Digital Responsibility

Additional: 7 Responsibility Goals - Open Operations - Green IT - cloud-native Software - Community Health - Sustainable Identities
Relax. Turn off your phone. Leave the office. Let Managed SaaS do the work for you.

Find out more on www.cloudical.io
Dear reader,

the past year brought new challenges we all couldn’t image to face. The world is different now. We were forced to rethink sovereignty, sustainability, and freedom. And we learned how fast we can be susceptible to blackmail via our dependencies. But the last year also revealed how digitization has changed our perceptions of the world with quick information, pictures running over the world, fake news which became weapons in the war in the Ukraine...

So, I was very happy to learn about the seven Digital Responsibility Goals – Jutta Juliane Meier is introducing to you. Digitization is a matter of fact, we use digital services in our daily life, and it will increase even further. But we need to learn how we could and should use it, where are limits of IT, where might be danger, how should we use AI? Digitization is only offering a toolset and we must use it consciously and responsibly. For me, the “goals” should be the base of all digital transformation – what do you think?

Every theory must prove itself in praxis, so we take a deeper look into “transparency” with the Open Operations Manifesto coming from the open operations community. The Manifesto brings in best practices for operating IT, transparency, and fault management. And you are welcome to join the community!

Responsible usage of cloud also needs to be sustainable and open source! What we all can do gaining this goal Holger Dyroff tells you: by developing a sustainable data strategy and using more green software. Green software enables quicker deployments, quicker processing, less storage, and less energy usage. Cloud-native software is an important step towards green software. What cloud-native software means and what it needs to develop cloud-natively Karsten Samaschke is presenting in his article.

Suvish Viswanathan and myself took one point of “responsibility” to reflect on: how could we deal responsibly with online data – our own and data from customers? And how we can combine responsibility with sovereignty?

Coming back to communities Georg Link is introducing a measurement for community health. Are you curious what that means? The example of the OpenInfra community will shed light on. And furthermore, I visited the Container Days in Hamburg and the Univention Summit in Bremen and brought back impressions, information, and the good feeling of open source community!

In the “Security” section Sarah Polan is thinking about sustainable identities. What are identities in the digital world and how do we handle machine’s identity?

Stay open, healthy and peaceful. And become curious about your own responsibility. What is your goal?

Yours,

Friederike
The Cloud Report is published quarterly at the beginning of January, April, July and October. The Cloud Report is available in two versions: the online edition which can be accessed online or via downloaded from the homepage, and the printed edition which can be subscribed at a price of 20 Euros a year via the personalized customer portal in which a personal account is set up. When you register, please enter the information to specify the execution in which you want to obtain the report. The subscription can be cancelled at any time via the personal access of the subscriber, by mail at the latest two weeks before the publication of the new edition via the e-mail address: sales@cloudical.io. We collect relevant personal customer data for the subscription. Individual issues can also be purchased without a subscription at a price of: 5 Euro, in this case too, relevant personal data will be used to fulfil the purchase contract. Further information can be found at: http://the-report.cloud/privacy-policy.
TABLE OF CONTENT

EDITORIAL
Editorial 1

FOCUS
The Digital Responsibility Goals 4
Open Operations Manifesto 10
Green IT
and sustainable data strategy 14
Unlocking the Potential of cloud-native Software 18

COMMENT
Putting an end to the exploitation of customer data 13
Carry Responsibility 23

THE HUMAN COMPONENT
Measuring Open Source Project Health 26

CONFERENCE REPORT
Container Days 30
Digital sovereignty – only with open source! 32

SECURITY
Big Data: Friend or Foe? 34
The Digital Responsibility Goals – WHY?

We need a human- and planet-centered digital transformation

Digital technologies improve people’s lives, but technological innovations and their use must therefore be geared to the wellbeing of people and society. Therefor the Digital Responsibility Goals define a framework and work toward a trusting, ethically sensitive, and sustainable digital transformation. In this article we introduce the 7 Digital Responsibility Goals.

Likewise, the internet and digital technologies bring negative side effects: for example, in many places in the world, especially in totalitarian states, the internet is restricted, regulated, monitored, and used for their propaganda; also, fake news and hate speech poison the atmosphere and make social discourse more difficult.

Oftentimes, critical decisions about the future of digital developments are made without a clear framework. Trustworthy, ethically sensitive, and sustainable guidelines that focus on the benefits for people are missing. The Digital Responsibility Goals (DRGs) aim to define this framework and work towards a truly human- and planet-centred digital transformation.

The Digital Responsibility Goals – HOW?

We provide a human-centered digital transformation framework

Leading organizations and companies are committed to the United Nations’ 17 Sustainable Development Goals (SDGs). Following the same logic, addressing the digital dimension, the 7 DRGs aim to guide decision makers, companies, and other stakeholders, such as researchers and users, to develop trustworthy digital products and services.
The DRGs provide an opportunity for various stakeholders and decision-makers from businesses, regulators, academia, and civil society to form a common agenda and plan a common course of action to deal with a human-centered digital transformation.

The Digital Responsibility Goals – WHAT?

We measure a human-centered digital transformation step by step

DRG #1 “Digital literacy” is at the forefront of the DRGs, as only knowledge, education, and comprehensive information can be the basis of all self-determination and decision-making. This includes access to the technological infrastructure itself. Goals 2 to 6 of the DRGs are oriented along the data value chain – from the security of the system (DRG #2 “Cybersecurity”) as the technological basis, to the protection of personal data (DRG #3 “Privacy”) – as a European promise – and fair handling, a new understanding of how to deal with non-personal information and data (DRG #4 “Data Fairness”), responsible collection, processing and understandable evaluation of data (DRG #5 “Trustworthy Algorithms”) and transparent communication of behaviour (DRG #6 “Transparency”). Finally, the DRGs form a big social bracket – the protection of our identity as well as the preservation of human agency (DRG #7 “Human Agency & Identity”) in the digital space (fig. 1). In the representation of DRGs, people and their identity are elevated from their previous positioning as marginal figures to the sovereign bracket and at the same time to the centre of digital transformation.

Developed in a consortium consisting of leading academics, NGOs, and industry experts, the DRGs cover 7 areas where we see scope for commitments that go beyond compliance with existing laws and regulations. We already described 5 guiding criteria per Digital Responsibility Goal to make the responsible behavior measurable (fig. 2).
Guiding Criteria and Examples

DRG #1 Digital Literacy and accessibility of technology as the fundamental basis for trust and acceptance of digital innovations: starting with the individual. Digital competence and access to digital products, services, and processes are prerequisites for the acceptance of digital technologies. They are the basis for all other goals of the Digital Responsibility Goals, enable the assessment of the trustworthiness of offerings, and put humans at the center. They are what makes the multi-layered human identity in the digital space possible in the first place.

These are the 5 guiding criteria of DRG #1 Digital Literacy:

- **DRG #1.1** The information offered for digital products, services, and processes must be designed individually and in a way that is suitable for the target group.

- **DRG #1.2** Access to digital products, services, and processes must be reliable and barrier-free.

- **DRG #1.3** The acceptance of digital products, services, and processes must be proactively considered in design and operation. This includes measures on equity, diversity & inclusion.

- **DRG #1.4** Education on the opportunities and risks of digital transformation is essential – everyone has a right to education on digital matters.

- **DRG #1.5** The education and information offered should be designed to create awareness of related topics such as sustainability, climate protection, and diversity/inclusion (for example along the UN SDGs) where applicable.

Example of a successful implementation of DRG #1: “DRG 4 GovTech”: In the design and operation of an authority website for the electronic application of a car license plate, principles of accessibility were implemented in accordance with DRG guiding criteria #1.2, for example in accordance with BITV 2.0 (Barrier-free Information Technology Ordinance). This includes perceptibility, usability, comprehensibility, and robustness for the relevant target groups.

DRG #2 Cybersecurity as the crucial foundation and basis of secure digital technologies. Cybersecurity arms systems against compromise and manipulation by unauthorized persons and ensures the protection of users and their data – from data collection to data utilization. It is a basic prerequisite for the responsible operation of digital solutions.
These are the 5 guiding criteria of DRG #2 Cybersecurity:

DRG #2.1 Developers, providers, and operators of digital products, services, and processes assume responsibility for cybersecurity. Users also bear some of the shared responsibility – awareness (see DRG #1) is essential here.

DRG #2.2 Developers, providers, and operators of digital solutions are responsible for appropriate security measures and are constantly developing them further. Products, services, and processes are designed from the outset to be resistant to compromise and abuse by unauthorized persons (security by design).

DRG #2.3 A holistic view and appropriate implementation are considered along the lifecycle, the value chain, and across the entire service or solution.

DRG #2.4 Developers, providers, and operators of digital products, services, and processes must account for how they provide security for users and their data – while maintaining necessary trade secrets and information security.

