Pain & Action Points from the Opendata.ch/2020 Forum

Pain & Action Points from the Opendata.ch/2020 Forum **1**
- Filling Female Data Gaps 1
  - Pain Points 1
  - Action Points 2
- Research Data Connectome 3
  - Pain Points 3
  - Action Points 3
- Serving, not Spying 4
  - Pain Points 4
  - Action Points 5
- Better use of data in SwitzerlandBetter use of data in Switzerland 5
  - Pain Points Workshop 1 6
  - Pain Points Workshop 2 6
  - Action Points Workshop 1 7
  - Action Points Workshop 2 7
- Solid basis for committed debate and innovation 8
  - Pain Points 8
  - Action Points 9
- Open decentralized data storage 9
  - Pain Points 9
  - Action Points 10
- Shared ecosystem of energy data in Switzerland 11
  - Pain Points 12
  - Action Points 12

Filling Female Data Gaps

A world where data is available without gender bias and gaps

Pain Points

- Pain point 1: Lack of diversity in teams, decision makers and contributors, which are responsible for governing, collecting, publishing, etc. data.
- Pain point 2: Lack of education regarding diversity - for everyone (men and FINT).
- Pain point 3: Lack of awareness about how to handle biased data for developing

algorithms. Models and algorithms are not conceived to include the bias from the beginning. Algorithms need to try to fix the biased data by e.g. checking the outcome. source

- Pain point 4: Data collected is not representative of minorities. Diversity in itself is complex. When we analyse data we try to simplify. That's a challenge in itself.

- Pain point 5: Data is published without precision of the "subjects" so if data is provided by a majority of "men" it will then be generalised and "applied" to all people.

Action Points

Stakeholders	**Action Points**
Meeting organisers (Co-)Working culture in companies, events, etc.	Asynchronous meetings (contribute at your own pace, in written or orally etc.) Opening up several ways of communication in order to overcome language barriers. Strengthening code of conduct to ensure harassment-free environments (incl. mentioning them at the beginning of events etc.). [potentially publish a handbook on working culture for diverse teams by an institution like witty works...]
NGOs / NPOs education institutions	Develop interventions at schools or programs for schools to motivate young children to get familiar with data science and strengthen diversity in this realm.
NGOs / NPOs	Child book with underrepresented groups as
NGOs / NPOs	Initiative like ruby on rails for data. Or expanding the offering of an initiative like ruby on rails for data.
NGOs / NPOs Media	Give visibility to role models of minorities working in data science.
Lobbying	Strengthen inclusion and diversity in the curriculum of Lehrplan 21.
PH/teacher education institutions	Make this topic part of the gymnasium teacher education in all cantons of Switzerland - use existing slots (PH Bern) /gain additional slots (Thurgau) to raise awareness for diversity and inclusion re data
Data publishers (state)	Develop a "diversity data catalogue": a list of datasets that are needed / useful / missing to tackle diversity and underrepresentation in Switzerland. This should be done with collaboration with the underrepresented groups.
Data collectors (state or NGOs)	Something like GDPR for data privacy, but for collecting data. Guidelines for ensuring diversity when collecting, presenting

	and publishing data.
Us as engaged citizens	Making the existings conferences for women in statistics & data science more popular / recognised.
	Best practices for presenting diversity issues in data publication or analysis.
NGOs / NPOs	Building communities of underrepresented stakeholders.
NGOs / NPOs	A group or representatives of minorities to reach out to as a company when working on data science projects. Potentially in collaboration with existing communities such as Frauenhäuser, Quartiertreff or Streikhaus. Or political parties

Research Data Connectome

To connect and organize (meta)data for research sustainably across disciplines, in order to make it widely accessible, interoperable and valuable.

Pain Points

Society:

- Data Literacy (quality, standards): Understanding of documentation (“language”, quality of data”).
- Trustfulness of sources, data, metadata, etc.

Community (Research, Education and Innovation):

- Data is not findable.
- Missing standards (e.g. metadata / data descriptions, quality).
- Infrastructure (interconnected).
- Cultural aspects of sharing (“fear of sharing”).
- Conflicts of interests between Stakeholders.
- Data Literacy (quality, standards).
- Trustfulness of sources, data, metadata, etc.
- Collaboration is a necessity for this (on a daily basis, not only on specific projects).

