ORIGINAL PAPER

No Algorithmization Without Representation: Pilot Study on Regulatory Experiments in an Exploratory Sandbox

Denisa Reshef Kera[1] · **František Kalvas**[2]

Abstract

The exploratory sandbox for blockchain services, Lithopy, provided an experimental alternative to the aspirational frameworks and guidelines regulating algorithmic services ex post or ex ante. To understand the possibilities and limits of this experimental approach, we compared the regulatory expectations in the sandbox with the real-life decisions about an "actual" intrusive service: contact tracing application. We gathered feedback on hypothetical and real intrusive services from a group of 59 participants before and during the first and second waves of the COVID-19 pandemic in the Czech Republic (January, June 2020, and April 2021). Participants expressed support for interventions based on an independent rather than government oversight that increases participation and representation. Instead of reducing the regulations to code or insisting on strong regulations over the code, participants demanded hybrid combinations of code and regulations. We discuss this as a demand for "no algorithmization without representation." The intrusive services act as new algorithmic "territories," where the "data" settlers must redefine their sovereignty and agency on new grounds. They refuse to rely upon the existing institutions and promises of governance by design and seek tools that enable engagement in the full cycle of the design, implementation, and evaluation of the services. The sandboxes provide an environment that bridges the democratic deficit in the design of algorithmic services and their regulations.

✉ Denisa Reshef Kera
   denisa.reshef@um.edu.mt

   František Kalvas
   kalvas@kss.zcu.cz

1   Centre for Distributed Ledger Technologies, University of Malta, Msida, Malta

2   University of West Bohemia in Pilsen: Zapadoceska Univerzita V Plzni, Pilsen, Czech Republic

1 Introduction

Regulatory sandboxes in the FinTech and LegalTech domains have pioneered an experimental approach to regulating algorithmic services that support participatory engagements of institutional stakeholders (Alaassar et al., 2020; Gromova & Ivanc, 2020; Madir et al., 2019). We used the model of live testing under supervision to accommodate exploratory goals that involve a variety of participants in the full cycle of design, implementation, and regulation of intrusive blockchain services. After experiencing a "polite" surveillance under intrusive satellite and blockchain services, the participants negotiated the relations between code, values, and regulations on a concrete case of a biased and discriminatory code.

In May 2020, we saw an opportunity to extend the original focus of the research that followed decisions on functional, although hypothetical, algorithmic services. The COVID-19 pandemic enabled us to compare the sandbox-based decisions on smart contracts with real-life decisions on contact tracing. We piloted an "ad hoc" longitudinal study that followed the attitudes toward contact tracing during the first and second waves of the pandemic in the Czech Republic (June 2020 and April 2021). Because of the increasing non-response in the cohort of 59 nonrandom participants, the findings are not representative but offer material for discussing the opportunities and limits of experimental policy research.

We claim that direct participation in the design and regulation of new services supports the trust in algorithmic services as a matter of agency and political representation. Instead of "ethical" algorithms defined by normative frameworks, guidelines, and recommendations, the exploratory sandbox offers an experimental environment for defining the stakes and power over the code and regulations.

The preliminary results from the exploratory sandbox Lithopy[1] (Kera, 2021), together with the surveys about contact tracing in this paper, suggest support for independence rather than government oversight and for hybrid, experimental regulations. Instead of reducing the regulations to automated code or insisting on strong regulations over the code, the participants with different levels of knowledge and agendas seem to support the experimental interventions they experienced in the sandbox. In our earlier work, we discussed design and policy experimentation as a possibility of "regulation through dissonance" (Reshef Kera, 2020b). It describes a process of negotiations in the sandbox, during which the participants with various motivations, knowledge, and attitudes test their comfort levels around the technology before making any decisions.

To summarize participants' attitudes to the regulation of intrusive algorithmic services, we use a famous slogan of the American Revolution and interpret their interventions as demand for "no algorithmization without representation." The intrusive services are like new algorithmic "territories" where, instead of relying upon existing institutions and universal values, including ethical algorithms, citizens explore their agency over both the regulation and the code.

---

[1] "Future of RegTech: How to Regulate Algorithms?" Tableau visualization: https://tiny.cc/lithopy

In the exploratory sandbox, participants redefine their sovereignty and agency as new "data" settlers that tentatively combine the on-chain algorithms with off-chain values, norms, and institutions. Instead of a universal model for regulating (with) algorithms, the sandbox supports situated, agonistic, and "good enough" solutions that remain hybrid and open for further modifications, engagement, and deliberation. The interventions support friction that slows down the technology and regulations to increase participation in a particular context.

