LUND UNIVERSITY

Growing a Data Sharing Community

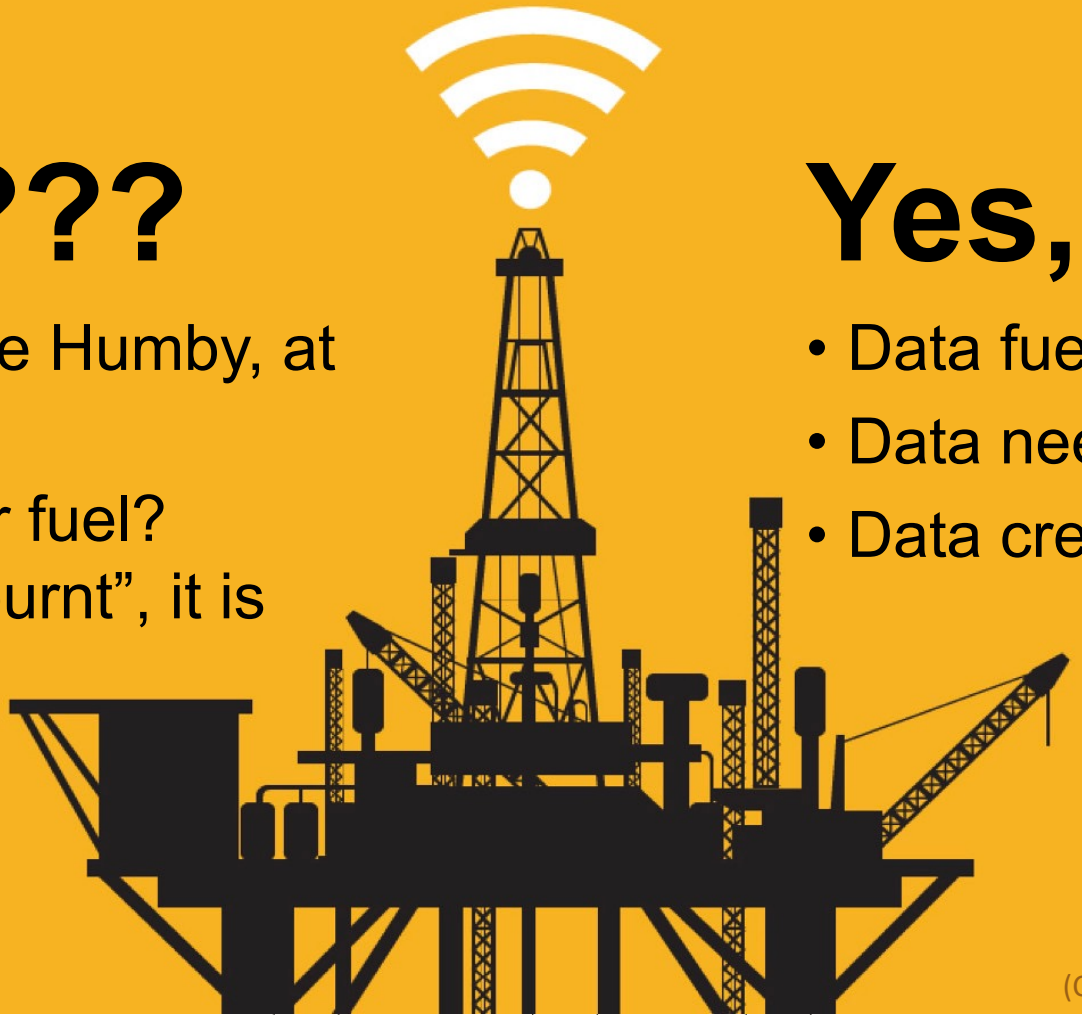
PER RUNESON, PROFESSOR SOFTWARE ENGINEERING

Co-authors
Johan Linåker
Hussan Munir
Thomas Olsson
Karin Rathsman

Data is the new oil!

Is it ????

- Claim by Clive Humby, at Tesco, 2006
- Lubrication or fuel? Data is not “burnt”, it is non-rival



Yes, it is!

- Data fuels businesses
- Data needs refining
- Data creates oligopolies

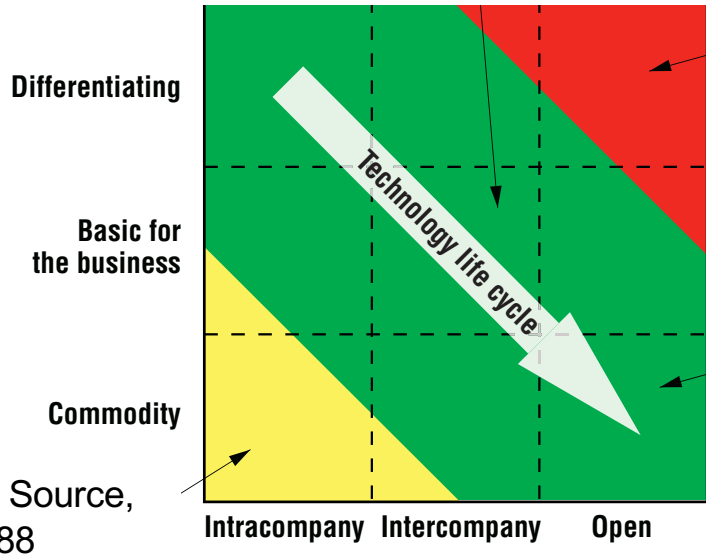
(CC BY-SA 2.0) Gerd Leonhard

Gerd

Data challenges and opportunities

- Costs for *data maintenance*, *quality assurance* and *annotation* is a challenge
- Data will gradually become *commodity* for some functionality

Data sharing?



Lundell *et al.* Commodification of Industrial Software: A Case for Open Source, *IEEE Software*, 26(04):77-83, 2009. https://doi.org/10.1109/MS.2009.88

LUND UNIVERSITY

2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)

Open Collaborative Data – using OSS principles to share data in SW engineering

Per Runeson
Dept. Computer Science, Lund University,
Lund, Sweden
Email: per.runeson@cs.lth.se

***Abstract*—Reliance on data for software systems engineering is increasing, e.g., to train machine learning applications. We foresee increasing costs for data collection and maintenance, leading to the risk of development budgets eaten up by commodity features, thus leaving little resources for differentiation and innovation. We therefore propose Open Collaborative Data (OCD) – a concept analogous to Open Source Software (OSS) – as a means to share data. In contrast to Open Data (OD), which e.g., governmental agencies provide to catalyze innovation, OCD is shared in open collaboration between commercial organizations, similar to OSS. To achieve this, there is a need for technical infrastructure (e.g., tools for version and access control), licence models, and governance models, all of which have to be tailored for data. However, as data may be sensitive**

companies. The cost of curating and maintaining data, will sooner or later exceed its business value.

One approach to adress this issue in software, is open sourcing what has no or little differentiating value anymore. Thereby, the maintenance costs may be shared by multiple companies using the commodity software. As a results, more differentiation can be achieved, and other positive side effects of open innovation may be gained [3], i.e., inflow of ideas and knowledge for innovation. In fact, studies show that the inflow of innovation may be the dominating gain even if the open sourcing was initiated to save costs [4].