Action Points

Stakeholders	Action Points
Government	- International alignment of licenses.
Data- and Service Providers	- Improve understanding. - Communications - Part of service design

Researchers	- Educating junior researchers - Using the “right language” for research dissemination
Institutes and Departments	- Trainings - Education Sessions
Research Support Services	- Trainings - Education Sessions
Data- and Service Providers	- Educating users
Researchers	- Provide feedback / needs to infrastructure providers. - Establish Communities of Practice.
Government	- already very active, use it!
Infrastructure Providers	- Establish useful infrastructure - Establish interconnections with data- and service providers - Collaborate.
Users	- Provide feedback / needs to data- and service providers. - Establish Communities of Practice.
Data Providers	- Communicate data - Integrate
Institutes and Departments	- Communicate - Teach users
Societies (of interest), Umbrella organisations	- Communicate - Teach users

Serving, not Spying

Ten years from now, I don't feel observed and manipulated through the Data being collected about me but I can rely on being served by data for better services, better policies, and a better environment.

Pain Points

- Pain point 1:
  The remuneration / business model of the digitalized society

- Pain point 2:
  decentralization on data collection at the level of companies and institutions

- Pain point 3:
  lack of Data literacy / awareness

- Pain point 4 :
  difficulty to control the multitude of “data collectors” & ways to collect Data (for ex. Cumulus karte..? Visa? Netflix..etc..), the legislation also cannot keep up

- Pain point 5 :
  the definition of Good vs. Bad usage of Data

Action Points

Regaining control:

- an international call for a trusworthy, consumer owned Data wallet is launched: A vault / Data wallet: the control is fully back into the hands of the consumer:

- an international infrastructure to administrate & redistribute data sets
- -Civil society/ governments should start the discussion and agregate the will of their citizens

Regaining trust:

- Data privacy protection is by default
- international instance & legislaton / citizens as warrant of the fairness / the point of view of the consument / indivdual
- the personal Data should be encrypted and act as half of the key : the companies and institutions don’t have direct access to the information
- Companies don’t collect any data
- open source -> transparency about data flows

changing the business models:

- consumers and companies have to enter a creative negociation process to invent new retributions that are not bound to data flows
- companies have to see that they don’t need to work personal information

Raising Data Literacy and awareness

- Reword personal data to personal information to make it less abstract
- make data Literacy a large public topic which is not only for nerds, but show the broad and accessible applications of Data

Better use of data in SwitzerlandBetter use of data in Switzerland

Public and private sector data is open, shared, and used for public and private good within a trustworthy data ecosystem ("Swiss / European DataSpace“).

This workshop was treated by two separate workshop groups and therefore has a higher amount of pain points and action points.

Pain Points Workshop 1

- Pain Point 1: Lack of documentation / metadata on “raw” open data
    - lack of information about the time at which some data were “true” (what if we had no information about the expiration date of the food we buy and eat??) as well as about their accuracy

- Pain point 2: data curation
    - data curation takes time (financial & personal resources required)
    - who can support e.g. researchers in data curation?
    - responsibility
    - education (data management)
    - deleting data that is not useful or outdated

- Pain point 3: Education in the use of open data - how can it be used? what are the dangers? how is data produced? what is the meaning of data and how can it help to make sense of the world?
    - the general public knows how to “digest” information (news, wiki, …); the ability to extract valuable information out of “raw” data is not yet widespread
    - Mistrust of the public in data use, sharing, processing...

- Pain point 4: Misuse of personal data by private companies / government makes individuals suspicious on the thematic
    - Insufficient governance of data handling (law lags behind innovation)

Pain Points Workshop 2

- Easy **understandable APIs** instead of .csv or .xls-Files
- Money is not only spent for impressive looking dashboards or frontends, but also for **data cleaning and testing**.
- **Data are not accessible.** For example: Patient data belong to the patients. They agree, that their data can be used by scientists. But the scientists treat the data by their own. Ownership switch
- **Data are not provided -lack of money**. Everybody saves on providing data. We save on data --afterwards we pay.
- **Demand** from data consumers is not strong enough - no **pressure** to start a data publishing organization.
-
- **OGD Licence vs existing laws:** open is clearly defined, whereas on opendata.swiss we have a mix up. Standard, re-usable licenses designed to provide consistent and broadly recognized terms of use are needed, swiss and european wide;
- **Data literacy** is low. train people
- **Data is difficult to find** when starting from a concrete goal or question
- Data is often difficult **to navigate** (how do I get from one data set to a related one, and how do I even find related data sets) - **Orchestration of data**