2 Background and Context of the Study

There is a democratic deficit in the design and implementation of algorithmic services (Danaher et al., 2017; De Filippi & Hassan, 2018; Introna, 2016; Shorey & Howard, 2016; Susskind, 2018), such as content filters, recommendation systems, various implementations of AI and ML in autonomous vehicles and robots, blockchain consensus mechanisms, and smart contracts. It leads to a clash between the legitimacy of off-line and non-digital institutions and norms with the efficiency of the algorithms and technical protocols. For example, the various risk assessment and prediction algorithms in justice, education, or health make epistemologically correct decisions that address no ethical and political values and aspirations (Alexander, 1992; Begby, 2021; Hajian et al., 2016; Johnson, 2020). They guarantee efficient, fast, and impartial decisions but lead to structural discrimination, stereotyping, and other excesses based on the inherited patterns in the ML training data and weighting. Similarly, with blockchain algorithms, the decentralized protocols promise an automated distribution of power but lead to various scams and forks that reveal the extreme inequality in the governance of the source code (Aziz n.d.; Vili Lehdonvirta, 2016). In summary, the processes of designing and implementing algorithmic services sacrifice the expectations of fairness, equity, or legitimacy for disruption and novelty.

To mitigate the excesses of this clash between efficiency and legitimacy, scholars proposed various normative frameworks, recommendations, and guidelines. They tried to resolve the issue of arbitrary but "hard-coded" decisions expressed in code either with ex post (via regulation) or ex ante (via design) interventions. Aspirational frameworks for AI ethics (Floridi et al., 2018; Winfield & Jirotka, 2018) or privacy-by-design (Cavoukian, 2009) and society-in-the-loop (Rahwan, 2018) proposals reconcile the social and cultural norms with code by stating the values or norms in advance and enacting them in the design process.

The guidelines, frameworks, and recommendations map the values–such as transparency or oversight–into the design constraints (Introna, 2016; Kroll et al., 2017; Lee et al., 2019, Burke, 2019; Shneiderman, 2016) or demand ex post changes to the services to meet the regulatory expectations (Fitsilis, 2019; Goodman & Flaxman, 2017; Hildebrandt, 2018; Yeung, 2018). By reducing the legitimacy and ethics to ex post or ex ante interventions, the frameworks overlook the issue of political participation and representation as a means of legitimacy in any social and (as we claim) technological process.

The main problem with ethical and social commitments in the algorithmic frameworks, principles, and guidelines is that they often claim universal values that conceal the vested interests of the stakeholders that define them. They also ignore the unequal power relations over the code of the emerging infrastructure and regulatory processes. To solve the issues with the democratic deficit, we need to make the potential clashes of interests and stakes transparent and even productive on every level of the design and policy. Instead of placing the pre-defined values, regulations, and aspirations outside the design and implementation phase, we need to connect them with the work on the code through participatory, experimental, and iterative interventions.

3 Regulatory and Exploratory Sandboxes

Inspired by the present regulatory sandboxes in the FinTech and LegalTech domains that advocate the experimental deployment of algorithmic services under supervision (Madir et al., 2019; Financial Conduct Authority & Authority, 2015; Alaassar et al., 2020), we designed an environment described as an exploratory sandbox. We adapted some of the functions of a typical regulatory sandbox to support more public engagement and deliberation, supporting the ideals of open government (Chwalisz, 2020). While a typical regulatory sandbox expands the notion of a testing environment from software development and computer security to include various regulatory institutions, the exploratory sandbox involves the general public. Both use a sample of participants and stakeholders, on which they follow the effect of the new services over a limited period to propose changes and reiterate. In the exploratory sandbox, the supervised interventions include also individuals and groups without a clear identity or agendas.

To experiment with the regulations over intrusive algorithmic services, we designed the sandbox to support changes in a simulated imaginary smart village of Lithopy. We tested it with a group of 59 participants belonging to the W.E.I.R.D. (Western, educated, industrialized, rich, and democratic) demographic (Henrich et al., 2010) at a Czech university in January 2020 before the start of the COVID-19 pandemic. The participants in the workshops reacted to functional, though biased, smart contracts that caused discrimination in the village. Subsequently, they responded in a survey in which they weighed the different regulatory options that we discussed in our earlier work (Kera, 2021).