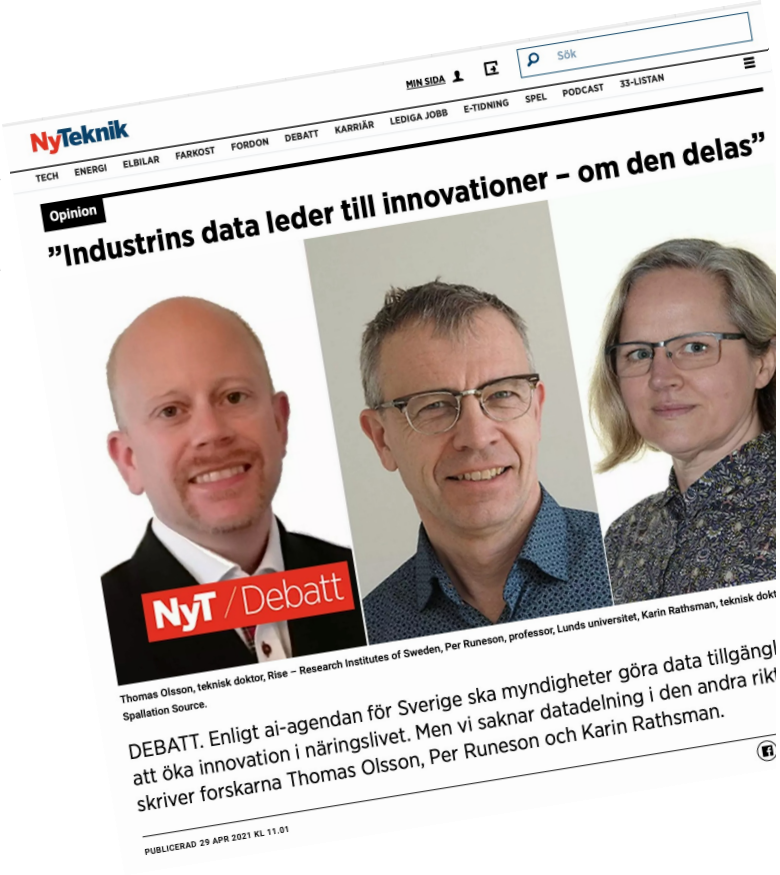
We therefore propose *Open Collaborative Data (OCD)* as

Data sharing?

"Value comes from data being brought together, and that requires organizations to let others use the data they hold"

https://www.bennettinstitute.cam.ac.uk/publications/value-data-summary-report/

LUND UNIVERSITY



NyTeknik

TECH ENERGI ELBILAR FARKOST FORDON DEBATT KARRIÄR LEDIGA JOBB E-TIDNING SPEL PODCAST 33-LISTAN

MIN SIDA Sök

Opinion

”Industrins data leder till innovationer – om den delas”

NyT / Debatt

Thomas Olsson, teknisk doktor, Rise – Research Institutes of Sweden, Per Runeson, professor, Lunds universitet, Karin Rathsman, teknisk dokt

Spallation Source.

DEBATT. Enligt ai-agendan för Sverige ska myndigheter göra data tillgängl att öka innovation i näringslivet. Men vi saknar datadelning i den andra rik skriver forskarna Thomas Olsson, Per Runeson och Karin Rathsman.

PUBLICERAD 29 APR 2021 KL 11.01

Ny Teknik April 2021

Vi uppmanar därför:

• Industriföretag, särskilt i processindustrin, att etablera datadelning som praxis för att möjliggöra innovation kring data.

• Vinnova, att i sitt uppdrag att stödja forskningsinfrastruktur av särskilt värde för svenskt näringsliv, inkludera data från driften av ESS och Max IV som innovationskatalysator för processindustrin.

• Myndigheter, i sina digitaliseringsuppdrag att utveckla öppen källkod och öppna data för at och öka transparens och innovationskraft för Sverige, samt minska inlåsningseffekter.

Genom att dela data, och därmed kostnaden för insamling, förädling och underhåll, kan såväl företag som myndigheter bidra till effektivisering.

Data bidrar då till öppen innovation som en motkraft mot att ”the big five” tar sig in i nya domäner.

Per Runeson, professor, Lunds universitet

Karin Rathsman, teknisk doktor, European Spallation Source

Thomas Olsson, teknisk doktor, Rise – Research Institutes of Sweden

LUND UNIVERSITY

Background and motivation
Open source software –> data

Open source in mobile devices – 2011

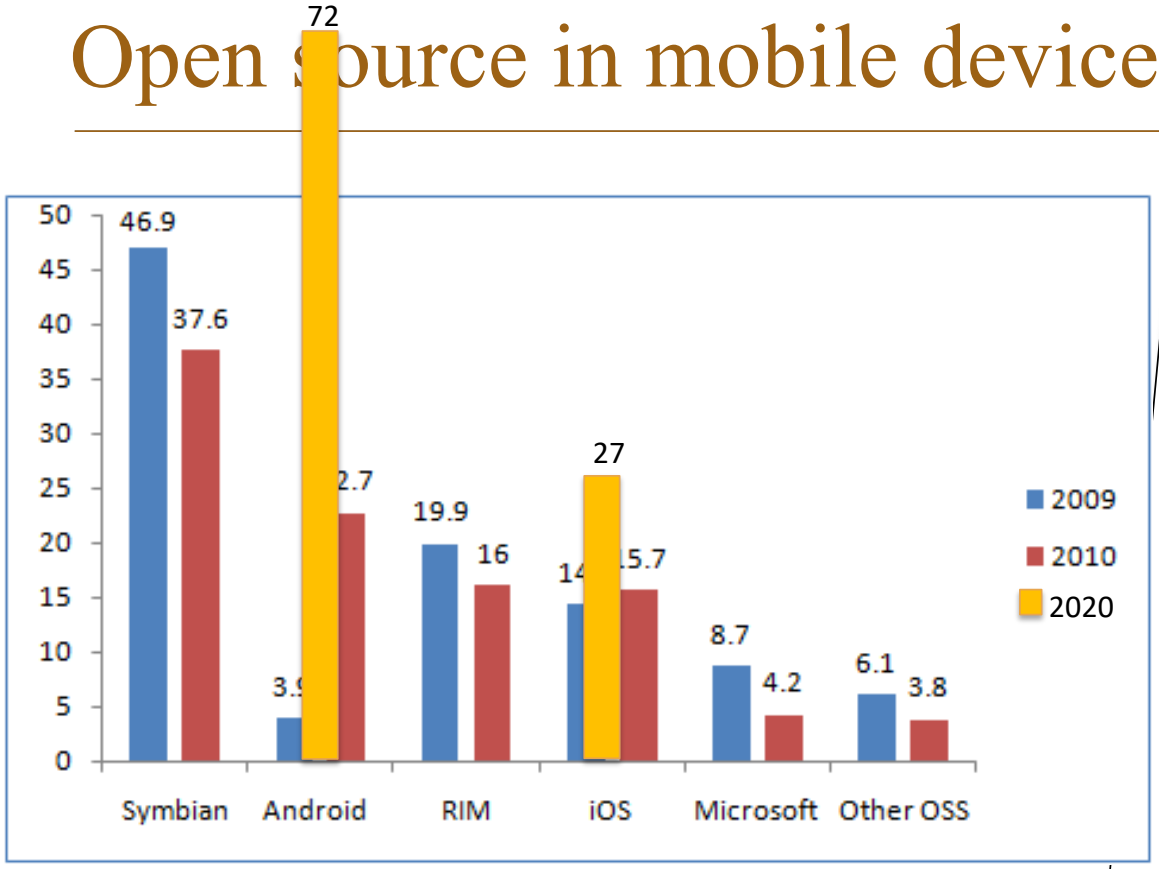


Fig. 1. Worldwide smart-phone Market shares (%) by platform in 2009/2010 (Gartner, 2011)