DRG #2.5 Business, politics, authorities, civil society, and science must jointly and collaboratively shape the framework for cybersecurity with appropriate objectives, measures, and targets. This requires open and transparent cooperation (for example according to principles of “responsible disclosure”).

Example of a successful implementation of DRG #2: “DRG 4 Finance”. A bank offering online services has been certified ISO 27000 to – in accordance with DRG guiding criteria #2.2 - demonstrate it possesses a robust security system based on appropriate measures to prevent unauthorized access to private information, internal systems, and networks. Ultimately this helps minimize the risk of security breaches, making the company more reliable and reputable in the eyes of potential customers.

DRG #3 Privacy as a European promise: Privacy is part of human dignity and a prerequisite for digital self-determination. Protection of privacy – with a consistent purpose limitation and data economy – allows users to act confidently in the digital world.

Privacy by design and default enable responsible data usage. Users are given control and providers must account for how they protect privacy.

These are the 5 guiding criteria of DRG #3 Privacy:

DRG #3.1 Operators and providers of all digital products, services, and processes must take responsibility for protecting the privacy of their users.

DRG #3.2 When dealing with personal data, strict pur-

pose limitations and data economy are observed.

DRG #3.3 Privacy protection is considered throughout the entire lifecycle. Privacy protection is the default setting.

DRG #3.4 Users have control over their personal data and its use – this includes the rights to access, rectify, erase, restrict processing, object, avoid automated decision-making, and ensure data portability.

DRG #3.5 Providers must account for how they protect users’ privacy and personal data – while maintaining necessary trade secrets and information security.

Example of a successful implementation of DRG #3: “DRG 4 ResponsibleTech”: An online search engine assumes responsibility for protecting the privacy of its users in accordance with DRG guiding criteria #3.1. Privacy protection is clearly anchored in the organization, and sufficient financial resources are available for additional expenses incurred as a result. Responsibilities for privacy protection in the organization are defined, with a clear mandate at the highest organizational level.

These are the 5 guiding criteria of DRG #4 Data Fairness:

DRG #4.1 When collecting data, proactive care is taken to ensure that it fairly reflects and represents the context in which it is collected.

DRG #4.2 In digital ecosystem structures, the mutual exchange of data between all parties involved must be clearly described and regulated (data governance). The goal must be fair participation in the benefits achieved through the exchange of data.

DRG #4.3 Developers, providers, and operators of digital solutions must clearly define and communicate the purpose (wherever possible) with which they use and process data (including non-personal data). Exceptions are approaches like “open data”.

DRG #4.4 Data is designed “FAIR”, especially for use cases relevant to society as a whole – “FAIR” stands for Findable, Accessible, Interoperable, Reusable.

DRG #4.5 Data providers must be equipped with mechanisms to control and withdraw their data – they shall be able to have a say regarding the usage policies.
Example of a successful implementation of DRG #4: “DRG 4 GovTech / DRG 4 Mobility”. In line with DRG guiding criteria #4.4 a municipal government has a dedicated strategy to ensure the use of data based on the “FAIR” principles. It takes a number of dedicated measures with the aim of bringing data including traffic information, environmental data, and economic indicators to the public and promoting its use.

DRG #5 Trustworthy Algorithms
“Trust by Design” – through trustworthy algorithms: Once the data has been collected, it must be processed with the goal of trustworthiness. This is true for simple algorithms as well as for more complex systems up to autonomously acting systems (AI = Artificial Intelligence for example).

These are the 5 guiding criteria of DRG #5 Trustworthy Algorithms:
- **DRG #5.1** Algorithms, their application, and the data-sets on which they are based are designed to provide the highest level of fairness and inclusion.
- **DRG #5.2** The individual and overall societal impact of algorithms is regularly reviewed, and the review documented. Depending on the results, proportional measures must be taken.
- **DRG #5.3** The results of algorithmic processing and their occurrence are comprehensible.
- **DRG #5.4** AI systems must be designed to be reliable and precise to be able to withstand subtle attempts to manipulate data or algorithms. It must be possible to reproduce results where possible.
- **DRG #5.5** AI systems must be designed and implemented in such a way that independent control of their mode of action is possible.

Example of a successful implementation of DRG #5: “DRG 4 Industry”: A startup that develops and markets AI tools for industrial applications implements measures to maintain fairness and inclusion in accordance with DRG guiding criteria #5.1. These includes active measures to increase diversity in developer teams and the establishment of an AI Ethics Board.

DRG #6 Transparency must form the basis to guide the actions of all stakeholders in the digital supply chain to create trust: openness and transparency making the difference. Pro-active transparency for users and all other stakeholders as to which principles underlie digital products, services, and processes, as well as transparency on the digital solution itself and its components, is created. Principled behavior is an important building block for building trust.

These are the 5 guiding criteria of DRG #6 Transparency:
- **DRG #6.1** To gain the trust of users, organizations establish transparency about their digital ventures and solutions – for the final digital products, services, and processes as well as the organization, business models, data flows, and technology behind them.
- **DRG #6.2** Transparency is implemented in interactive communication (for example, between providers and users), and mechanisms for interaction are actively offered.
- **DRG #6.3** The use of digital solutions is designed to be transparent wherever there is a digital interaction between people and the digital solution takes place (for example, the use of chatbots).
- **DRG #6.4** In addition to transparency for users, transparency should also be provided for professionals – while maintaining the necessary business secrets and information security.
- **DRG #6.5** Organizations must outline how they will make transparency verifiable and thus hold themselves accountable for their actions in the digital space.

Example of a successful implementation of DRG #6: “DRG 4 Health”: In a tool for diagnostic imaging in line with DRG guiding criteria #6.1 it is made transparent to doctors upon use that image recognition and analysis is used for diagnostic purposes in healthcare. Furthermore, this is also clearly communicated to relevant patients in the doctor-patient conversation.

DRG #7 Human Agency & Identity are critical guideposts and the precondition for digital development. Digital products and services must be human-centric, sustainable, inclusive, and developed under human oversight. It’s about each and every one of us. Even in the digital space, we must protect our identity and preserve human responsibility. Preserving the multi-faceted human identity is a prerequisite for any digital development. The resulting digital products, services, and processes are human-centered, inclusive, ethically sensitive, and sustainable, remaining in human care at all times. Only in this way can digital technology promote the well-being of humanity and have a sustainable impact.

These are the 5 guiding criteria of DRG #7 Human Agency & Identity:
The preservation of the multifaceted human identity is a basic requirement and must be the basis for any digital development. The resulting digital approaches are always user-centric – they respect personal autonomy and dignity, limit commoditization, and open new perspectives.

Sustainability and climate protection must be part of digital business models and implemented in practice (especially in accordance with the UN Sustainable Development Goals).

Digital products, services, and processes promote responsible, nonmanipulative communication. Where possible, communication takes place unfiltered.

Digital technology always remains under human authorship and control – it can be shaped throughout its deployment.

Technology may only be applied if it is of use to individuals and mankind and promotes welfare.

Example of a successful implementation of DRG #7: “DRG 4 ResponsibleTech”. In line with DRG guiding criteria #7.5 a technology company conducts an impact assessment on the effects of the technology of facial recognition. Discovering the risk of malicious and unfair use, it decides to clearly limit the use of that technology to dedicated, risk-mitigated use cases and transparently communicates that decision.

Summary and Outlook

Digital technologies have the potential to improve people’s lives. At the same time, the internet and digital technologies bring negative side effects. Technological innovations and the use of innovative technologies must therefore be geared more to the wellbeing of people and society in the future than is already the case today.

As a benchmark for players in the digital space, the Digital Responsibility Goals define this framework and work toward a trusting, ethically sensitive and sustainable digital transformation among decision-makers. The Digital Responsibility Goals pursue an integrative and combined approach of all relevant actors in order to promote trust in digital technologies and business models. The Digital Responsibility Goals are accompanied by a guideline to measure the degree of successful value-based digital transformation, to strengthen the responsibility of digital play and to give participants clear guidance for their digital strategies.

Based on the logic of the UN SDGs, the DRGs are also about presenting very complex tasks for society as a whole in a simplified framework for action using target images as orientation, initiating responsible action at all target levels, and enabling and accompanying their progress in a comprehensible manner. Acting along the goals is of course desired at all levels and does not have to be done chronologically from 1-7. Rather, the aim is to create awareness and sharpen attention for challenges and opportunities in the necessary fields of action – to shape our sustainably responsible digital space for our society.

As a target picture to shape a sustainable human-centered digital transformation, the DRGs offer an opportunity to promote greater responsibility in the digital space across sectors. Responsible behavior and responsible leadership all along the data life cycle is at the core of establishing trust. By adhering to the framework of the DRGs, and implementing it with a dedicated management system, building trust will no longer be a random by-product, but a pro-active and targeted achievement.