The exploratory sandbox supported a direct experience of negotiating the code and regulations that provided the initial data on the importance of participation and representation in the design and regulation of algorithmic services (Ibid). It helped the participants to understand the architecture of the platform (permissioned blockchain and consensus mechanism), the design of the blockchain applications (smart contracts with readable JavaScript code), and the interface (dashboard connecting the data flows). While experiencing the technology, participants also familiarized themselves with the regulatory tools (audits, codes of conduct, soft and hard laws, and ISO norms) to make individual decisions on how to prevent discriminatory and biased codes. We visualized the preferred strategies as a tableau story about the

"Future of RegTech: How to Regulate Algorithms"[2] and later compared them with the data from the surveys on contact tracing published as a series of reports[3] and also datasets on Zenodo.[4]

The iterative work on the efficiency but also the legitimacy of the regulatory and design decisions enabled the participants to connect their work on the code with negotiations of values and norms. Instead of the teleology of better Reg or Tech (), the exploratory sandbox thus supported pragmatic and experimental engagements (Reshef Kera et al., 2019; Reshef Kera, 2020d) that responded to the immediate challenge of a discriminatory code in one intrusive algorithmic service. To determine how the decisions in the sandbox relate to real-world situations, we used a longitudinal pilot study on contact tracing as exemplary pervasive surveillance and automation over Bluetooth. We followed the decisions and regulatory expectations about a COVID-19 contact tracing application in surveys conducted in June 2020 and April 2021.

The sandbox participants expressed similar attitudes and regulatory expectations about hypothetical services in Lithopy and COVID-19 contact tracing in 2020. While there were differences between the two algorithmic services in terms of their functions and contexts, they nonetheless shared a similar challenge: the service claimed norms and values without engaging in the process of checking and testing if they hold. While they supported institutional engagements in the oversight, they lacked direct participation and political representation in the design and policy processes.

4 Hypothetical and Real Algorithmic Services

The exploratory sandbox Lithopy extended the model of stakeholder engagement and institutional oversight that defines regulatory sandboxes (Bromberg et al., 2018; Fan, 2017; Herrera & Vadillo, 2018). Instead of involving stakeholders that represent some pre-defined groups and agendas, the sandbox involved participants with open agendas and interests. The institutions representing the public interests set up most of the regulatory sandboxes, but anyone interested in design and regulation of near future services can operate an exploratory sandbox. Instead of protecting any pre-defined values and interests, the sandbox supports anyone to define their regulatory expectations, values, and visions on the go while confronting a problematic service.

Most of the participants in the Lithopy sandbox negotiated their regulations and code as individuals (Reshef Kera, 2020d). They refused to accept the

---

[2] Future of RegTech: How to Regulate Algorithms? Tableau visualization: https://tiny.cc/lithopy

[3] Report I. June 2020 visualization with R of e-Rouska responses: https://les.zcu.cz/eRouska.html; Report III. January 2020 visualization with R of Lithopy responses: https://les.zcu.cz/eRouska3.html; Report II.

  Comparison of June 2020 and April 2021 eRouska responses: https://les.zcu.cz/eRouska2.html

[4] Zenodo data on eRouska surveys https://zenodo.org/record/5949422

imaginary roles of stakeholders in the village in the stakeholder-oriented 2019 design of the sandbox (Reshef Kera, 2020a) and switched to their identities as citizens in the village, which informed the 2020 design of the sandbox. Instead of expecting any unity, the participants became comfortable with what we describe as "regulation through dissonance" (Reshef Kera, 2020c). There was a willingness to experiment with both code and regulations while acknowledging the limits of their knowledge or control.

Based on the Lithopy experience, we claim that exploratory sandboxes democratize the decision-making processes and support negotiations of diverse groups about concrete scenarios (Kera, 2021; Reshef Kera, 2020a). It is a training and testing ground for understanding the issues of power, stakes, and ownership when facing a new code or missing regulations. The sandbox confronted the participants with hypothetical but functional surveillance over persistent satellite and drone data that triggered smart contracts for citizenship, marriages, and property rights. It then included a small piece of code that discriminated against Czechs who wanted to buy property in the village and challenged the participants to decide upon regulations. The early 2020 sandbox showed that even participants that reject intrusive services provide nontrivial ideas on how to regulate them (Kera, 2021), which inspired us to continue the research.