Open-Source Software Implications in the Competiti
Mobile Platforms Market

Salman Qayyum Mian[1], Jose Teixeira[2], and Eija Koskivaara[3]

[1] Nokia Siemens Networks (NSN), Linnoitustie 6,
02600 Espoo, Finland
Salman.Mian@uta.fi
[2] Turku Center for Computer Science (TUCS), Joukahaisenkatu 3-5 B,
20520 Turku, Finland
Jose.Teixeira@tse.fi
[3] Turku School of Economics (TSE), Rehtorinpellonkatu 3,
20500 Turku, Finland
eija.koskivaara@utu.fi

Abstract. The era of the PC platform left a legacy of competitive strategies for the future technologies to follow. However, this notion became more complicated, once the future grew out to be a present with huge bundle of innovative technologies, Internet capabilities, communication possibilities, and ease in life. A major step of moving from a product phone to a smart phone, eventually to a mobile device has created a new industry with humongous potential for further developments. The current mobile platform market is witnessing a platforms-war with big players such as Apple, Google, Microsoft in a major role. An im

Triggers of Openness – why engage?

- Access to skilled workforce
- Faster development speed
- Low license costs and switching costs
- Flexibility in tool usage and adaptations
- Shared cost with the ecosystem
- Governing ecosystem

Feature Article

How Companies Use OSS Tools Ecosystems for Open Innovation

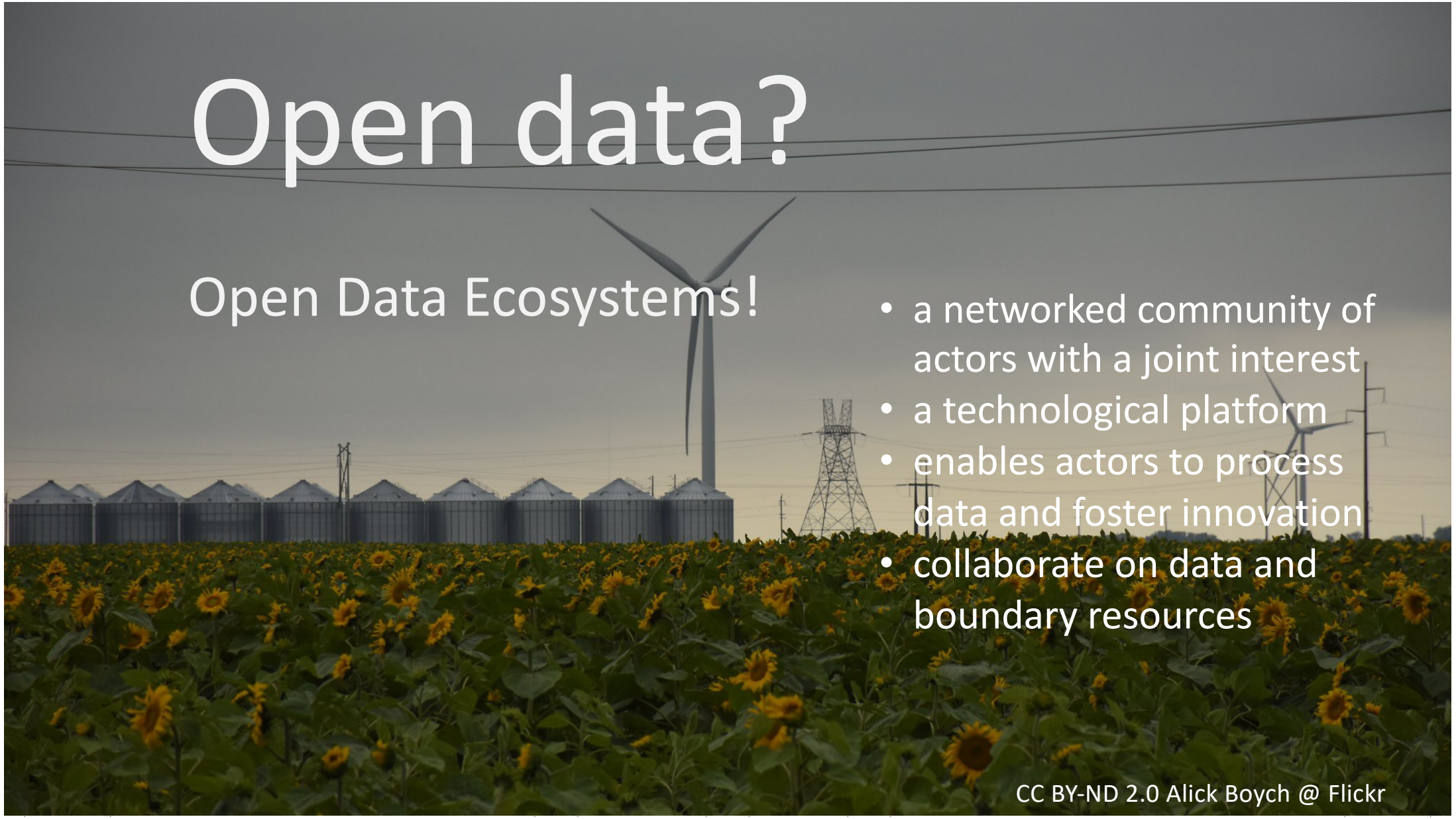
Hussan Munir
Lund University

Per Runeson
Lund University

Krzysztof Wnuk
Blekinge Institute of Technology

Abstract—Moving toward the open innovation (OI) model requires multifaceted transformations within companies. It often involves giving away the tools for prod development or sharing future product directions with open tools ecosystems. Mo from the traditional closed innovation model toward an OI model for software development tools shows the potential to increase software development compete and efficiency of organizations. We report a case study in software-intensive compa developing embedded devices (e.g., smartphones) followed by a survey in OSS communities such as Gerrit, Git, and Jenkins. The studied branch focuses on develo Android phones. This paper presents contribution strategies and triggers for openn These strategies include avoid forking OSS tools, empower d the ecosystem, steer ecosystems thr differentiation

