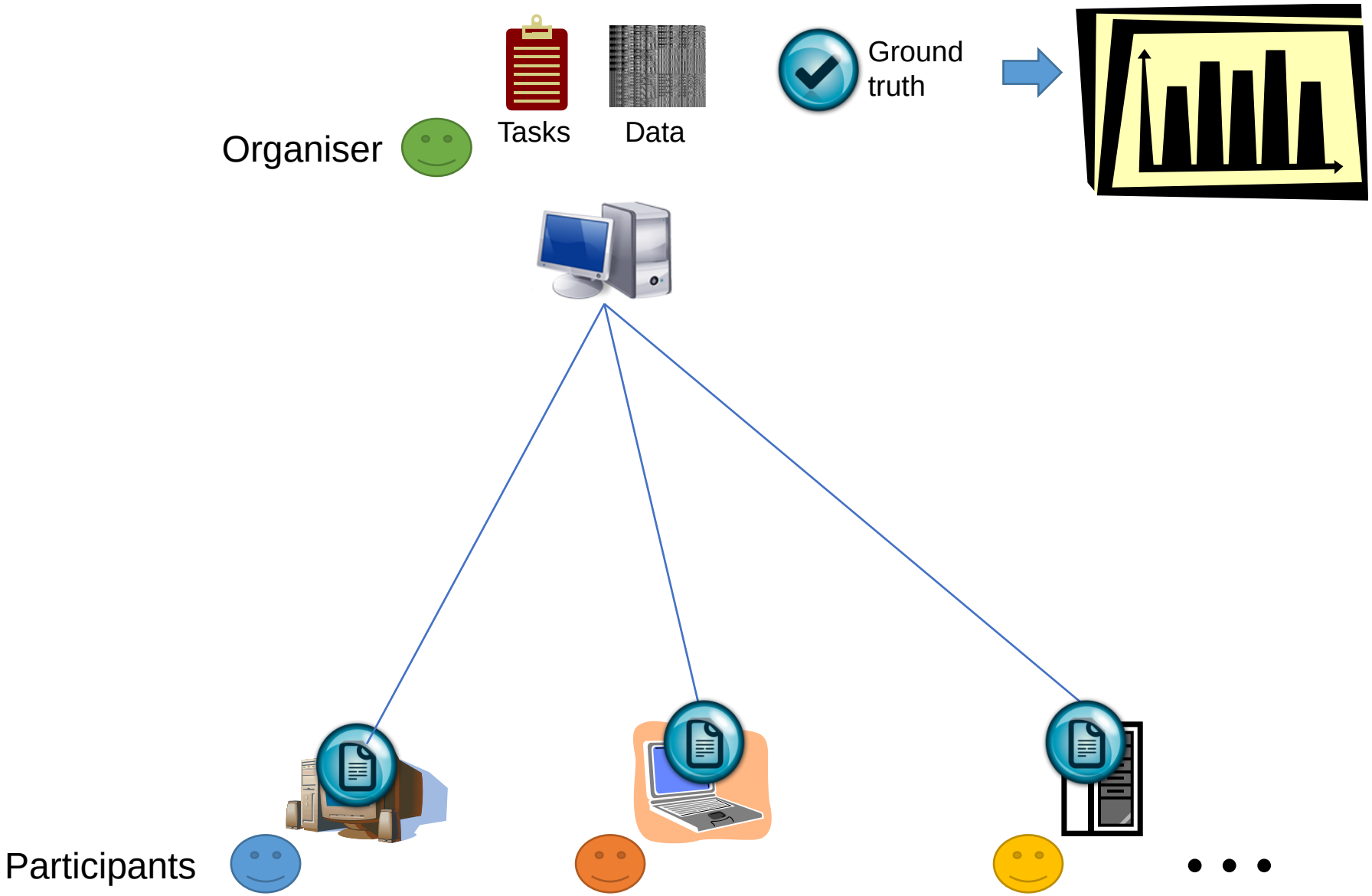
INEX

Initiative for the Evaluation of XML retrieval

TAC



Kyle Mcdonald: http://www.flickr.com/photos/kylemcdonald/6187343093/



Kyle Mcdonald: http://www.flickr.com/photos/kylemcdonald/6187343093/