The COVID-19 pandemic contact tracing provided an opportunity to compare the feedback on the hypothetical services in the sandbox with June 2020 decisions on installing an actual application that monitors and reveals the infection status via proximity measured over a Bluetooth service. Despite the differences in their functions, both services present the challenge of accepting persistent everyday surveillance. In 2020 and 2021, we decided to follow how the decisions to install or not install relate to various concerns and regulatory expectations about the oversight.

In June 2020 and April 2021, we sent surveys to the sandbox participants to probe what type of regulations they expect in contact tracing applications. About two-thirds of the participants in 2020 rejected the intrusive algorithmic services in both cases as a hypothetical possibility and real application and expressed a low trust in the oversight, including the technical promises of open-source solutions or Bluetooth technical "fixes" that preserve privacy. The decisions to install or not install an intrusive algorithmic service seemed to be driven by deontological, principle-based, rather than utilitarian concerns, for example, appeal to surveillance that "saves lives." We could not follow the responses on an individual level because the original study supported the complete privacy of the participants. However, when we analyzed the responses of the entire cohort, they seem consistent with the January 2020 responses about the sandbox (Kera, 2021).

Both groups that accepted (one-third) and rejected (two-thirds) the intrusive but "polite" surveillance shared similar reasons for their actions. They expected regulations based on an independent rather than government oversight with elements of self-regulation that we interpreted as a call for experimental and iterative design and policy interventions. While the ones that accepted the intrusive service believed that the independent oversight was already happening and trusted the claims of the platform developers, the two-thirds that rejected the service was skeptical about the present oversight.
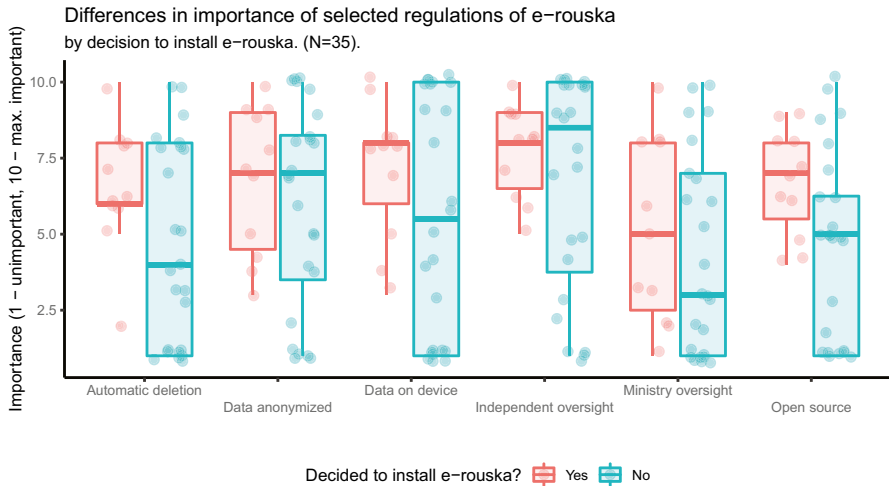
Fig. 1 Attitudes toward e-rouska regulatory expectations: regulation over anonymized data, automatic deletion, data in mobile, open- source solution, Ministry of Health, and independent oversight

Instead of a strong thesis on what makes algorithmic services acceptable or unacceptable to the public, our claim is moderate. With this pilot study, we mapped the different reasons for accepting or rejecting algorithmic services and the different levels of trust in the oversight. We interpret them as something that changes depending on the experiences with agency, participation, and representation in the design and policy process. In the next chapters, we will discuss the insights from the three surveys and then focus on the claim that we can increase the trust in the oversight by engaging the participants directly in the design and policy decisions about the services.

5 Contact Tracing Regulatory Expectations

The app e-Rouska[5] was introduced in April 2020 as a state-of-the-art solution for contact tracing during the first wave of the COVID-19 pandemic in the Czech Republic. The app promised anonymization, privacy, and control over the data, and it gained enormous media support; however, a majority of citizens remained reluctant to install it. According to media reports, about one million in a country of 10 million had installed it by October 2020 (Kenety, 2020), and it was officially described as a failure in February 2021 (Nemcova, 2021) without ever achieving the intended goals.