https://doi.org/10.1109/MITP.2019.2893134



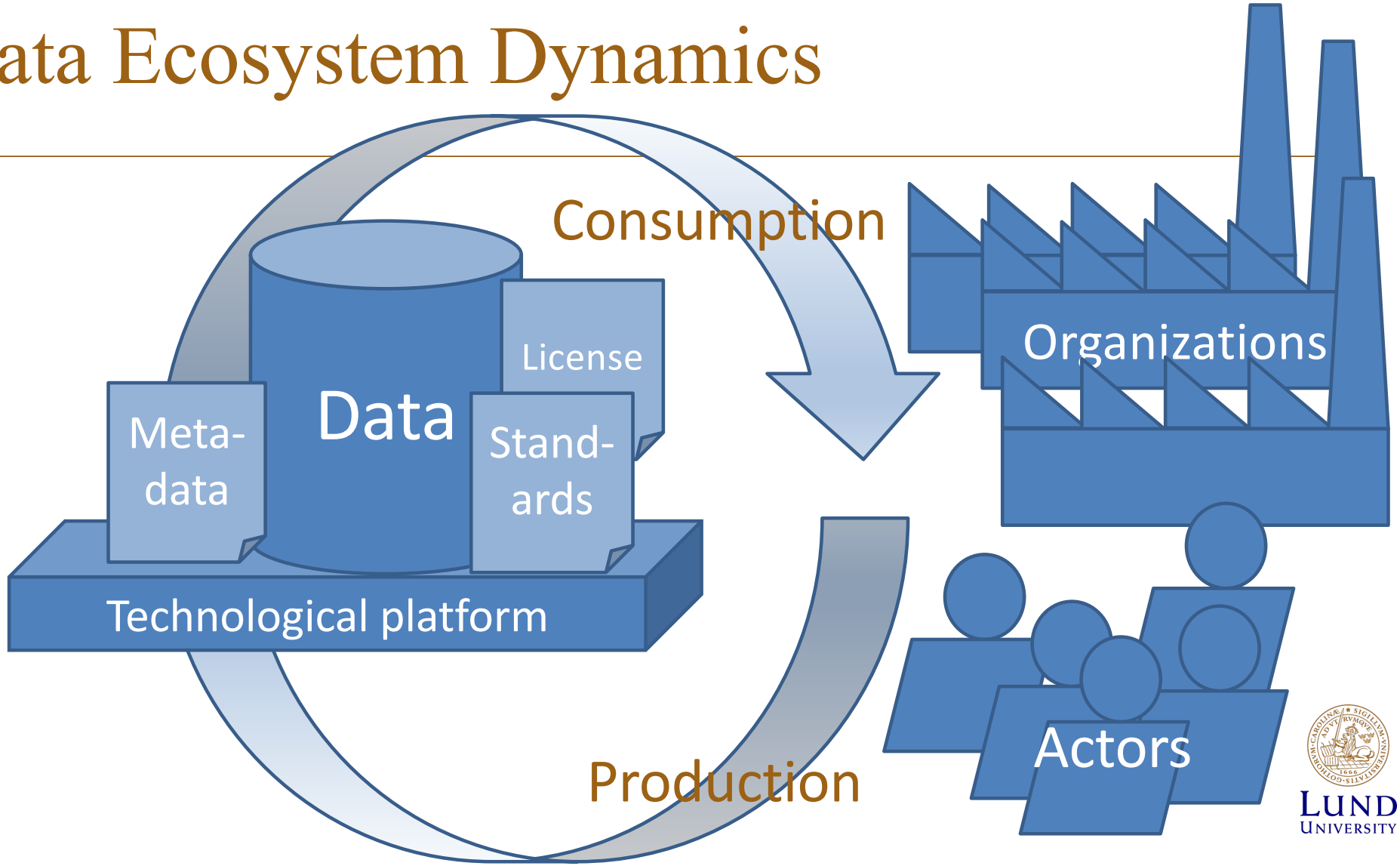
Open data?

Open Data Ecosystems!

- a networked community of actors with a joint interest
- a technological platform
- enables actors to process data and foster innovation
- collaborate on data and boundary resources

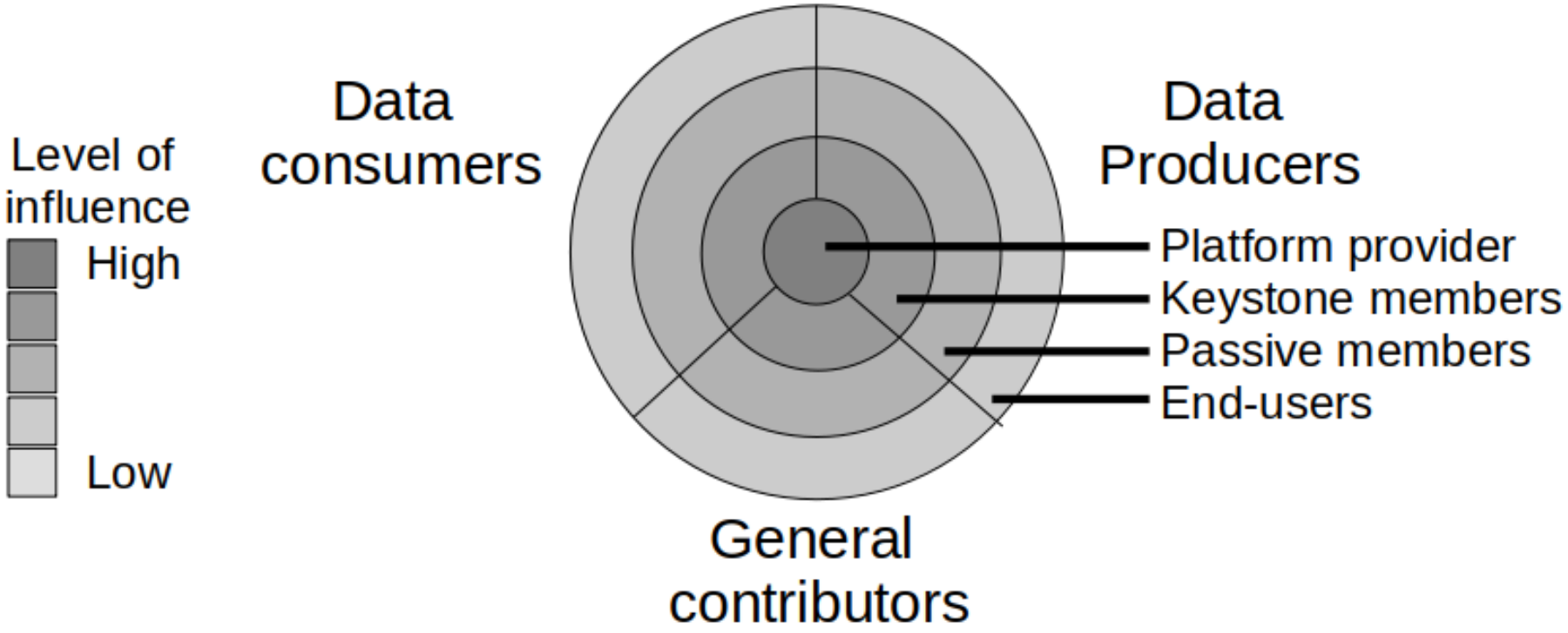
CC BY-ND 2.0 Alick Boych @ Flickr

Data Ecosystem Dynamics



LUND UNIVERSITY

Data Ecosystem Roles



https://doi.org/10.1007/978-3-030-57599-1_22

Commons

- Joint resource
- Mutual dependence
- Governed collaboration

Tragedy of the Commons:
What is best for the individual perspective is bad from the joint perspective.

Ostrom's advice for Commons

Commons are best goverened by rules defined by the **users themselves**, supported by **sanctions** aginst rule breakers. This solution requires some **joint interest** among the users and **communication** means to agree on optimal solutions.