Any development towards more digital responsibility does not happen in isolation inside an organization only, but also takes place in an ecosystem if not the wider environment and on a societal level, therefore: “Our activation plan depends on an interactive and engaging ecosystem, that wants to put the user, the human, back in the centre of digital innovations. It’s about collaboration and it’s about putting perspectives together from different sectors to find solutions that can create a more sustainable and trustworthy world. We need responsible leadership. More than ever.” – Jutta Juliane Meier, Founder & CEO, Identity Valley.

Sources

1. https://identityvalley.org/drg
Open Operations Manifesto

We – the founding and supporting organizations¹ – proclaim our primary objectives to be transparency along with the sharing of knowledge and are in the process of building a community of practice – Open Operations.

Open Source technology has become increasingly mature. However, the greatest software is useless if it cannot be operated with full confidence. For this to happen, it is imperative that skills and knowledge be built up, fostered, and retained. In an increasingly competitive market for skilled people and increasingly complex IT systems, this is becoming an ever greater challenge for governments, institutions, and companies. How can we operate digital offerings in a self-determined, secure, and supremely excellent manner?

The answer to this must be the collectivization of operational knowledge, just as it has been practiced for many years with software code. This manifesto provides the foundational principles for these goals and invites other organizations to join this movement.

We are building a community of practice

From time to time, we can become mentally stumped and may need a sparring partner to think in other, more innovative ways. Open Operations is building a community of practice to keep the barrier to market entry low and create a thriving environment for easy and convenient exchange. Regardless of the individual level of knowledge, everyone is encouraged to contribute and thereby enriches this community of practice with their knowledge and individual views. Communities of practice connect through various channels, serving both asynchronous and synchronous communication needs in the form of pair operating or on a larger scale to solve these challenges cooperatively. We are hence building a network by encouraging organizations to foster and connect internal and external communities of practice.

We share knowledge

The availability of knowledge and skilled engineers is the limiting factor for many organizations to adopt, leverage, and successfully operate complex technology. For true self-determination in the ever-growing digital domain, we need to build skills and share the gained knowledge freely and in unlimited ways, including in the realm of operational
knowledge — which is exactly the way we started with software code almost half a decade ago. We encourage organizations to collectivize operational knowledge and share it with the entire Open Operations community.

We firmly believe that failures make us experts. In the principle of learning from our mistakes the lessons learnt add to our expertise. Thus, the way we handle errors is how we become better. We feel that handling our mistakes in the best way possible means communicating about them openly and without shame. We are convinced that the culture we build around failures is fundamental for an innovative environment: not being afraid of making mistakes opens the door to try out new approaches, without fear of failing. We encourage people to speak out – not just when they are successful, but also when they fail. We are sure that this is indeed one of the best ways to prevent ourselves and others from making the same mistakes. We hold that to achieve the state of complete transparency, the history of failures and incidents should be recorded and preserved. Our goal is that others can truly experience trust in this way, rather than just being told they can.

We are transparent about our operational processes

We share our internal processes for the sake of transparency. We firmly believe that transparency leads to better and more reliable processes. We are convinced that publishing well-documented operational processes offers a great opportunity to gather feedback and gain inspiration from the wider public.

Likewise, we are convinced that by sharing our operational experiences, we can be of inspiration and benefit to others. In today’s business environment, most companies face the same challenges. By sharing knowledge the same way, the Open Source community shares software, we are turning a common knowledge base into a reality, allowing everybody to benefit.

We aim to get away from being a big black box and seek
to become a source of shared knowledge. We consider the disclosure of internal processes to be an essential step in this direction.

About this manifesto

The Open Operations Manifesto is initiated by the Sovereign Cloud Stack Community\(^2\) to foster transparency, share knowledge, and build a community of practice within the domain of IT operations. It invites organizations to adhere to these principles and to take an important step forward both in their current operations and into the future.

We encourage you to join our mailing list\(^3\) and attend our regular meetings. If you want to contribute to our manifesto or spread the word by signing it, please read through the README\(^4\) or reach out to us at info@openoperations.org.

Open Operations Manifesto – Based on commit c4b068 – Released under CC BY-SA 4.0

Sources
1. https://openoperations.org/#supporting-organizations
2 https://scs.community/
3 https://scs.sovereignit.de/mailman3/postorius/lists/list.openoperations.org/
4 https://github.com/SovereignCloudStack/open-operations-manifesto/#readme
5 https://creativecommons.org/licenses/by-sa/4.0/

Authors:
Eduard Iltrich,
Felix Kronlage-Dammers,
Katharina Heinritz,
Cemil Degirmenci,
Maximilian Wolfs,
Ralf Heiringhoff,
Kurt Garloff,
Friederike Zelke
We must no longer accept that online services misuse customer data for monetary purposes under the guise of "customer experience". A transparent data policy is urgently needed.

Every time customers visit a website or use an app, they must expect the operator to store their data and use it for purposes they hardly ever knowingly consent to.

Sometimes the intended use is harmless and even helpful – such as using preferences and past behaviour to improve navigation and enhance the customer experience. Consumers thus benefit from personalized feeds, tailored suggestions or even loyalty rewards.

Most of the time, however, unclear cookie descriptions and dark-pattern designs only serve to mislead users, and spy on them. The operators will collect data and sell it to third parties. In the best-case scenario, this will be to advertisers or direct marketing agencies that want to influence customer needs. In the worst case, the data ends up in the hands of malicious parties, on black market forums, or with aggressive spammers.

The reasons for this illicit activity are almost always financial. It is an easy extra income that increases sales and perhaps makes investors happy. It is the consumer who loses out.

One popular tactic to obtain data is through the application of third-party tracking cookies across websites. Through the interaction of hardware, software and smart applications, trackers can collect valuable information about all user activities, and can paint a picture about a person’s interests, preferences, and even their location. We need to be aware of this: there is no such thing as anonymity on the internet. In fact, data collectors are so sophisticated that they can paint a surprisingly accurate picture of exactly who we are by combining all the data traces we leave behind. It is no coincidence that this approach is called fingerprinting.

Fortunately, consumers themselves can play an important role in stopping this. As users become more and more data-conscious, they increasingly demand control and transparency over the use of personal information, putting more pressure on operators and data collectors to act ethically and with data privacy in mind.

Companies need to ask themselves: is it really still worth risking one’s own reputation to gain financial advantages through the exploitative use of data? We at Zoho, for example, advises companies not to use third-party trackers and to have a fully transparent data usage policy towards their customers.

In fact, distancing oneself entirely from any form of data analysis, collection or use can be a worthwhile strategy. This creates a brand image that values customer trust over short-term financial gains. By removing third-party trackers, the success of online campaigns may not be as easy to monitor, but the benefits to the customer experience could make it well worth it. We removed all third-party cookies from our websites years ago and instead developed our own solution to track online marketing performance. As an alternative, a feedback and rating system that enables honest and open communication with customers could allow them to actively express their opinion on a particular service instead of making assumptions based on indirect and dishonestly gathered information.

Another innovative way companies can improve the customer experience is by investing in AI-powered automation. Corresponding software then not only ensures quick access to important customer data, via specially developed CX tools, but also significantly reduces the time spent on back-office tasks. This allows employees to concentrate on more important tasks, such as personal contact with existing or potential customers.

Data-driven insights are not the only way to improve the customer experience. Only brands that act fairly towards their customers will remain competitive and sustainable in today’s competitive markets.
Green IT and Sustainable Data Strategy

Sustainable Data Strategy

Sustainability can be understood in different ways. In the narrow sense, sustainability means climate neutrality, ecology, green IT, and so on. In a broader sense, however, sustainability has a much greater social value. Solutions that are developed and used openly, transparently, and collaboratively, that promote sovereignty and independence, are also sustainable. In software, for example, open source solutions are sustainable. Well-documented open source code prevents the same solutions from having to be developed repeatedly and increases the security and quality of the software, as in most cases the code has been reviewed by several developers. But transparent solutions also offer sustainability and flexibility in terms of processes and hardware solutions. Let us take the example of an ideal data centre that is climate neutral or even generates green energy. If the basic structure in combination of hardware and software can be communicated openly and transparently, then this solution can be transferred to other data centers. Data centers are always dependent on their specific environment regarding the surrounding climate or waste heat utilization, but transparently communicated experiences and information nevertheless help to build data centers in a climate-positive way at some point.

In addition, if you interpret sustainability more broadly and always look at what opportunities arise from this approach, then a sustainable data strategy requires that organizations use and support open source, that solutions are developed with open standards so that they are able to share this solution sustainably with other organizations as well, so that they do not have to develop the same thing again on their own. This is one of the prerequisites for ultimately achieving a sustainable data strategy.