Tasks

Data

Ground truth

Data Repository

Code Repository

Publication

Tasks Data

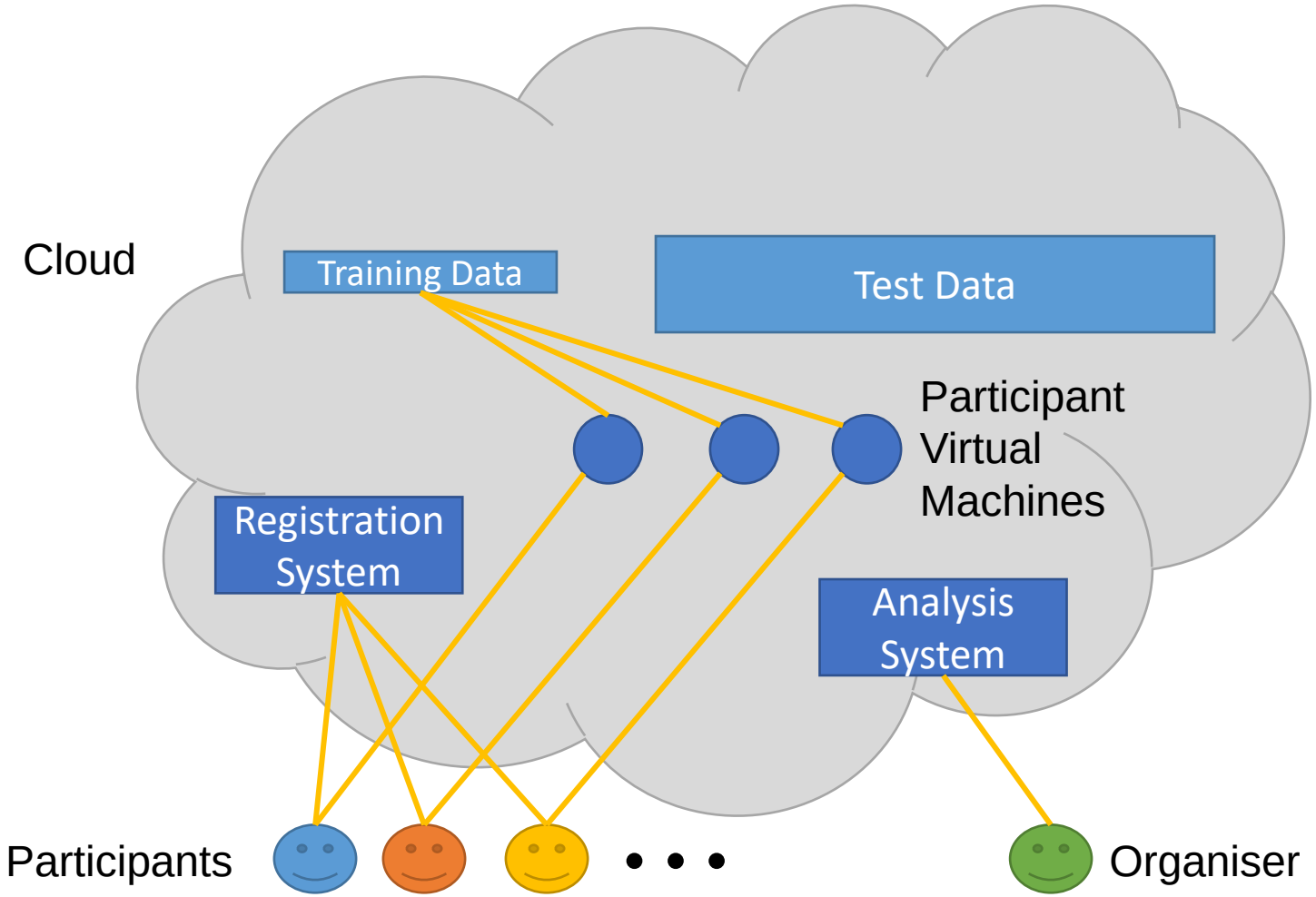
Ground truth

Kyle Mcdonald: http://www.flickr.com/photos/kylemcdonald/6187343093/