Ekonomipriset 2009

Kungl. Vetenskapsakademien har beslutat utdela Sveriges Riksbank
till Alfred Nobels minne 2009 till

Elinor Ostrom

Indiana University, Bloomington, IN, USA,

"för hennes analys av ekonomisk organisering,
särskilt samfälligheter"

Oliver E

University of Ca

"för hans an
särskilt föret

LUND UNIVERSITY

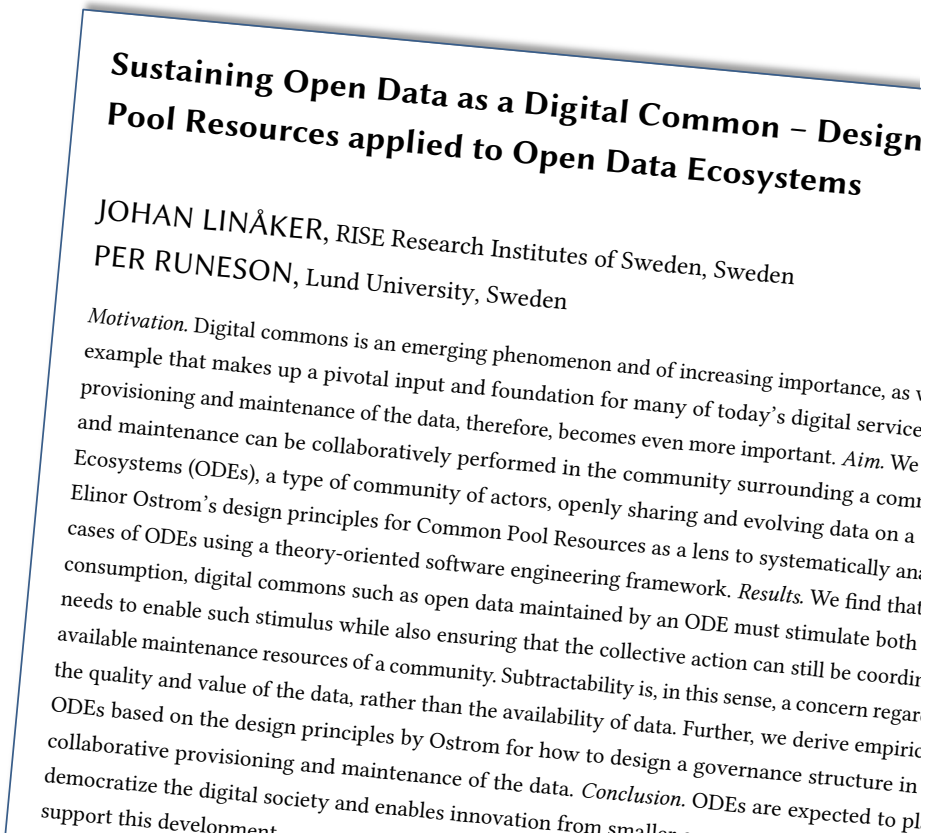
Källa Wikipedia, uppslagsord Allmänningens dilemma, Elinor Ostrom,

Data Ecosystems as Common Pool Resources

Recommendations

1. Boundaries emerge from vision
2. Licenses and processes must balance perceived risks
3. Define rules in dialogue among members
4. Monitoring should be performed by platform provider and members
5. Sanctions to be decided jointly
6. Neutral actor provides trust
7. Recognize and embrace relevant partners
8. Interoperability and internal sub-groups

DOI 10.1145/3555051.3555066



Sustaining Open Data as a Digital Common – Design Pool Resources applied to Open Data Ecosystems

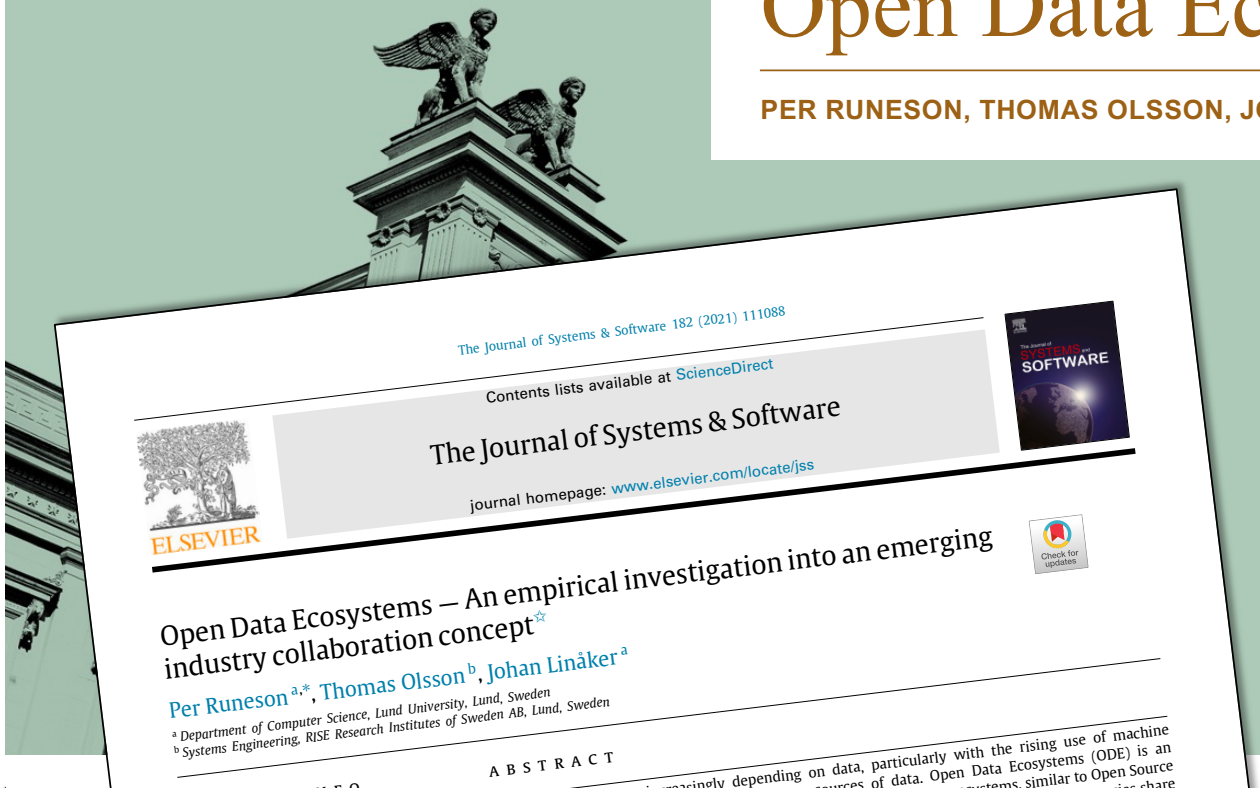
JOHAN LINÅKER, RISE Research Institutes of Sweden, Sweden
PER RUNESON, Lund University, Sweden

Motivation. Digital commons is an emerging phenomenon and of increasing importance, as v
example that makes up a pivotal input and foundation for many of today's digital service
provisioning and maintenance of the data, therefore, becomes even more important. *Aim.* We
and maintenance can be collaboratively performed in the community surrounding a comm
Ecosystems (ODEs), a type of community of actors, openly sharing and evolving data on a
Elinor Ostrom's design principles for Common Pool Resources as a lens to systematically ana
cases of ODEs using a theory-oriented software engineering framework. *Results.* We find that
consumption, digital commons such as open data maintained by an ODE must stimulate both
needs to enable such stimulus while also ensuring that the collective action can still be coordin
available maintenance resources of a community. Subtractability is, in this sense, a concern regar
the quality and value of the data, rather than the availability of data. Further, we derive empiric
ODEs based on the design principles by Ostrom for how to design a governance structure in
collaborative provisioning and maintenance of the data. *Conclusion.* ODEs are expected to pl
democratize the digital society and enables innovation from smaller
support this development

LUND UNIVERSITY

Essential concepts of Open Data Ecosystems

PER RUNESON, THOMAS OLSSON, JOHAN LINÅKER



The Journal of Systems & Software 182 (2021) 111088

Contents lists available at ScienceDirect

The Journal of Systems & Software

journal homepage: www.elsevier.com/locate/jss

Open Data Ecosystems — An empirical investigation into an emerging industry collaboration concept☆