---

[5] Contract tracing application https://erouska.cz/

We can see the reasons for the reluctance to install in Fig. 1 based on the June 2020 survey that also confirms the January 2020 pre-pandemic rejections of hypothetical intrusive services in Lithopy discussed in the next chapter. On the boxplot visualizations in the report I[6] (Fig. 1), we can follow the weights of the various regulatory expectations that played a role in the decision to install (red) or not install (green) the application. To be able to compare the responses to Lithopy services and contact tracing app in the same cohort, we mapped the January Lithopy survey (report III[7] and tableau[8]) into technical and non-technical regulatory expectations from contact tracing (reports I and II[9]).

In the June 2020 and April 2021 surveys, we used information on what the developers claimed e-rouska does to protect their privacy as technical interventions, and we added two more categories (government or Ministry of Health based and independent oversight) as non-technical interventions. Technical interventions in Lithopy on the level of "architecture" and "application" became what the developers claim contact tracing app provides (i.e., data anonymization, data remaining on the mobile phone, automatic deletion, open-source technology). We then mapped the non-technical interventions in Lithopy such as "audit," "industry codes of conduct," "market," and "government" into "Ministry of Health" versus "Independent oversight" in the contact tracing.

The group that installed contact tracing ("adopters") was small (one-third) but homogenous in their responses to different interventions or regulatory expectations about the intrusive app. The adopters trusted the claims of the developers and appreciated that the app was open source. They also considered it important that the data are anonymized and remain on users' devices as something guaranteeing their security. In the group that refused to install contact tracing in June 2020 (24 out of the 35 participants from the original cohort), we can notice more diverse attitudes toward the different regulatory interventions, ranging from 1 (not important) to 10 (maximum importance).

The rejectors seem to have two extremely divergent views on what constitutes good governance of such intrusive services in all six categories: "data anonymization," "saving data on the mobile phone of the user," "automatically deleting data in 30 days," "open-source technology," "oversight by an independent institution," or "government (Ministry of Health)." It seems there are two different personas of "rejectors": the first, who refuses to install such applications in principle and expresses low trust in all regulatory instruments, and the second, who trusts the interventions but does not believe the present app meets the criteria. This distinction in the two personas of the rejectors is clear in the case of interventions that place the data under the control of the user on his or her mobile (so they can decide when to

---

[6] Report I. June 2020 visualization with R of e-Rouska responses: https://les.zcu.cz/eRouska.html

[7] Report III. January 2020 visualization with R of Lithopy responses: https://les.zcu.cz/eRouska3.html#Plot

[8] Tableau Lithopy visualization: https://tiny.cc/lithopy

[9] Report II. Comparison of June 2020 and April 2021 e-Rouska responses: https://les.zcu.cz/eRouska2.html
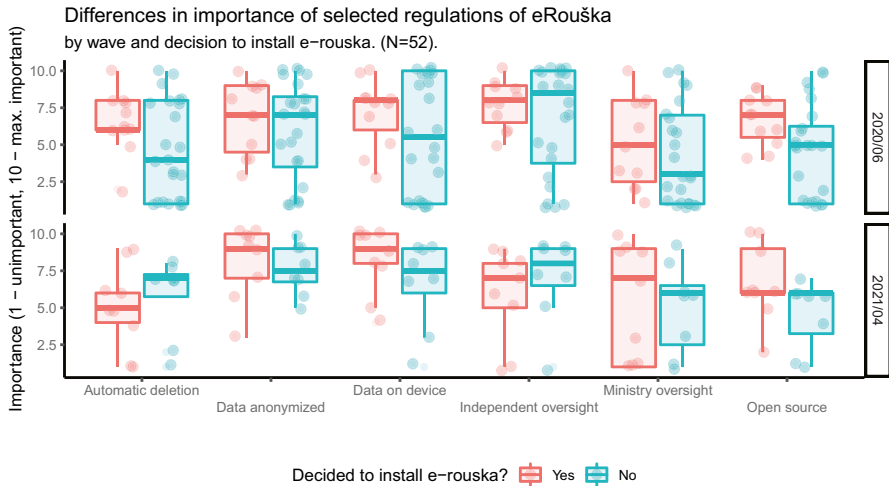
Fig. 2 Comparison of survey data on contact tracing from June, 2020 to April, 2021

share if infected by the virus). The rejectors marked this as equally important and unimportant (the median is about 5.5). While the ones who think this is an important criterium did not trust the present application delivering upon the promise, the ones who think it is not important refuse such applications in principle.