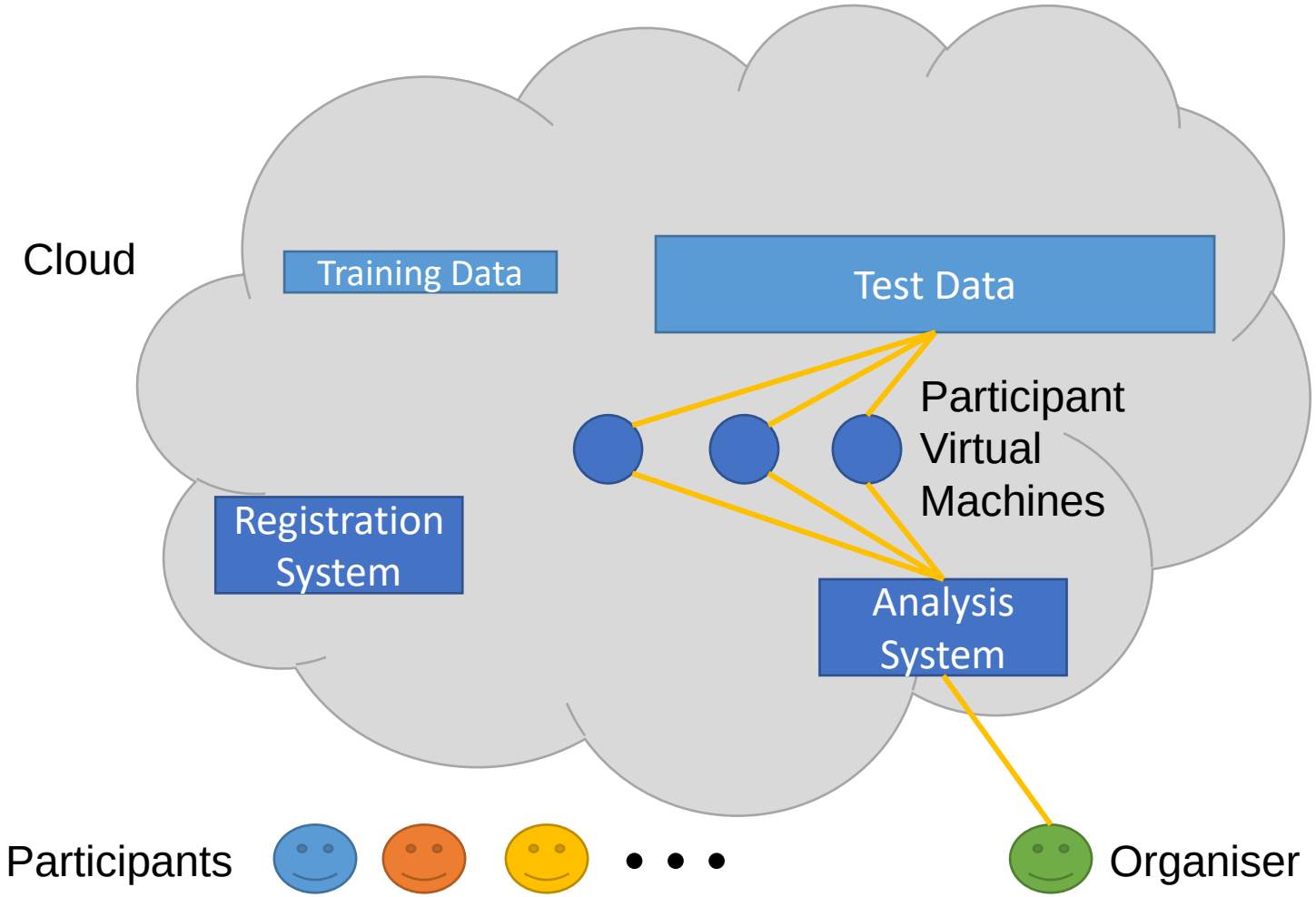
VISCERAL
Anatomy Benchmarks

Whole body labelling in 3D
medical imaging data

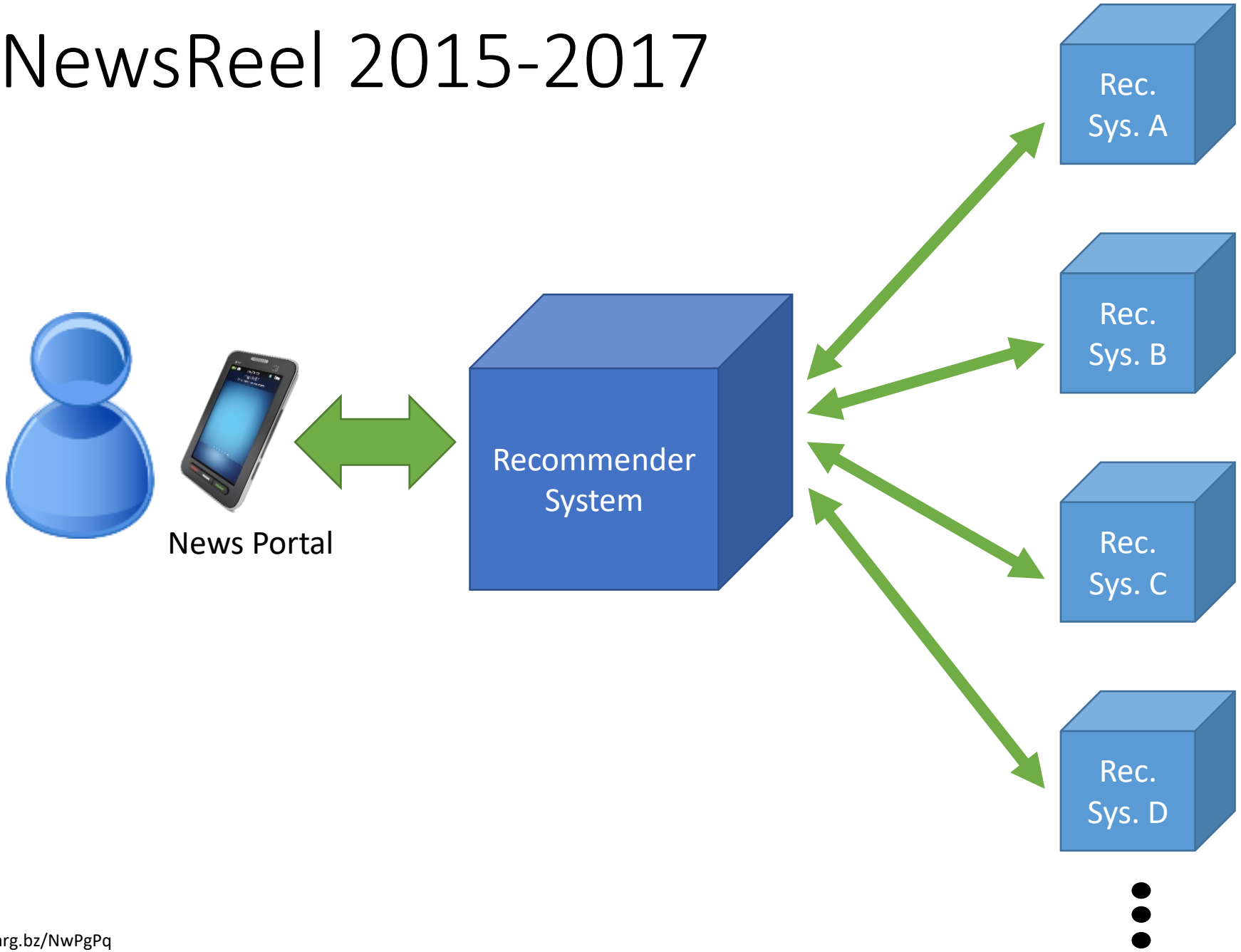
Training Phase



Evaluation Phase