At this year’s it-sa in Nuremberg, ownCloud presented its contribution to Green IT and the concept of sustainable data strategy. This article takes a closer look at how they want to support climate-neutral IT in a sustainable way. We at the cloud report have the issues of net zero, sustainability, and carbon neutral very much at heart – that is why we are always glad to present successful cases.
Sustainability also means that organizations are able to comply with the policies and regulations that they set for themselves as an organization, that are required by legal regulations or that are established in agreement with business partners or customers. This is also part of a sustainable data strategy, which allows you to comply with very different rules, store them transparently for all parties involved and adapt them at any time. This is particularly relevant for organizations in regulatory industries.

Green IT

Sustainability therefore means long-term and ecological solutions, but also openness, transparency, flexibility, being able to work in hybrid environments, or rather to work in whatever environments the user wishes. Green IT should always be sustainable in a broader sense, even if the focus is initially on the ecological side. Green IT is often about the question of the carbon footprint or, for example, about the question: How performant am I with the software when I upload 1000 files? How much electricity do I really consume to do that? Fortunately, this is easily measurable nowadays. This is where the newly introduced ownCloud Infinite Scale² solution comes in. Thanks to it, the user is able to upload an amount of small files up to ten times faster. Moreover, higher speed helps saving energy. If I upload faster, if I have to burn fewer CPU cycles to do certain things, I logically save energy at the same time.

Green IT and sustainable data strategy thought together

True carbon neutrality requires close cooperation between data centers and software providers. Much is demanded of
hardware providers in terms of carbon footprint, but software developments can and must be energy-efficient in the long term, so that the interaction of software and hardware can result in significantly lower energy consumption. After all, in the long run, providers of digital services will only work sustainably if carbon neutrality is not achieved through climate certificates.

In a broader sense, sustainability has a greater social value. Solutions that are developed and used openly, transparently, and collaboratively, that promote sovereignty and independence, are also sustainable.

One of the challenges of open source software providers with open standards is that they thus have open offerings that can be operated virtually anywhere. Whether the software is really used in a climate-neutral way always depends on the environment in which it is applied. There are numerous examples of green software, but it is important to use it responsibly, because even with climate-active software you can waste energy if you use it incorrectly or operate it in an energy-inefficient environment. There are some positive examples of open source projects and software, but there are still too few software vendors who think and develop in this direction. The first open source software has been awarded with the Blue Angel. This is the software Okular, a PDF viewer from the KDE Project. This shows that it seems to be possible for software in general to be awarded a Blue Angel. The underlying criteria can become a basis for software development in the long term.

An important step in this direction is definitely to be aware of which partners organizations want to work with and to develop criteria on how the respective organization wants to achieve true carbon neutrality. Many providers claim to be carbon neutral, but only achieve this by buying climate certificates, thus ‘green washing’. This is where the broader meaning of sustainability comes into play. As a user of green open source software, I have the technical freedom to run it in the data center of my choice, I am not dependent on a particular provider and, accordingly, not dependent on an environment. I have the option of choosing an energy-efficient data center that meets my criteria. That is the decisive factor for Green IT.

From 2030, data centers in Europe must operate in a climate-neutral manner. So, such questions will basically also become relevant for data center operators, especially if it plays a decisive role which software is used or offered. They will have to ask themselves: Which software fits into my climate concept? How can I build my overall environment to save energy sustainably? These questions will not only continue to be important in climate issues, but will also become more relevant in financial terms, as the development of the global energy market will still be tense.

Transparency is measurable
In order to achieve the previously described, the overall system must be measurable, and the measurement data must then also be communicated transparently. The total energy consumption is basically measurable, individual applications are only measurable if only this one application is used in a certain period of time. This would also have to be compared with the energy consumption in standby mode. However, in principle, it will be possible to determine the consumption of individual applications: uploading and downloading certain data sets, collaborative editing, and so on. Measurement criteria will have to be developed for this in the future. Standards must be developed on how to implement this so that it becomes comparable.

Data maintenance
A sustainable data strategy means being able to really understand what happens to the data and when, as well as defining exactly when to delete it, for example. Strategic deletion of data then has to do with Green IT in a broader sense; everything that is deleted does not consume any storage space and thus no energy. However, this is a simple connection and only takes effect with a long-term data strategy.

However, control over one’s own data always has something to do with having an overview and keeping it up to date, which means that this cleaning up is often forgotten. However, data has the characteristic that it always becomes more and more and one only rarely deletes things. This can be problematic, as deletion is actually necessary for many regulations. Almost all NDA agreements that deal with confidentiality state that both sides have to delete the data when the business relationship is over. This is rarely checked. And how often the deletion of data really takes place in practice, including the backup tapes, is also a question that arises here. On the other hand, there are retention obligations that have to be complied with. Data maintenance is therefore a challenge for every organization that must be built into every data strategy.
For example, the ownCloud solution is able to set deletion periods. Metadata can be used to define the period of time after which certain files are to be deleted or placed in an archive. For the archives, too, time periods can be defined. For example, if documents have to be kept for 10 years, this can be marked accordingly and they will then be for 10 years. Other archives bring up the files after two years, and if no one has touched the files during this period, they can be deleted. A bit like the “Simplify your Life” method, where once a year you put everything you don’t need in a box. If you realize after a year or two that you haven’t even opened the box, you can most likely just return it to the circular economy, which is more sustainable than throwing it away for things from the household. For data, a different strategy should be used, as they should really be deleted.

As a user of green open source software, I have the technical freedom to run it in the data center of my choice, I have the option of choosing an energy-efficient data center that meets my criteria. That is the decisive factor for Green IT.

Companies should think about data strategy from the very beginning and create a system where employees work in a cloud-based and collaborative data room that is independent of any individual. This means that they do not only have their own data, which is only maintained by one employee and, if necessary, never deleted. However the different workspaces share the data room and then also tidy it up accordingly, and do so together, so that joint archiving strategies can also be developed, which then apply to the entire data room. This also includes a common version management, a common recycle bin for data, a common archiving system and cleansing metadata. This is one approach to a sustainable data strategy.

The metadata can be used to predefine whether something should be deleted straight away or resubmitted after two years. If something has been deleted, however, there is a certain amount of time to find and restore the files. After a defined period of time, it goes into an area where only the support team can restore everything, and again after a defined period of time it is finally deleted.

Currently, ownCloud is working on making data maintenance and processing “smarter” with machine learning. This means that in the future it will be possible to automate much more than today. The system will learn what exactly people frequently touch and what not at all, then report back what has not been used and for how long and, if necessary, make recommendations for archiving. The next future step for ownCloud will be to link this metadata with each other. It will be about linking the various metadata and the various pieces of information that are available and then supporting them in even more intelligent decisions that are not only based on simple policies.

Another aspect of data maintenance is version thinning. When documents are created collaboratively, there are always different versions of the documents that are created, and in most cases not all of them are needed, unless it is something to do with legal documents, where the history of creation would be relevant. Otherwise, if there is a version two, the different variations of version one can be deleted. Implementing version thinning complements active data maintenance in an organization and also supports low energy needs in the long run.

Sources
1. https://www.itsa365.de/de-de/it-sa-expo-congress
2. https://owncloud.com/infinite-scale/
3. https://okular.kde.org/de/

Holger Dyroff
COO and Managing Director at ownCloud

Friederike Zelke
Editor in chief of the cloud report
friederike.zelke@cloudical.io
Unlocking the Potential of cloud-native Software

In this article, we cover some of the most important aspects of developing and running cloud-native software, as well as what to take care about before going cloud-native. We’ll also delve into cloud-native architectures, technologies, development, and discuss the advantages of cloud-native software in general. Lastly, we’ll look at the change in mindset required for adopting cloud-native solutions.

What is cloud-native software?
But before we dive deeper into the world of cloud-native software, let’s discuss what cloud-native software is: A cloud-native application built on a cloud-native architecture, running on cloud-native infrastructures.

Cloud-native infrastructures are infrastructures meant to be breathing. This is a fundamental difference compared to traditional, static infrastructures. These infrastructures were designed to be there, to remain stable, to have a defined sizing. In comparison, cloud-native infrastructures are laid out to be volatile and adoptable: When there is more demand, more infrastructure will be provisioned dynamically. When there is a new version of an application, it will be rolled out on dynamically provisioned infrastructure. When demand slows down, infrastructure will be de-provisioned. The infrastructure adopts to the requirements of the workloads running on top of it.

The cloud-native architecture enables the application

Cloud-native software is a rapidly growing trend in the world of software development, and for good reason. This approach to building and running software is optimized for cloud computing environments and offers a number of benefits that can help businesses of all sizes and types to compete in today’s digital economy.
Cloud-native applications are designed from ground up to take advantage of the capabilities of the cloud environment. This includes support for auto-scaling, which allows applications to automatically and dynamically scale up and down in response to customer demand. It also includes support for auto-healing, which allows applications to detect and repair any errors that occur in the system.