This changed in the second wave survey from April, 2021, when it shows that the rejectors became more homogenous in their attitudes toward the interventions (Fig. 2), such as "data on the device," "automatic deletion," and "data anonymization." The April 2021 feedback by the adopters of contact tracing paradoxically became less homogenous, especially in the way they perceived the government interventions (oversight by Ministry of Health)—Fig. 2. The participants who installed the application in June 2020 (11 out of 35 that responded) marked all criteria as important (median above 5) except for the support and oversight of the application by the government (Ministry of Health). They seem to agree on this with the rejectors (the lowest median value is 5 in the adopters that installed the app.) Both groups agreed that independent oversight is the most important intervention, but in the group that installed the app the median was 8, while the rejectors' score was even higher (almost 9).

The adopters and rejectors in 2020 seem to agree on two important values driving their decision to install or not install: "data anonymization" and "independent oversight." We defined the independent oversight in the survey as a "continuous and regular oversight of the system (security and data management) by an independent organization with transparent methodology that follows the safety issues on the mobile but servers with which the mobile interacts." When compared with the second survey from April, 2021 (report II) in Fig. 2, we see that this value remained similar and essential for all groups in both periods. In the case of the rejectors, we can notice a growing interest in government interventions (from median 3 in June,

2020, to 6 in April, 2021); however, the group that installed the application lost interest in this intervention (decreasing from 7 to 5), so that even if it had previously played a role, it was not essential later in the pandemic.

In the written feedback, we can see that the low trust in government institutions (Ministry of Health) among both the adopters and rejectors of contact tracing relates to their view of independent oversight as essential. Among many functions of such oversight, the participants describe the importance of random security testing that provides quick feedback:

- "(The ideal) is an independent agency selected in a competition that will randomly check that user data are not being misused";
- "I don't understand the technology well, but (ideally) some form of quality control over the security of the system by whitehat hackers (developers who will not misuse the bugs they identify)";
- "Independent supervision by someone who will be impartial";
- "Government organized tender to select an independent company to ensure that the data are not misused";
- "Regular supervision of how the system operates with transparent methodology that monitors data security. In an ideal world, an application could be very effective if all citizens were able to use it and everyone used it properly."

Several of the participants also emphasized the importance of involving the public in oversight and improving the literacy of the citizens in this respect:

- "(The ideal) is a system of feedback from a wide range of independent IT professionals (or organizations) and the public";
- "To some extent, a supervision from the users themselves over the app or someone neutral";
- "The application is supposedly secure and decentralized, but if it is not open source and because I'm not a skilled programmer, I can't be sure that it doesn't contain a backdoor."

The interest in independent oversight remained strong in the second survey, but the responses were more pragmatic—not as much a wish list and opinions—but expressing the acceptance of the application as a necessary evil or even joking about its intrusive surveillance: "No, I don't think anyone is interested in my location. I still spend most of my time on the couch at home now $<U+0001F914> <U+0001F603>$– thinking and smiling face."

One important category on which both groups of adaptors and rejectors relatively agreed concerned the "anonymized data." The group that installed contact tracing was more homogenous in its responses, while the rejectors were spread in equal parts across the 1–10 scale, and the median remained similar in both groups (around 7.5). In the second wave, the April, 2020, survey suggests that anonymization of data became central for the group that installed the app (median 9) and for the rejectors who became more homogenous in their responses (median 8). Strangely, the two groups disagreed about the possibility of automatic deletion of data (as another

example of a technical solution promising anonymity); for the rejectors, this intervention rose in significance (from 5 in June, 2020, to 7.5 in April, 2021); however, for the groups installing contact tracing, the "automatic deletion of data" fell significantly from 7 in June 2020 to 4.5 in April,2021.

The rising interest in government (i.e., Ministry of Health) oversight among the rejectors was surprising (from 3 in June, 2020, to 6 in April, 2021), while in the group that installed the app, it remained ambiguous and possibly divisive, rising from 3 in June, 2020, to 4.5 in April, 2021. In the June 2020 survey, the rejectors expressed strong distrust in the government as a regulatory and oversight body (median 2.5) but also in the user's control over the mobile phone device (automatic deletion of data after 30 days). This shows skepticism toward government but also technical solutions as safeguards of privacy except for the category of "data anonymization." (The median in June, 2020, was 7.)