CLEF NewsReel 2015-2017



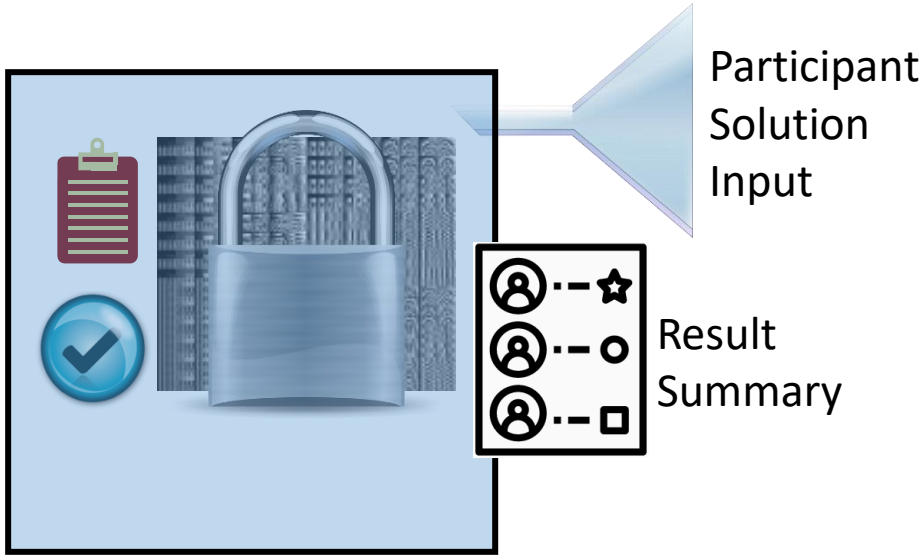
http://mrg.bz/NwPgPq

Evaluation-as-a-Service

Bringing the Algorithms to the Data

Evaluation-as-a-Service Stakeholders

Data/Task Provider



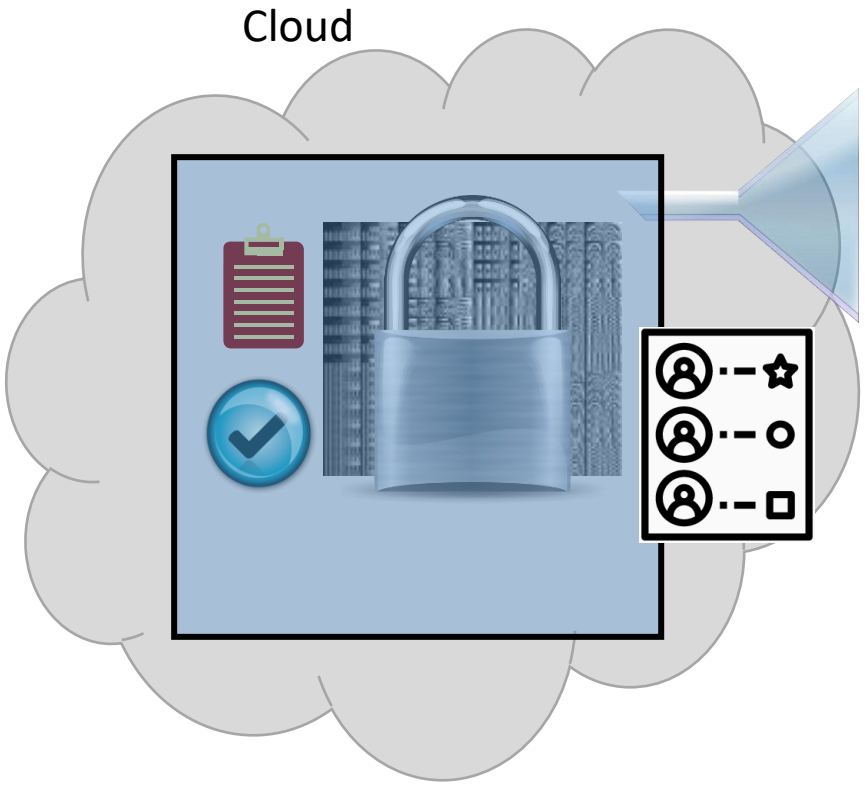
Stakeholders

Data/Task Provider

Organiser

https://thenounproject.com/term/leaderboard/2484528/

https://www.needpix.com/photo/download/29862/funnel-blue-flow-cone-liquid-fluid-tunnel-equipment-pour