**Advantages of cloud-native software**

One of the key advantages of cloud-native software is its ability to scale up or down as needed to meet changing business demands. With traditional software, businesses
often have to invest in expensive hardware and infrastructure to support new users or services. But with cloud-native software based upon cloud-native architectures, running on cloud-native infrastructures, businesses can easily and quickly add more resources as needed, without having to make significant upfront investments. This means that businesses can be more responsive to changing market conditions and customer needs and can better compete overall.

**Flexibility**

Another benefit of cloud-native software is its flexibility. Because it is designed to run in a cloud environment, it can be easily deployed across multiple locations and devices. This allows to take advantage of the global reach of the cloud to have workloads being available at locations where their customers reside.

It also means that businesses can more easily adapt to changing market conditions or political changes, as they can quickly and easily move their workloads. Or businesses can decide to have the application on infrastructure being controlled by them, in their own private cloud environments. Without having to change the application itself.

**Cost-Efficiency**

In addition, cloud-native software can also be more cost-efficient than traditional software, since they typically can be rolled out with a smaller footprint than traditional software. When the load increases, new infrastructures and resources can be provisioned – but other than with traditional software on traditional environments, this can happen when it is needed, not upfront. This aspect can be especially beneficial for small and medium-sized businesses, which often have limited budgets and resources.

**Open Source based**

One of the biggest issues with traditional software is that is often based on closed source components or vendor specific infrastructures. Think of software required to run on a specific operating system, such as Windows, working best only with the database server from the same vendor (SQL Server). Even before serving a single customer, businesses have to handle huge upfront payments, or when they choose a subscription-based model, high monthly fees.

Cloud-native software is built upon an entirely different ecosystem. Here, most, if not all, components of infrastructure and middleware stacks would be based upon

*Fig. 2: Instead of monolithic architectures, cloud-native applications are built using microservices*. 
open source software. Businesses would be able to choose the point in time when they want to pay the vendor – for support, or managed services, or maintenance, but not for licenses, as it is the case with traditional closed source software. Typically, open source-based infrastructures are way more inexpensive than their closed source counterparts, which increasingly often are based upon open source components themselves.

And don’t forget about the inherent security aspect associated with open source software: since the software is open source, everyone can have a look at it and identify security issues. This transparency leads to bugs and issues most often being fixed way faster than with closed source software.

By using open source software as basis for cloud-native applications and infrastructures, businesses can avoid vendor-lock, meaning they can avoid being tied to a specific vendor’s product on a specific vendor’s infrastructure or platform. This gives more freedom to change providers or software options if needed.

Cloud-native software embraces (and enforces) Change

However, to fully take advantage of the benefits of cloud-native software, organizations will need to change their mindset and approach to software development and operations, such as adopting a more agile and DevOps-oriented culture that is focused on rapid iteration and continuous delivery (fig. 1). It also implies investing in the right tools and practices to manage and monitor cloud environments, and to automate as many processes as possible.

Change in operations

In terms of operations, teams will need to be equipped to handle the dynamic nature of cloud environments, with a focus on monitoring and automation. This will require a shift away from traditional, manual processes towards more automated and cloud-native tools and practices. Operations teams need to learn about automation, about tools such as Ansible, Terraform and GIT. They will need to overcome their traditional approaches of doing things by hand – everything needs to be automated; everything needs to be executed based on versioned scripts. This is a huge change in approach and mindset.

The operations processes themselves need to be adopted as well, in fact, they need to be rethought entirely: Since cloud-native software is modular and decoupled, and since cloud-native infrastructures are based on the concept of volatility, traditional operations approaches don’t fit anymore. Operations does need to become proactive, not reactive as it was in the past. And since the sheer number of components being involved can be way bigger than within traditional environments, approaches such as Centralized Metrics and Centralized Logging need to build the foundation of an operational environments.

Which then in turn requires deeper and better integration with development teams, to ensure proper exposure of metrics and logs to the underlying platform. This means promoting collaboration and communication across teams and encouraging experimentation and learning.

It also means fostering a culture of continuous improvement and encouraging employees to embrace new technologies and ways of working.

And it also have to include adopting the right security measures to protect data and applications in the cloud and to ensure compliance with relevant regulations. This alone would be the topic for its own article and conversation.

Change in Architecture and Development

When it comes to software development, cloud-native software requires a different approach than traditional software development. Instead of monolithic architectures, cloud-native applications are built using microservices, which are small, independently deployable services, running inside containers, communicating with each other over APIs and Message Queues (fig. 2 and 3).

This allows for faster development and deployment cycles, and makes it easier to scale and update individual components of the application. But it comes with a cost: More services imply more complexity on the infrastructural level. More services also imply more chances of things breaking. And it implies more challenges when rolling things out or updating components.

This then in turn leads to not only technical-, but also mindset-based changes: To handle these complexities and challenges, services need to be written more resiliently, with a mindset of embracing the ever-changing nature of cloud environments. New frameworks and tools have to be learned, error-handling has to happen differently and more intensively, metrics and logs need to be exposed, deployment processes need to be automated end-to-end, and everything needs to be configurable from the outside world. Finally, softer and often underrated aspects of a developments cycle, such as documentation of components, need to be improved upon as well.

Change in Culture

Since cloud-native environments and software imply a multitude of technical changes, the culture within an organization needs to change as well: away from the error-preventing and static thinking of the past, embracing change and fostering a climate of error-management. Errors will happen, in fact, they should happen, since we only master complexity by knowledge and experience, which can only be gained from handling errors as well as successes.
Also, as already pointed out, communication and cross-team interaction need to change a lot: since the complexity of cloud-native applications is different on all levels when compared with traditional applications, communication and interaction need to happen constantly. Limiting conversations to silos and preventing holistic approaches, would be a fundamental flaw, leading to huge challenges in operations, development, security, and longevity of infrastructures and applications. Constant change and iterations need to be the center of culture in an organization working with and within cloud-native environments, not standstill or egoistic approaches.

**Closing thoughts**

Overall, cloud-native software can bring significant benefits to businesses and organizations, including technical capability such as increased scalability, flexibility, cost-efficiency and the benefits of open source software and avoidance vendor-locks. But to fully realize these benefits, businesses will need to change their mindset and approach to software development, operations, and interaction, adopting a more agile and DevOps-oriented culture.

This will require investing in the right tools and practices and encouraging collaboration and experimentation across teams. By embracing cloud-native software, businesses can be more responsive to changing market conditions, more cost-efficient, and better compete in today’s fast-paced digital economy.

Which will turn into a huge benefit not only technically or culturally, but also financially and regarding long-time sustainability of a business.

---

**Fig. 3: Microservices, which are small, independently deployable services, running inside containers, communicating with each other over APIs and Message Queues**

---

**Karsten Samaschke**

Founder and CEO of Cloudical

karsten.samaschke@cloudical.io
Why the idea of the 7 Digital Responsibility Goals¹ immediately convinced me is because, among the important ideas behind, they name an aspect that is implicit in all statements, definitions, commitments of and to digital sovereignty, but is all too rarely also named: responsibility.

The open source software park is now large and sometimes confusing, but some projects are so sophisticated that the software can be used for many requirements in all areas of society. And especially if you operate and host the software yourself, many things can be used confidently and securely.²

The associated buzzwords are now everywhere: open source as part of digital sovereignty guarantees transparency, is fundamentally secure, increases the ability to control, enables the ability to design, protects against unwanted dependencies. That’s all well and good and right. Nevertheless, there is a but.

**Gaining Sovereignty**

However, with all the control and design ability over and of our data, processes, accesses, tools we want to keep and retain, we also always retain responsibility for them. We are responsible for maintaining and securing our data, we have to define and establish processes, we have to install security mechanisms and keep them up to date, we have to decide which technology, which software we need, use and operate with everything that goes with it.

For any company or public institution, the decision to use open source has to be a conscious one. It is not possible to go to a website, click on a few buttons and start working, as is possible with some proprietary software if one has deposited payment options. Rarely are there pre-integrated software packages, which means that a user of open source must know in advance which applications are needed for which requirements. There are projects and groups of companies working on such offerings like the Phoenix-Project in Germany³. But this market is growing slowly.

In my eyes, this required knowledge is an advantage, it is always good to know what I want to do and how I want to implement it. In the best case, this is already linked to processes, because the tools must not only meet the formal requirements, but also fit the way of working. If I know what I am looking for, I can also better evaluate the search results and find suitable solutions. And since there is no website that offers all available open source tools and applications with names, functionalities and integration possibilities, the search is definitely easier the more precisely I know what I want.

So, when I have found the right applications, I can download software packages. And then? Any IT team that enjoys modern technologies and is interested in challenges will be happy about this situation: How much memory is needed not only to store the software but also to use and operate it? Is an additional database needed? Does the database talk to the application? Do different applications talk to each other when I want them to? Can I integrate different applications with each other? Do I have to search for and install updates myself? Do updates change the configuration? Who can use the applications? Who needs access? What additional hardware is needed? What can be automated? ... And what if there is no IT team available, whether they enjoy it or not?