Action Points Workshop 1

Stakeholders	**Action Points**
media	FAQs, “Open Data Bot” (AI, Knowledge Graph)

public administration	data education/training; coordination (e.g. Cogis); data knowledge (metadata)
opendata.swiss	data broker, data librarian, data clustering - foster cooperation between data user and data provider
Government	Unified Open Government Data License
data user	clear data requests (obligation)
investors	sustainable data investments (data pipeline)

Summing up action point “Swiss Data Space”:
"Foster cooperation between data user and data provider. On the data user side this means education (data literacy), good data use cases und sustainable data investment (not only “dashboards”). On the data provider side this means good data publication (incl. data knowledge and metadata), and unified open data licences. And third data brokers/data librarians/data exchange infrastructures are needed (e.g. opendata.swiss) so data users and data providers can match their data offers and needs."

Action Points Workshop 2

Stakeholders	Action Points
data producers	● be transparent about known flaws in data (it’s better to state/take on than “hide”) ● Actively seek feedback from data users, to provide better documentation, metadata, and to be able to curate data better ● dedicate more resources to data documentation (not just capture) ● provide support to the users if required ● provide versioned data
libraries	● raise awareness ● courses (coffee lectures, workshops, courses) ● direct support in implementing policies (national, international) ● developing metadata standards, at least giving best advice and support ● give best practice examples
standardization body	● standards shouldn’t only concern syntactic/formal aspects, but also “semantics” (too often, only keys are standardised whereas values are not) ● should provide tools to measure compliance
Schools &	● Integrate data literacy into curriculum, probably on all

Universities	school levels ● showing our data world - it's everywhere ● educate children and adolescents by bringing together different perspectives, showing best practices ● create possibilities for youth to develop their own approaches (e.g. through hackathons, pilot programs, accelerators etc)
"Success Story tellers" / media	● provide data, code and hands-on tutorials ● management tools: e.g. knowledge graphs ● data journalism
Private companies	● Educate about the use/benefit of open data and shared data ● Conduct pilots with sharing their data with the ecosystem and the public ● Collaborate with the public sector on defining standards/governance for the use of data
Regulators	● Define what's allowed, and what's not allowed, to do with data, and for what purpose (eg. for research, for profit, for media use etc.) ● CH: align with EU legislation (GDPR) for a dynamic regulatory space

Synthesis:

- **Tools for compliance with standards would be valuable.**
- **All of the stakeholders (particularly schools, university, libraries) can be "success story tellers" - use the resources/tools available**
- **Media & journalists have the power to frame the narrative**

Solid basis for committed debate and innovation

Data creates a critically discussable basis for political analysis and joint shaping of our future just as they enable innovations in everyday life - free access and responsible use are indispensable for an enlightened information society.

Pain Points

- Pain point 1: Obligation to data publication (current data, quality data), for public and private data (at least metadata to private data and its accessibility)
- Pain point 2: free accessible infrastructure for dealing with (big) data
- Pain point 3: free access to knowledge re use of infrastructure and data handling (data skills, data literacy)

- Pain point 4: Culture of data trust thanks to methodological competencies (understanding of data production and methodical limits of meaningfulness [Aussagekraft])

Action Points

Stakeholders	Action Points
NGO	Political “hackdays” (analysing and reproducing public information, Abstimmungsbüchlein reverse engineering)
Public institutions (libraries, archives, educational institutions, data holders)	Provide Infrastructure (as an alternative to GAFA) open for everybody
Media	Publication of data sources, scripts for visualisation
Public / private institutions	Offer further education vouchers [Bildungsgutscheine]
Public administration / executive	Create an open data strategy (bottom-up)

Open decentralized data storage

Data is a new form of currency: It is too powerful to be controlled by centralized private entities - let's build a decentralized system to store, share and use data.

Pain Points

- Lukas: FAGAM / big data players setting the rules and evading responsibility. No global response / control. Many data breaches with little consequences. No awareness how powerful data is.

- Ambrogio: Data is “somewhere” and, even if I personally am aware that I produce and send data, I do not know where it is and what it is used for. There is generally no control and where I can get control or demand a removal of my data, the process is convoluted and painful to follow (that means: I will not ask for it)

- Christian: There are some early stage initiatives (->MyData proto operators) that already offer such services, but no one has meaningful traction yet and usability is still an issue.

- Oleg: When we think “data”, we think “database”. We don’t really have an overview of where the data resides that is critical to our day-to-day experience, GDPR laws are a door opener but hardly the solution to the problem. On a state level, we can only

glimpse at what data drives policymaking, and usually late after the cards have fallen. Important data is siloed, secured-by-obscurity, unusable out of context. The penalties for the misuse of our data, for the lack of transparency on an institutional or governmental level, are not a serious obstacle to the determined.