Stakeholders

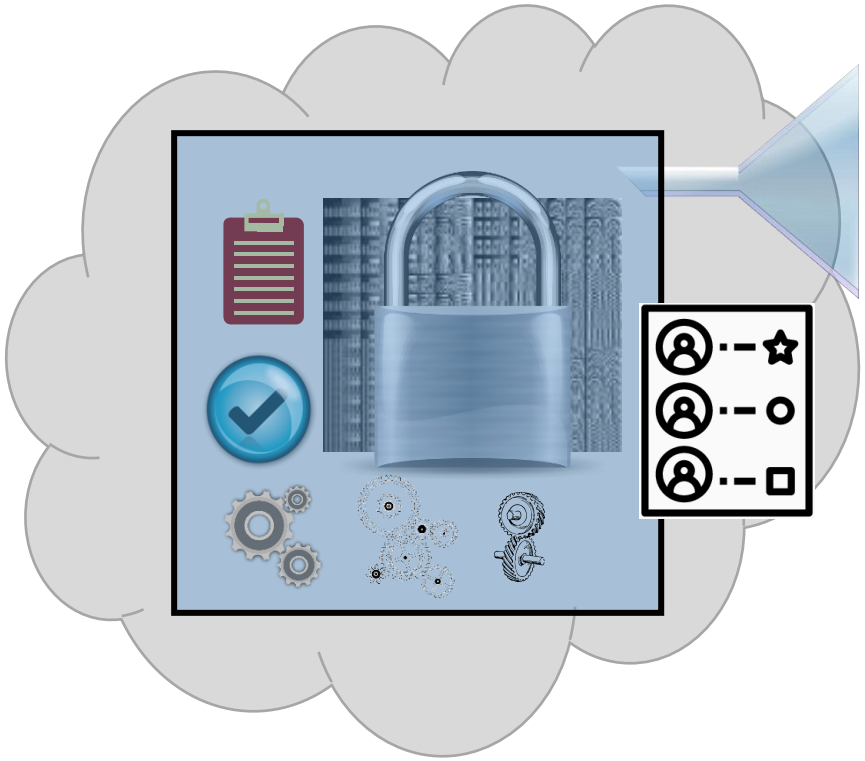
Data/Task Provider

Organiser

Infrastructure Provider

https://thenounproject.com/term/leaderboard/2484528/

https://www.needpix.com/photo/download/29862/funnel-blue-flow-cone-liquid-fluid-tunnel-equipment-pour