And in every case, even if something fails, I bear the responsibility. And all I wanted to do was work confidently.

At this point I would like to say that I think it is basically a very good thing to bear responsibility for one’s own actions, processes and things. It should just not become overwhelming.

I can understand that companies with small IT teams or the public sector, which often have administrators...
rather than engineers in their IT departments, are reluctant to switch completely to open source tools. The knowledge needed to build and operate them in a performant, productive, fail-safe, secure, well-integrated and customized way is huge and still not widespread enough. And if all this is to be operated in one’s own data center, the effort to procure and maintain the hardware components is considerably higher. And now I have to bring in the buzzword shortage of skilled workers.

But nevertheless, open source should and is being used more and more. Which is good because everything positive that is said about it is true. So how is that possible? Above all, without losing control.

This is mainly thanks to the communities of the individual projects, which in recent years have also increasingly focused on usability, productivity, integrability, security and up-to-dateness. New projects and communities have also been founded that take up and solve some of the problems mentioned here by building software packages in which well-integrated, sensibly coordinated open source tools and projects can be combined and used together: VanillaStack, SCS, nextcloud, owncloud, ... These software packages nevertheless remain completely open source, so they facilitate use but lose none of the advantages.

The operational challenges remain here as well, but it makes it easier to get started, and a community has also formed around this topic that wants to open source operational practices and processes: open operations.4

Sharing Responsibility

These are steps into the right direction. However, this still does not really help organizations without open source experts. And these organizations should not be denied access to control, sovereignty and creative possibilities. That is why a well-functioning service system has developed around the open source ecosystem. Companies have specialized in building open source environments and toolsets for organizations and adapting them to their needs. The open source world is so complex that not only individual software tools are offered, but also virtualized environments, operating systems, platforms, databases, networks - the entire IT stack can be (and in most cases already is) covered with open source technologies. Along with the companies that take care of the “toolbox”, hosting providers have also evolved to provide the servers and environments on demand and keep the tool-park running. In fact, the offer in this sector is now so broad that every need can be met. But also as with software, an organization should know beforehand what it needs and wants! There are hosters who only provide the hardware, which in many cases is a good, sensible offer, but all software operation remains with the customer. And there are cloud service providers who take care of everything - and still implement everything with open source technologies.

Let’s summarize briefly: A great deal can now be implemented with very good, mature open source software, and
the variety of possible applications, including infrastructure and platform software, is enormous. This makes it possible for all organizations to base their IT on open source, which is for me the only basis for digital sovereignty. With the right partners and providers, open source can also be stored and operated in compliance with data protection laws.

Some questions may still remain open: Do I retain control over my work if I store my data in a third-party data center? Am I sovereign if an external provider runs my software (which they could even do in the organization’s own data center if, for example, a federal state has a state data center but not the engineers to run a cloud solution, for example)? Am I only digitally sovereign if I do everything myself?

Here I come back to responsibility: If I am so overwhelmed with the technical challenges of running software and hardware myself that it doesn’t work, I am definitely not sovereign and I am not living up to the responsibility towards my organization either. Sovereign, responsible action in the area of technology is to find good, secure solutions that fit my needs. And this solution can be that I work with partners who understand my needs and work with me to find the appropriate implementation, who offer standardized open source environments and applications in an open source operating system, with whom I can work in confidence, who allow me access and control over my data at all times and who communicate transparently. But it must also be possible to switch to other partners at any time if the relationship of trust or the service is no longer appropriate. It is responsible to decide how much control and practical responsibility I want to relinquish. And if I use secure open source services in order to be able to use the latest technology, I still remain sovereign, because I can revise this decision and retain the control and responsibility that I can bear.

Responsibility is an essential part of sovereignty.

Sources:
1. https://identityvalley.org/drg
2. https://the-report.cloud/use-your-freedom-of-choice
3. https://www.dphoenixsuite.de/
4. https://openoperations.org/

Friederike Zelke
Editor in chief of the cloud report
friederike.zelke@cloudical.io
Measuring Open Source Project Health

In recent years, we learned the hard way that open source software requires substantial maintenance work. Just because the source code is open to inspection does not mean anyone is looking. While not the first incident, Heartbleed raised the awareness that even widely used open source software needs to be maintained. Since then, supply chain attacks have become commonplace. Software security is getting much attention, with legislators in the US and Europe drafting laws to improve cybersecurity that will affect all software, including open source software. We must remember the lesson we already learned with Heartbleed: We need healthy open source projects if we want high-quality and secure software.

The partnership of OpenInfra Foundation and Bitergia

To make the following discussion tangible and more interesting, I will use the OpenInfra Foundation as an example. I chose the OpenInfra Foundation because Bitergia is their Official Metrics Partner, which gives me relevant experience and insights. We shared these examples at the OpenInfra Summit 2022 in Berlin, which is why I was invited to write this article. Following, we will explore what it takes to have a healthy open source project and what metrics we can use to understand project health.

For Context, allow me to introduce Bitergia and the OpenInfra Foundation. Bitergia is an open source company specializing in Software Development Analytics. 10 years ago, the founders commercialized their research into how to analyze software development and released their tooling as open source. The tools are called GrimoireLab, a project within the CHAOSS Community. CHAOSS is short for Community Health Analytics for Open Source Software and a community with the Linux Foundation. It is a community of practice for open source professionals, researchers, and anyone interested in the topic. To be upfront, I am a co-founder and board member of CHAOSS. I will mention CHAOSS again later because it is the place to find resources, guidance, and tools for analyzing open source project health.

The OpenInfra Foundation emerged from the now 12-year-old OpenStack project. OpenStack is a cloud platform developed by over 450 companies and today runs on 40+ million cores in over 300 data centers. That is more data centers than the largest proprietary cloud provider, AWS, which has about 125 data centers. The OpenInfra Foundation hosts other projects too, including Kata Containers, StarlingX, and Zuul. It has two remarkable things in its approach to stewarding open source projects. First,
the projects are provided with a software forge consisting of only open source tools (see open-dev.org) that projects can choose to use and avoid proprietary services. Second, OpenInfra believes in a philosophy of open source with The Four Opens as guiding principles:

- Open Source
- Open Design
- Open Development
- Open Community

Bitergia and the OpenInfra Foundation share a passion for open source and therefore partnered up to build healthy projects. After a brief definition of Project Health, we will look at examples.

We need healthy open source projects if we want high-quality and secure software.

Defining Open Source Project Health

Open Source Project Health is a project’s potential to continue producing quality software. We may also refer to it as the health of the open source community or the project’s sustainability. I researched this for my Ph.D. and confirmed what other researchers also found: It takes three things to have a healthy project: Quality Code, Sufficient Resources, and an Active Community.

Quality Code is the desired outcome of software development. It includes best practices for source code to be well structured, human-readable, sufficiently documented, and free of bugs. Given the complexity of modern software, these may not be fully achieved, but we can have processes and policies in place to help projects get closer to this ideal. For example, OpenInfra strongly recommends practices such as extensive code review, automated testing, and source code linting, just to name a few.

Sufficient Resources will be different for each open source project. Some projects may only need a source code repository and an issue tracker, which are provided for free through services like GitHub, GitLab, or Gitee. A project that produces software for Mainframes or the public power grid will need special hardware to test the software. In the example of OpenInfra, projects are provided with a suite of open source tools that enable collaboration. The foundation works with its members for additional resources if the projects need them. For example, infrastructure donors are companies running OpenStack clouds, donating cloud resources to the OpenStack project infrastructure. Those resources are mostly used in the automated testing framework to support OpenStack development efforts.

Active Community is the result of people working together on the software. A project without activity is not updated, not adapted to the changing environments, and not helping users to resolve issues. Ideally, an active community has many types of contributions, including code and non-code contributions. It also matters who the contributors are.

This is the topic for the remainder of the article, exploring four aspects: Contributions, People, Organizations, and Inclusion. I will look at each aspect and the metrics available to understand these.

How to measure community activity

Contributions are the building blocks of work in an open source project. Code contributions advance the source code and are logged in a version control system such as git. The change history makes it easy to count and measure code contributions which has historically led to a bias to recognize this type of contribution over others. Other types of contributions include code reviews, which are essential to maintaining the source code’s quality and educating other contributors. Beyond these, bug reporting and triaging are contributions logged in an issue tracker or similar system. Many projects also have communication channels, such as mailing lists or chat platforms like Slack, which have a communication archive. In contrast, some contributions don’t get tracked at all but are essential to the health of a project. For example, organizing
events, giving talks about the technology at a conference, socializing with other members, or managing finances. The OpenInfra Foundation is the fiscal host of its projects, manages background activities, organizes events, and ensures the environment is right. All activity levels across logged contribution types are captured and visualized in a Bitergia Analytics Platform. This dashboard shows trends of engagement and helps to understand how projects are doing over time.