Per Runeson [a,*], Thomas Olsson [b], Johan Linåker [a]

[a] Department of Computer Science, Lund University, Lund, Sweden
[b] Systems Engineering, RISE Research Institutes of Sweden AB, Lund, Sweden

ABSTRACT

increasingly depending on data, particularly with the rising use of machine ... sources of data. Open Data Ecosystems (ODE) is an ... systems, similar to Open Source ... share

Emerging data ecosystems

JobTech
- Labor market
- Job ads
- Public-driven
- Organization-centric

ESS-CSDL
- Industry 4.0
- Alarm data
- Business-driven
- Organization-centric

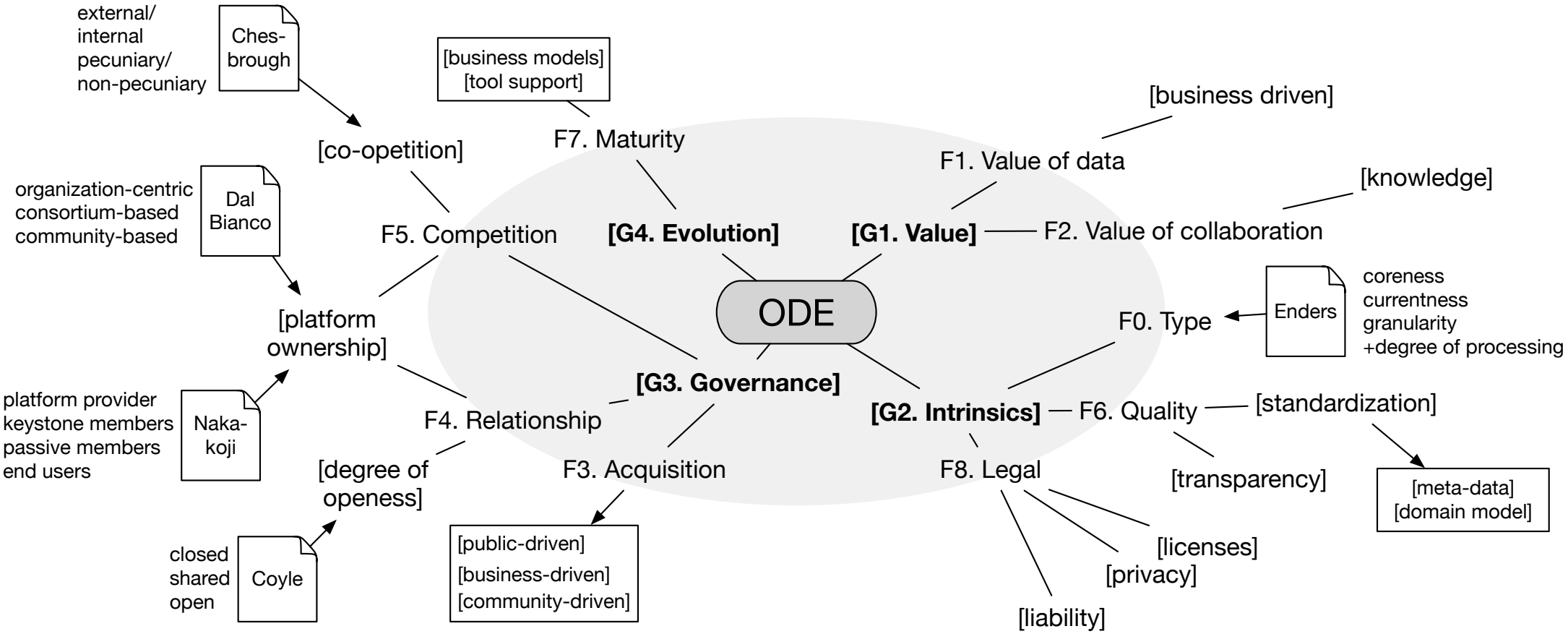
RoDL
- Automotive
- Traffic video
- Business-driven
- Consortium-based

JobTech Development

ess EUROPEAN SPALLATION SOURCE

101 0010
110 1111
100 0100
100 1100
RoDL

LUND UNIVERSITY



Open Data Ecosystems – an empirical investigation into an emerging industry collaboration concept

LUND UNIVERSITY

G1. Value

The value of data (F1) and the value of collaboration around the data (F2) are two sides of the same coin. One or the other may be the primary value, but they are highly intertwined.

CC-BY 2.0 Mike Lawrence @Flickr

LUND UNIVERSITY

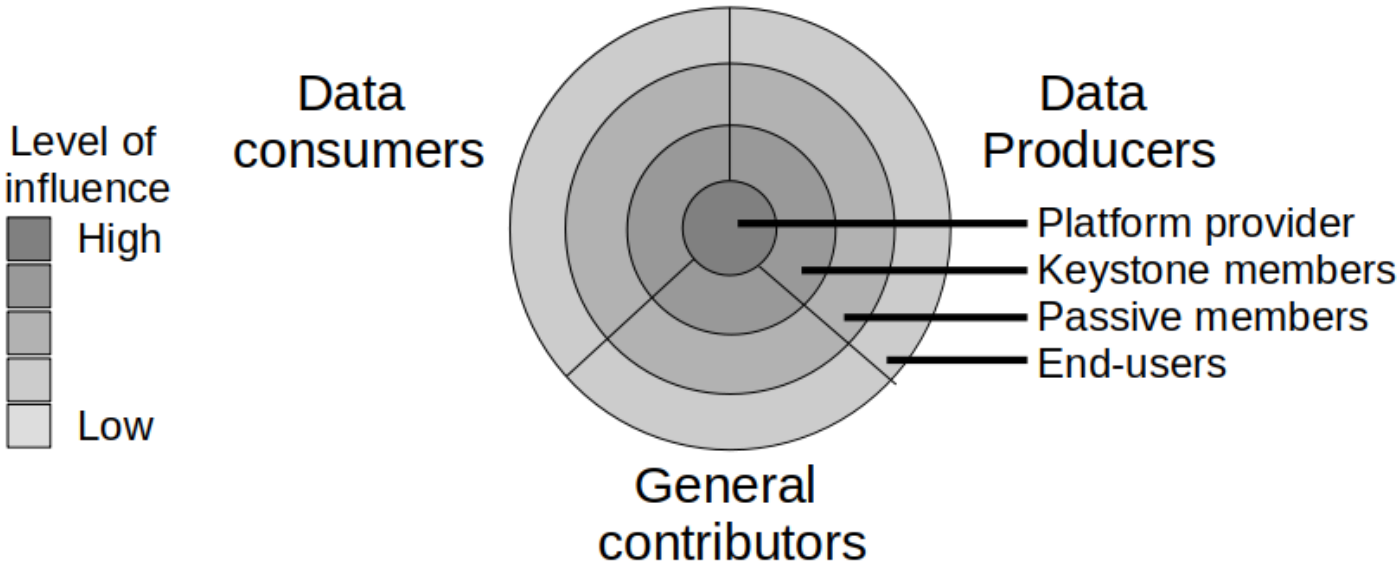
G2. Intrinsics

Intrinsics,
or internal characteristics of data

- data type (F0)
  - coreness
  - currentness
  - granularity
  - degree of processing
- data quality (F6)
  - correctness
  - provenance
  - meta-data

- legal aspects (F8) is tightly connected to data, although they also connect to governance of the ODE.
  - licenses
  - privacy
  - liability