- Moritz: Fragmented data ecosystem, outdated structures and still a lack of transparency and knowledge among the general population. Private sector actors invest most in structures and develop know-how. Accountability issues.

  (Fragmentiertes Datenökosystem, veraltete Strukturen und noch mangelnde Transparenz und Wissen in breiter Bevölkerung. Privatwirtschaftliche Akteure investieren am meisten in Strukturen und entwickeln Know-how. Accountability issues.)

- Marc: Data are analog and digital, they are in different places, at public administration and companies, just with me. Too many data and which data is really necessary.

Action Points

Stakeholders	**Action Points**
State / government	• **Regulate big tech companies on a global level (FAGAM) to make them implement distributed digital tools to enable wider participation.** • Demand more transparency from parliamentary/political processes for democratic access to participative decision-making - enabled by real-time and historic access to indicators on which decisions in government are based. • Financial support for startups tackling the issues.
Industry associations	• **Industry associations should start working together more closely to develop a common vision** • **Start realizing the value for their members of working towards a more participatory and inclusive data ecosystem.** • Make it easier and more compelling to work with civil society groups like Opendata.ch to achieve these goals.
Political decision-makers	• **Provide the legal basis for clear data portability rules and for a decentralized data infrastructure.** • **Enable (cantonal) pilots for digital participation.**
Big players	• **Listen and respond to user demand for better data ownership and portability.** • **Control interdependency by thinking more organically about data rather than building more giant (and unsustainable) data warehouses.** • Brace yourself, invest in, and prepare to step aside and be a form of globalized government, as the next generation of digital upstarts makes your core business obsolete.
Start-ups and innovators	• **Leverage competitive advantage for next-gen companies that bank on decentralized data storage.**

	• **Design user-friendly services that implement high-privacy, high user-control solutions.**
Civil society	• **Raise awareness, start discussions, inform political bodies and the population at large of data privacy issues and opportunities.** • Twisting “fear” narrative into a positive one, where citizens see opportunities. • **Reviewing solutions put forward by other players - recognizing and popularizing the best efforts.** • Leave old thinking patterns, be curious, be open, experiment with new approaches
Standardization bodies	• Become more agile, more efficient.
End users	• **Take ownership of the topic** • **Improve your own digital literacy** • Don’t be “lazy”, don’t renounce to your rights too easily! • Install Metamask ! Use Tor ! Write a dat:// website ! • Act in responsibility, “I know what I do”

Shared ecosystem of energy data in Switzerland

Data about energy production and consumption is shared openly by Swiss utility companies, promoting open innovation and helping to reach the "Agenda 2030" climate goals.

Pain Points

- Pain point 1: Open data could open up the market to **new competitors** (e.g. from abroad), so incumbents are against it.
- Pain point 2: Finding/creating i**ncentives to share data**; lack of clarity regarding **usage scenarios** / use cases / business cases
- Pain point 3: Lack of **data management** know how and processes within companies
- Pain point 4: **Data incongruencies** between different stakeholders (cf. power dynamics)
- Pain point 5: **Competition regulations** as a hindering factor (monopolies are not completely free to act on the market)
- Pain point 6: Many small **monopolistic** stakeholders on their respective territories
- Pain point 7: Lack of sensitization of **political decision makers** (exposure to regional lobbying?)

Action Points

Stakeholders	Action Points
BFE digital innovation office	- communicate existing (international) standards for energy data - Publicity for successful use cases / success stories - Study on needs of utilities regarding opening up their data
VSE Verband Schweizer Energieversorger	- Clear guidelines on which data can be openly published (gemäss Energiegesetz / Strom VG / Strom VV / VSE Richtlinien und Handbücher)
Research centers, z.B. EIFER	- identify use cases for open data in the energy sector
2-3 Branchen-Newsletters / Messe “Innovationsforum” in ZH	- Success stories kommunizieren
Innovation labs within utilities: SGSW, Energie 360°, Romande Energie, larger “Stadtwerke”, ewz, EKZ, WWZ, all of the EZS partners	- Identify and test use cases for open data (Prototype, PoC, MVP…)
Energie Zukunft Schweiz	- Use cases identifizieren und Projekte lancieren (z.B. geeignete Partner finden)
People in charge of IT at utilities	
Kleinere lokale Energieversorger	
Grosse Energieversorger	
Startups	- identify use cases for open data in the energy sector
Swissgrid	