People make contributions and self-select to be members of the community. A concern for healthy projects is that the knowledge required to maintain and advance the software is spread across multiple people. As a negative example, a project with only one contributor is entirely dependent on that person, and they may stop maintaining the project at any time. In a bad case, that person can break the software on purpose, and there is no check from others (yes, this has happened). The Bus Factor is a key metric that shows how dependent a project is on one, a few, or many people. The OpenInfra Foundation keeps track of this metric through the Bitergia Analytics Platform and, if needed, works with the project to highlight areas that need additional people and helps onboard new contributors. If a project does not have enough interest, it may also be officially discontinued, which is an honest and fair move that signals to users that they need to find an alternative.

Organizations play an essential part in the health of open source projects as users, employers, and sponsors. Organizations benefit from open source software with reduced time to market, lower costs to build software, and lower license fees. With this vested interest in open source software, organizations allow their developers to contribute back, maintain, and advance the software. This also gives organizations influence over the project’s direction and helps align the internal roadmap with the work in the open source project. The metaphorical elephant in the room is that when an organization de-prioritizes a project, that project will be in jeopardy if
that one organization employs all project members. Bitergia created the Elephant Factor, which works the same as the Bus Factor, but instead of people, it looks for organizations. With a focus on getting more organizations involved in projects, OpenInfra has diversified the organizations contributing to the Kata Containers project⁷ (fig. 1).

Inclusion refers to a project’s approach to enabling and empowering contributors irrespective of their background or identity. This is a core concern of the Open Community principle. At a broader scope, inclusion is a major concern for open source because surveys show that a majority of open source contributors are white men. Investigations have shown that a toxic culture in projects is a frequent reason that actively excludes minorities or makes for a very unwelcoming environment. Unfortunately, the homogeneous contributor base also self-perpetuates through the unconscious bias in how software is written, workflows operate, and documentation is structured. It takes a concerted effort to identify the issues and to create a welcoming environment. The OpenInfra Foundation has a dedicated Diversity Working Group that works with the Foundation’s Board of Directors to incorporate diversity policies and create programs that reduce barriers and create an inclusive and welcoming culture. Every few years, the Foundation surveys its community to measure progress and identify areas of improvement.

Advice and Resources

We are at the end of the article. We discussed open source project health and saw examples from the OpenInfra Foundation of how that can be measured and promoted. I want to leave you with some advice and resources.

First, if you are thinking about measuring your open source project’s health but are concerned with the complexity, do not fret; start with the easy things to measure and answer some basic questions. Those answers will spawn new questions, and you can build on your experience to progress your metrics journey.

Second, you are in good company. Check out the CHAOSS Project to find resources, metric definitions, open source software for dashboards, and a community of practice.

Third, remember that the metrics are in service of the community. Be open and honest with the community when introducing metrics to stave off concerns early. For example, OpenInfra and Bitergia met with the community to review metrics. At the OpenInfra Summit 2022, we organized a Metrics Corner to showcase the metrics dashboards and discuss them with the project’s contributors. This was a huge success and we look forward to hosting the next Metric Corner at the OpenInfra Summit 2023 in Vancouver. Join us June 13-15!⁹

Fourth, collect metrics early to establish a baseline. The baseline is important to see changes in your community and to see if policy changes have the desired effect. A word of caution, do not try to benchmark against other, even similar, communities because each community has different ways of working, is in a different context, uses different tools, or has other reasons to produce metrics that are hard to compare. Each community is unique, and understanding its project health is a journey that can start with metrics but for sure requires a conversation with community members that can validate what is actually happening in the community.

Sources:

5. https://openinfra.dev/
6. https://.opendev.org/
7. https://openinfra.dev/members/#infrastructure
8. https://katacontainers.io/

Georg J.P. Link
georglink@bitergia.com

Georg is an Open Source Strategist. Georg’s mission is to make open source more professional in its use of community metrics and analytics. Georg co-founded the Linux Foundation CHAOSS Project to advance analytics and metrics for open source project health. Georg has an MBA and a Ph.D. in Information Technology. As the Director of Sales at Bitergia, Georg helps organizations and communities with adopting metrics and making open source more sustainable. In his spare time, Georg enjoys playing board games, Anno 1800, reading fiction, and hot-air ballooning.
The Container Days have become a permanent fixture in the event calendars of all container enthusiasts. Alongside KubeCons, they are the largest international event on Kubernetes and container technology. In 2021, the focus of the participants was more on the virtual event, because only 300 participants were allowed on site due to the Corona regulations. In 2022 it was the other way round, with around 800 open source enthusiasts coming to Hamburg. The joy of seeing each other again was huge and the thirst for knowledge enormous.

The event was a mixture of community meeting, conference and family reunion. And there was something for everyone in the exhibition as well as in the sessions: I met two employees of the company Firedrill, for example, who offer virtual fire protection exercises and run their services in a self-hosted Kubernetes clusters. They were there to learn more, to be able to ask questions and to meet the real experts - because they don’t have access to the communities (yet). And they were thrilled, along with me, that there were introductory sessions at such an established event as a matter of course. There were overview sessions, but also first steps and best practice sessions. But there were also advanced sessions for the “real experts”, where even they could learn something new, for example the guys from Cloudpunks who use the Container Days to learn about the container landscape in Germany but also to introduce themselves to the community. It is precisely this mixture that makes the event relevant for the entire container open source community. Of course, the communities are about the further development of the projects, but it must also always be about inspiring new contributors and being confronted with difficulties or ambiguities through questions from users. That’s why the broad offer to learn, but also to use the time for beer and exchange is so important!

In the meantime, Kubernetes has matured so much that many of the sessions dealt with concrete use cases or challenges such as edge computing or multi-clusters. A major focus was on security at all levels from infrastructure to supply chain, certificates to multi-cloud. Another focus was on operations, especially GitOps and monitoring. Fittingly, there was also a lot of talks about cloud-native development.

Especially in the area of git, I learned something new and great: There is a real alternative to GitHub and GitLab: Gitea! Check it out, try it and become part of the community - this is part of the path to sovereign open source projects.

Even though many participants were in Hamburg, the event remained hybrid, all sessions were streamed and recorded and are all available online now too via the Container Days YouTube channel! Here are my personal recommendations (please bear in mind that I am not a real expert):

1. About Monitoring and Observability: Michael Friedrich: From Monitoring to Observability and Matthias Haeussler: What’s going...
on in my cluster? (fig. 1)\textsuperscript{6}

- About GitOps: Jimmi Dyson: Kubernetes as a Service with GitOps\textsuperscript{7} and Lian Li: GitOps for the people\textsuperscript{8}

- About Cloud Security: Mike Milner: Cloud Security from Scratch\textsuperscript{9} and Viktor Gamov: Zero Trust Security with Service Mesh\textsuperscript{10}

- About cloud-native Development: Jim Sheldon: Container-native pipelines with Drone CI\textsuperscript{11} and Lukonda Mwila: How To Make Your Application Developers Fall In Love with Kubernetes and Cloud Native Applications\textsuperscript{12}

This year the event took place in a new location: Kampnagel in Hamburg – theatre, ballet and event location\textsuperscript{13}. But with a clear message: No Racism, no war – inclusion, openness, sustainability. For me this so fitting for an open source community event where these topics are important or should be important! For me personally, the catering was also great, everything vegan or vegetarian and 100% fresh – very fitting, very tasty.

All in all, I only can say: Come to Hamburg and experience this community for yourself. Get involved and make container technology come alive with everything that goes with it!

Sources:
1. https://firedrill.de/
2. https://www.cloudpunks.de/#about
4. https://www.youtube.com/@ContainerDays
5. https://www.youtube.com/watch?v=UlkhScJ-ZsD8&list=PLHHKcdBlprMdMzUXZ6ho0OPTik-TamLwa&index=3
6. https://www.youtube.com/watch?v=SbggN-GYSsg&list=PLHHKcdBlprMdMzUXZ6ho0OPTik-TamLwa&index=49
7. https://www.youtube.com/watch?v=IS4Z8x4hs&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=60
8. https://www.youtube.com/watch?v=jepLx_bY9vM&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=42
9. https://www.youtube.com/watch?v=SZoIzWc9II&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=22
10. https://www.youtube.com/watch?v=tmKRUyjAcow&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=24
11. https://www.youtube.com/watch?v=PXMc3rU7N4&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=4
12. https://www.youtube.com/watch?v=LjzrqFvLzs&list=PLHHKcdBlprMdMzUXZ6ho0OPTikTamLwa&index=20&t=2s
13. https://kampnagel.de/
The Univention Summit in Bremen traditionally opens the year of open source companies in Germany. They discuss with representatives of the public administration how they can jointly make Germany digitally sovereign and advance the digital transformation. Peter Ganten, Managing Director of Univention and Chairman of the OSBA, summarized the current situation in his keynote and thus also determined the goal of the summit:

The crises of the last few years, but above all the war in Ukraine, have massively questioned the sovereignty of European states and are forcing them to think about security and global dependencies in a completely new way. For the European states, this applies not only to the area of energy, but also to the entire IT landscape, in which there are just as many global dependencies, and which is similarly relevant to society. And it is also similarly vulnerable to oil pipelines, as the global network depends on cables in the oceans that could quickly become targets of attacks. This awareness of vulnerability must also lead in the area of IT to securing it, making it sustainable and transforming it into an open, reliably available basic infrastructure that is then also no longer in the hands of a few.