G3. Governance – platform provider

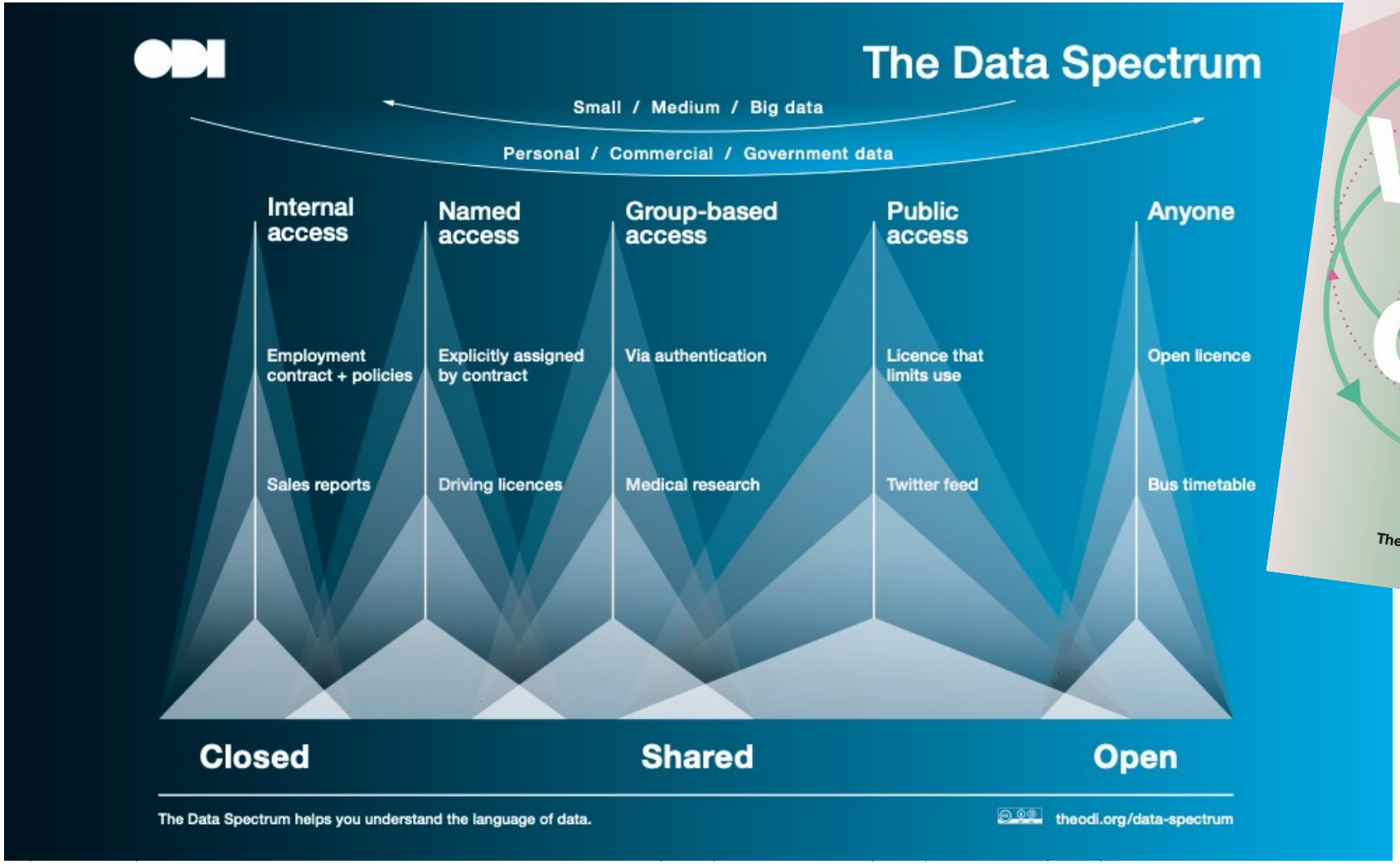


There is a need for an independent platform provider to ensure trust

Initiation may be public-driven, business-driven, or community-driven

LUND UNIVERSITY

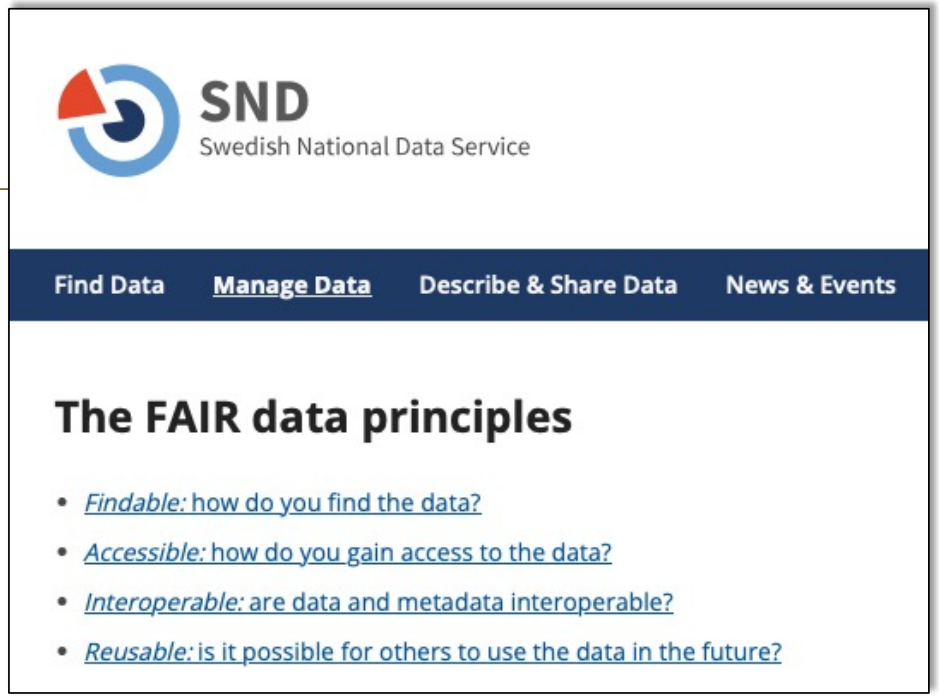
G3. Goverenance – How open is open



LUND UNIVERSITY

G3. Governance – How open is open?

- FAIR Research Data
  - Findable
  - Accessible
  - Interoperable
  - Reusable



"As open as possible and as closed as necessary

https://snd.se/en/manage-data/prepare-and-share/FAIR-data-principles

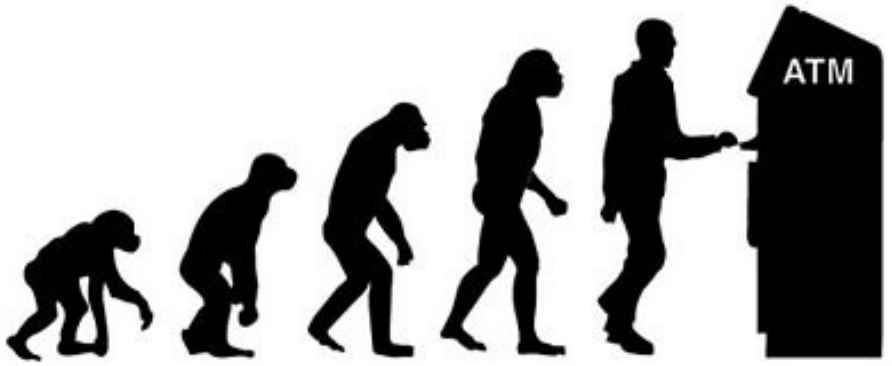
LUND UNIVERSITY

G4. Evolution

The concept of and strategies for open data ecosystems are still in their infancy