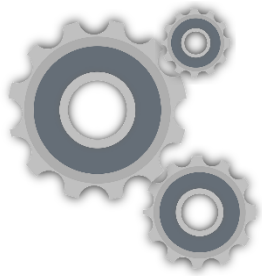
Stakeholders

Data/Task Provider

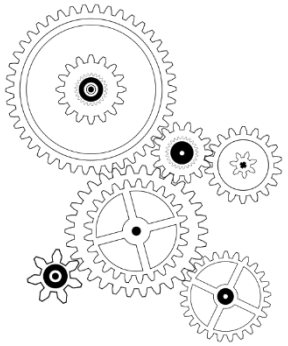
Organiser

Infrastructure Provider

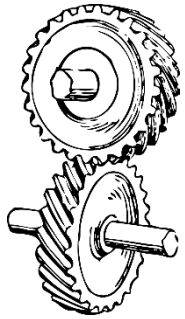
Participants



Participant 1 solution

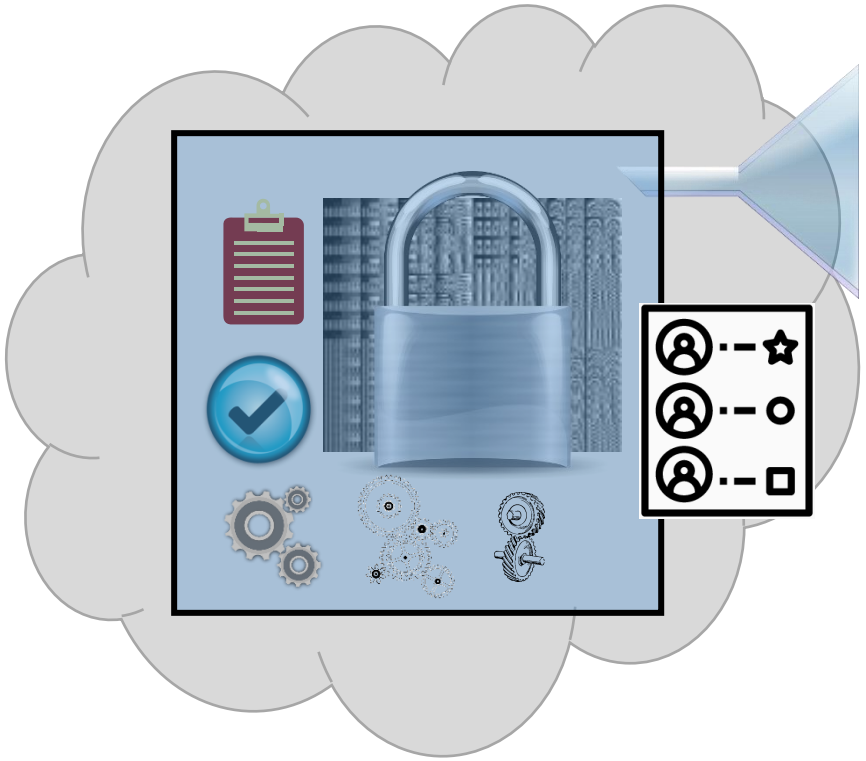


Participant 2 solution



Participant 3 solution

https://pixabay.com/vectors/gear-cog-wheel-tools-rack-wheel-307780/
https://pixabay.com/vectors/gears-wheel-rotate-mechanical-37306/
https://thenounproject.com/term/leaderboard/2484528/
https://www.needpix.com/photo/download/29862/funnel-blue-flow-cone-liquid-fluid-tunnel-equipment-pour



Sustainability?

Data Repository

Code Repository

Publication

Evaluation-as-a-Service for the Computational Sciences: Overview and Outlook

FRANK HOPFGARTNER, University of Sheffield
ALLAN HANBURY, TU Wien, Complexity Science Hub Vienna
HENNING MÜLLER and IVAN EGGEL, University of Applied Sciences Western Switzerland (HES-SO)
KRISZTIAN BALOG, University of Stavanger
TORBEN BRODT, plista GmbH
GORDON V. CORMACK and JIMMY LIN, University of Waterloo
JAYASHREE KALPATHY-CRAMER, Athinoula A. Martinos Center for Biomedical Imaging at Massachusetts General Hospital and Harvard Medical School
NORIKO KANDO, National Institute of Informatics
MAKOTO P. KATO, Kyoto University
ANASTASIA KRITHARA, National Center for Scientific Research "Demokritos"
TIM GOLLUB, Bauhaus-Universität Weimar
MARTIN POTTHAST, Leipzig University
EVELYNE VIEGAS, Microsoft Research
SIMON MERCER, Independent Consultant

Journal of Data and Information Quality, October 2018, Article No.: 15, https://doi.org/10.1145/3239570

European Data Strategy

Shaping Europe's digital future

POLICY

A European Strategy for Data

The success of Europe's digital transformation over the next five years will depend on establishing effective frameworks to ensure trustworthy technologies, and to give businesses the confidence and means to digitise.

The Data Strategy and the White Paper on Artificial Intelligence are the first pillars of the new digital strategy of the Commission. They all focus on the need to put people first in developing technology, as well as on the need to defend and promote European values and rights in how we design, make and deploy technology in the real economy.

https://ec.europa.eu/digital-single-market/en/european-strategy-data

The economic value of data sharing

- Data access and reuse can generate **social and economic benefits of 1% to 2.5%** of GDP[1].
- The new measures could **increase the annual economic value** of data sharing by up to €7-11 billion by 2028[2].
- In addition, the new rules will have a **wider impact on the EU economy and society** as a whole:

€1.3 trillion

in increased productivity in **manufacturing** through **Internet-of-Things data** by 2027[3]

€120 billion

of savings per year in the **EU health sector**[4]

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=71186

Health data: Providing better healthcare, improving personalised treatments, helping cure rare or chronic diseases.

Mobility data: Saving more than 27 million hours of public transport users' time and up to €20 billion a year in labour costs of car drivers thanks to real-time navigation[5].

Environmental data: Combatting climate change, reducing CO_2 emissions and fighting emergencies, such as floods and wildfires.

Agricultural data: Developing precision farming, new products in the agri-food sector or new services in rural areas.

Public administration data: Delivering better and more reliable official statistics, contributing to evidence-based decisions.

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=71186

Ensure that the EU Data Strategy facilitates data access for science

TU WIEN Informatics

COMPLEXITY SCIENCE HUB VIENNA

DIO
Data Intelligence Offensive

Univ.-Prof. Dr. Allan Hanbury

Institute for Information Systems Engineering
TU Wien
Favoritenstraße 9-11/194-04
1040 Vienna
Austria

Telephone: +43 1 58801 188310
Mobile: +43 676 978 0991

e-Mail: allan.hanbury@tuwien.ac.at