Overall, the last year has shown that the digital market worldwide is suffering just as much from the global crisis as other parts of the market, with most of the large cloud providers recording losses last year and having to lay off many employees, for example. However, the further development of open source software and its products has grown in comparison and become more established.

In the public administration in Germany, however, there is now a growing awareness that digital sovereignty is an essential building block of national sovereignty. The digital strategy and also the founding of ZenDiS (Zentrum für Digitale Souveränität – Centre for Digital Sovereignty) show that it has been understood that it is important to be able to control what happens with the IT system, who owns it, who has access to it and where data may
flow to. Especially when it comes to public administration, when it comes to secret data or data of citizens that need to be protected, it is necessary to have comprehensive knowledge of the systems. At the same time, however, we have learned in recent years that Germany in particular is lagging behind many of its European neighbours in the digitization of administrative processes and the interaction between administration and citizens, but also in the digitization of the education system, and that the people acting in Germany must be in a position to take the reins of action back into their own hands, to be able to shape things themselves.

With this in mind, important players from open source and the public sector have come together in Bremen, because digital sovereignty can only be achieved together, only if there is transparent communication and a diverse open group of active participants can new dependencies be avoided.

Many projects in Germany prove that the approach of relying on open source programs works excellently and that there are partner companies that support this approach, develop it further and use their expertise for public administration. The great example of the Summit is the development of the sovereign workplace (Fig. 1), which is being developed by dataport on behalf of the federal government and with the involvement of many open source products. In fact, a preliminary version of the complex software is already available, the dPhoenix suite, which offers the most common office applications integrated with each other. In the long term, this should be able to run either self-hosted in e.g. state-owned data centers or operated by digital sovereign hosting partners. The sovereign workplace will also have connections to the specialized procedures necessary in public administration and thus hopefully fulfill all the requirements of a digitally sovereign administration.

Overall, the Summit showed me that such events are necessary to enable an open exchange between the administration and developing companies, even away from concrete projects or questions. The modern IT landscape requires so much expertise that the administration or even a single company in Germany or Europe cannot provide the necessary services. This can only be achieved together with openness, transparency, expertise, and open source.

Sources:
1. https://www.dphoenixsuite.de/

**Fig. 1**: The sovereign workspace includes a whole office package and special software for administration.
As we move forward in a digital world, it is difficult to ignore the amount of data, frequently without consideration for the consumer or subsequent deliberation on the part of the consumer, large tech companies are compiling. While we have a tendency to distrust and malign big data and the organizations that harvest data, the obvious answer may be to stop using these services or technologies as a result may have wider-ranging implications which make an abrupt halt of these services less apparent. If used ethically, data has an important role to play in society. During the recent Covid pandemic, data and data sharing were critical to diagnosing the virus at a much earlier stage in the pandemic. Conversely, when companies do not adhere to legislation or subsequently employ practices that violate a common code of ethics, whether due to blatant disregard or because the development practices of the organization do not provide adequate observability and control, impending action should be required. The question is then how can we as technologists ensure that our systems are adequately and sustainably designed to allow for such control and elevated security, particularly as we embark into ever increasingly complex technological landscape.

As the industry looks to optimize infrastructure, one key way to do this is through reusing services, such as databases and logic across resources. With the advancement of 5G and the increase in the use of Internet of Things devices, the implications of sharing services are becoming increasingly complicated and require that security controls, that were once sufficient, such as using an IP address to root identity, are no longer adequately secure nor do they respond to the need for the sustainable deployment and development.

Identity for Machines, too
As technology evolves and more of our day-to-day life becomes digital, there is an increasing concern for the safeguard of digitally bound assets, such as bank accounts, that rely on human identity for access. For human authentication, we rely on strong authentication methods linked to what we know, what we have, or who we are, and in the case of multifactor authentication, a
combination thereof. Frequently, this authentication will be federated across multiple systems, resulting in a Single Sign On (SSO).

The use of SSO is fairly well established within organizations and the benefits are understood. What is becoming increasingly apparent is that with the adoption of shared services and Service Oriented Architecture, that identity must also be translated to machines in order to ensure an adequate security posture and ultimately permit more control and sustainable development. It’s easy to speculate why application identity federation hasn’t become more prominent, whether it be the static nature of infrastructure, until more recently or the pervasive use of monolithic architectures or a combination; regardless, the current state of application identity leaves us with a fragmented, decentralized, highly manual approach.

While most cloud providers have native identities for their resources, unifying these into a single, auditable, controlled workflow without a centralized brokerage is complex, not human readable, and not scalable due to the complexity required for automation. By leveraging a brokerage mechanism, we can guarantee a unified workflow over all the required application layers, from network to application across multiple resources.

**Why Does Machine Identity Matter?**

Examining the history of identity controls and the associated security controls, the industry has heavily relied on IP addresses. In a world dominated by physical servers, monolithic applications, and predominantly physical controls, a trust model tied to a static IP address and inherited trust provided the required security assurance. As the adoption of Cloud continues to increase, the controls that we previously relied on, are no longer adequate. This is due to a shift from a primarily physical infrastructure to a largely logic based infrastructure and the imminent fault domain associated with resources. As a result, the ability to implement identity and security controls in a scalable, secure manner is no longer tenable.

Looking at the principles surrounding the 2nd Industrial Revolution, the concepts can eas-
ily be adapted to the IT landscape, particularly adopting a factory system and division of labor. In the multi-hybrid cloud world, the difficulty appears that we are attempting to industrialize and homogenize highly customized platforms. If we want to transfer the benefits of industrialization to the digital world, we need to consider what that assembly line looks like on each CSP or Data Center level, and also how we industrialize those intersections. Just as the standardization through assembly lines permitted easier maintenance of machines due to homogenization and the ability to more easily exchange parts, the use of machine identity, when adequately brokered, provides a similar flexibility by ensuring that services are adequately federated and authenticated by mapping identities to the subsequent resources.

How Can Identity Scale?

All of the major Cloud Service Providers provide their own flavor of identity, networks have another set of identities, applications their own types of identities, and users, as well. Generally, these identities are bound by highly customized solutions on an application by application basis. As seen in industrialization, by creating a single workflow surrounding Identity, the possibility of engineering the intersection of these resources becomes highly automatble and repeatable, which in turn makes the process and the overall ecosystem more sustainable in the event of change, whether required or chosen, and more secure as all resources are mutually authenticated based on their identities.

By leveraging identity brokerage, we have more flexibility, secure authentication on all levels of our application stack, and better observability across systems. Because brokering allows a many-to-many relationship, the operational complexity is reduced as a result of managing fewer one-to-one relationships, resulting in added value due to more focus on declarative, relational authentication and authorization.

Should We Broaden Our Identity Scope?

While Gaia-X focuses largely on the human implication and controls of software engineering and data governance, the importance of the machine identity and the subsequent control required to adequately maintain those identities should not be overlooked. Given the complexity and volume of connected devices in an ever increasingly dynamic environment, the manner in which we control these machines, permit access becomes more critical. To fully secure our infrastructure, and as a result, ensure data security in a manner that we can force change across entire ecosystems, adequately managing machine identity and their access to resources is a fundamental step.

Sources:
vanilla

the all-inclusive
easy to install
open source
cloud stack
Join Us for the ContainerDay Security – Rising trends, best practices & more!

ContainerDay Security is a special one-day community event taking place at CreativSpace in Hamburg on March 8th, 2023! It offers plenty of networking and knowledge sharing opportunities among members of the Kubernauts community. You can expect expert sessions and roundtables about cloud native security best practices and use cases. The event’s location is an open loft space where you can connect with new people in a welcoming environment.

ContainerDay Security is the perfect event for anyone interested in keeping up with the latest trends and developments in the Security field. The agenda includes 8 engaging talks, one hands-on workshop and 4 discussion groups. You will learn how to run your Kubernetes-based workloads securely, hear about actual attack scenarios, conduct container security scanning, get an introduction of several security tools, and receive further helpful security insights and best practices.

Join us on March 8th, grow your network and learn from cloud native experts. Book your spot today!

www.containerdays.io/security-day/