Need for knowledge:

- how to integrate ODEs into an organization's business model
- tools to support ODEs and enable data sharing should be developed and standardized



(CC BY-NC 2.0) ThomasThomas@Flickr

LUND UNIVERSITY

Findings for data ecosystems

Value

Focus on business value in the data or collaboration

Intrinsics

Data coreness, currentness and granularity
Standardize format and legal framework

Governance

Level of openness and platform ownership
Relationship and competition must co-exist
Data acquisition incentives

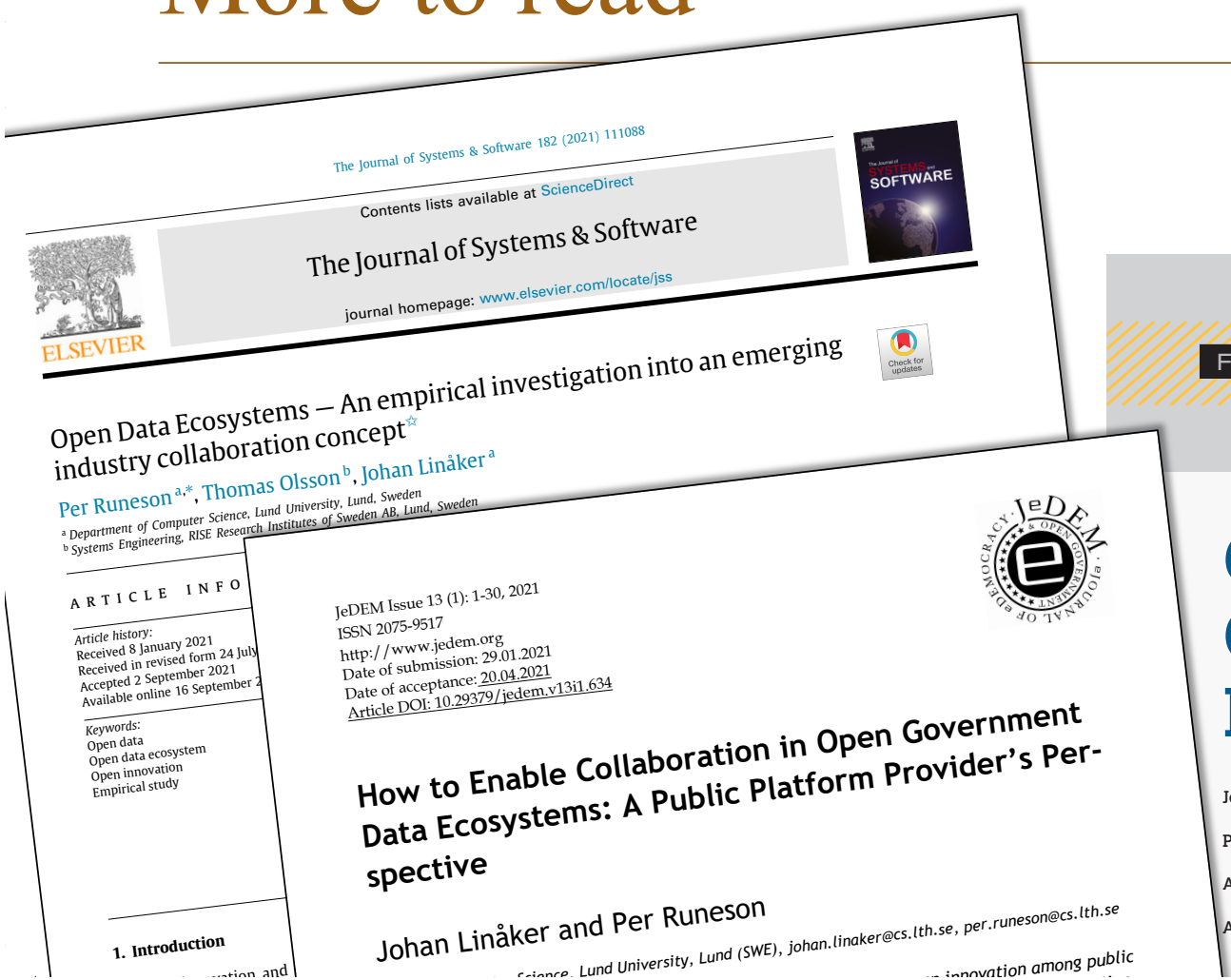
Evolution

Advance business models and tool support

CC BY-SA 2.0 Jocelyn Kinghorn @ Flickr

More to read

https://doi.org/10.1016/j.jss.2021.111088
https://doi.org/10.29379/jedem.v13i1.634
https://doi.org/10.1109/MS.2021.3118123



The Journal of Systems & Software 182 (2021) 111088

Contents lists available at ScienceDirect

The Journal of Systems & Software

journal homepage: www.elsevier.com/locate/jss

Open Data Ecosystems — An empirical investigation into an emerging industry collaboration concept☆

Per Runeson[a,*], Thomas Olsson[b], Johan Linåker[a]

[a] Department of Computer Science, Lund University, Lund, Sweden
[b] Systems Engineering, RISE Research Institutes of Sweden AB, Lund, Sweden

ARTICLE INFO

Article history:
Received 8 January 2021
Received in revised form 24 July
Accepted 2 September 2021
Available online 16 September 2

Keywords:
Open data
Open data ecosystem
Open innovation
Empirical study

1. Introduction

JeDEM Issue 13 (1): 1-30, 2021
ISSN 2075-9517
http://www.jedem.org
Date of submission: 29.01.2021
Date of acceptance: 20.04.2021
Article DOI: 10.29379/jedem.v13i1.634

How to Enable Collaboration in Open Government Data Ecosystems: A Public Platform Provider's Perspective

Johan Linåker and Per Runeson

Lund University, Lund (SWE), johan.linaker@cs.lth.se, per.runeson@cs.lth.se

innovation among public

FOCUS: GUEST EDITORS' INTRODUCTION

Collaborative Aspects of Open Data in Software Engineering

Johan Linåker, RISE Research Institutes of Sweden

Per Runeson, Lund University

Anneke Zuiderwijk, Delft University of Technology

Amanda Brock, OpenUK

More to come:
B2B Data Sharing for Industry 4.0 Machine Learning

Kowalkowski

Ahmed

Prof. **Per Runeson**, PhD Student **Konstantin Malysh**, Software Engineering, LU
Prof. **Christian Kowalkowski**, PhD Student **Tanvir Ahmed**, Industrial Marketing, LiU

Business models (LiU)

Two disruptive and interrelated transformations:

1) **digitalization** changes sociotechnical systems,
2) **servitization** entails the shift from selling products to 'product-as-a-service' business models

Collaboration tools (LU)

Git, Jenkins and Gerrit, provide a low-threshold entry o open source software (OSS). Data ecosystems need "an underpinning technological platform".

Runeson

Malysh

ELLIIT
Excellence Center at Linköping – Lund in Information Technology
LINKÖPING UNIVERSITY
LUNDS UNIVERSITET
BLEKINGE TEKNISKA HÖGSKOLA
HÖGSKOLAN I HALMSTAD
WASP WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM
LUND UNIVERSITY

per.runeson@cs.lth.se
http://www.lth.se/digitalth