The adopters and rejectors responded differently to the importance of open-source solutions as a safeguard and intervention preventing the excesses of surveillance and algorithmic rule. In the group that installed the app, we can see a high level of trust in the open-source solutions (median 7.5) dropping in April, 2021, to 6 while rising among the rejectors from 5 to 6.5. Both groups, however, seem to agree on the importance of saving the data on the user's phone and anonymization, while the issue with deletion became less important for both groups between June, 2020, and April, 2021. The group that installed the app seems to express more resignation, and perhaps skepticism, toward technical interventions in April, 2021, while the rejectors seem more willing to demand the technical but also institutional interventions.

6 Sandbox Smart Contracts Expectations

While intrusive surveillance became an essential tool of the pandemic response in June, 2020, the January, 2020, sandbox presented hypothetical but functional, near future services. Smart contracts in the imaginary village of Lithopy used satellite and drone data to automatize the blockchain transactions, such as becoming a citizen, marrying someone or buying and selling property. They offered a direct experience with regulating and changing the code on the example of a biased and discriminatory contract on property ownership that participants tried to regulate (Kera, 2021).

The participants interacted with the blockchain services over a dashboard that visualized the various data and actions needed to trigger the smart contracts. After testing the services over the dashboard, they would confront a discrimination in one of the contracts that forbade Czechs from buying a property. The systemic discrimination was hidden as a small piece of JS code documented on GitHub. From a list of regulatory and technical interventions as potential solutions, the participants ranked the preferred ones.

The rejectors of algorithmic rule in the Lithopy sandbox (two-thirds—73%) expressed strong privacy concerns and even called for discontinuation and ban of such services in the imaginary village, which did not appear in the case of contact tracing (Kera, 2021). There was a surprising connection between the self-assessed
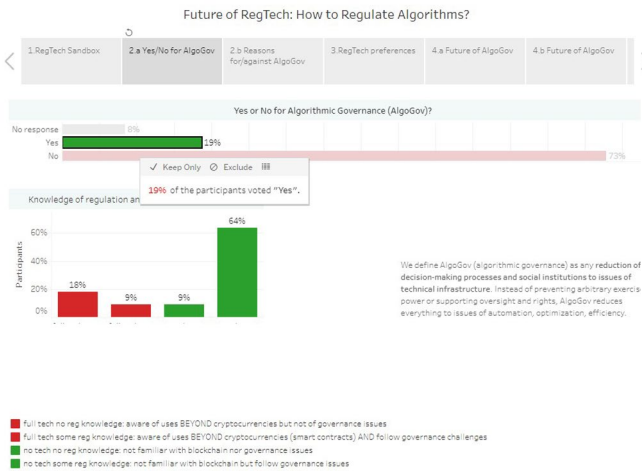
Fig. 3 Adopters of Lithopy services

lack of knowledge of technology and some knowledge of regulatory issues that led to the acceptance of algorithmic services, which we did not measure in the case of contact tracing. The majority of the adopters (Fig. 3—tableau story slide 1) claimed to have no previous knowledge and experience with blockchain, cryptocurrency, or smart contracts but some knowledge of regulatory and governance issues (the green "no tech/no reg" and "no tech/some reg" categories on the slide 1 in the tableau story).

When we look at the visualization in Fig. 3 from tableau story slide 1, a typical supporter of the algorithmic rule is someone with no knowledge of blockchain technology but awareness of governance and regulation issues (64%). It is someone open and curious about the possibility of regulation on both technical and institutional levels and enjoying the experience in the sandbox. In their feedback, the majority of participants in the green group, both rejectors and adopters, called for changes in the Lithopy services rather than their ban (Fig. 4—tableau story slide 3). They supported (and highly supported) any technological and non-technological interventions except self-regulation over market forces (Kera, 2021).

The strongest opponents of algorithmic services (Fig. 5) remained the self-assessed experts in cryptocurrency with no knowledge of governance and regulatory issues that seem to hold extreme libertarian views. They prefer technologies they already knew (cryptocurrency) as solutions to all governance issues.

The tentative support for algorithmic services based on surveillance was mainly driven by individuals that claimed no knowledge of the technology ("no tech/no reg" and "no tech/some reg" on slides 2 and 3 in the tableau). In the 42% of participants (Fig. 6, tableau story slide 3) that self-identified with some knowledge of governance issues, we noticed almost three times more support for the algorithmic rule and technological interventions than in the group with the knowledge of the technology.
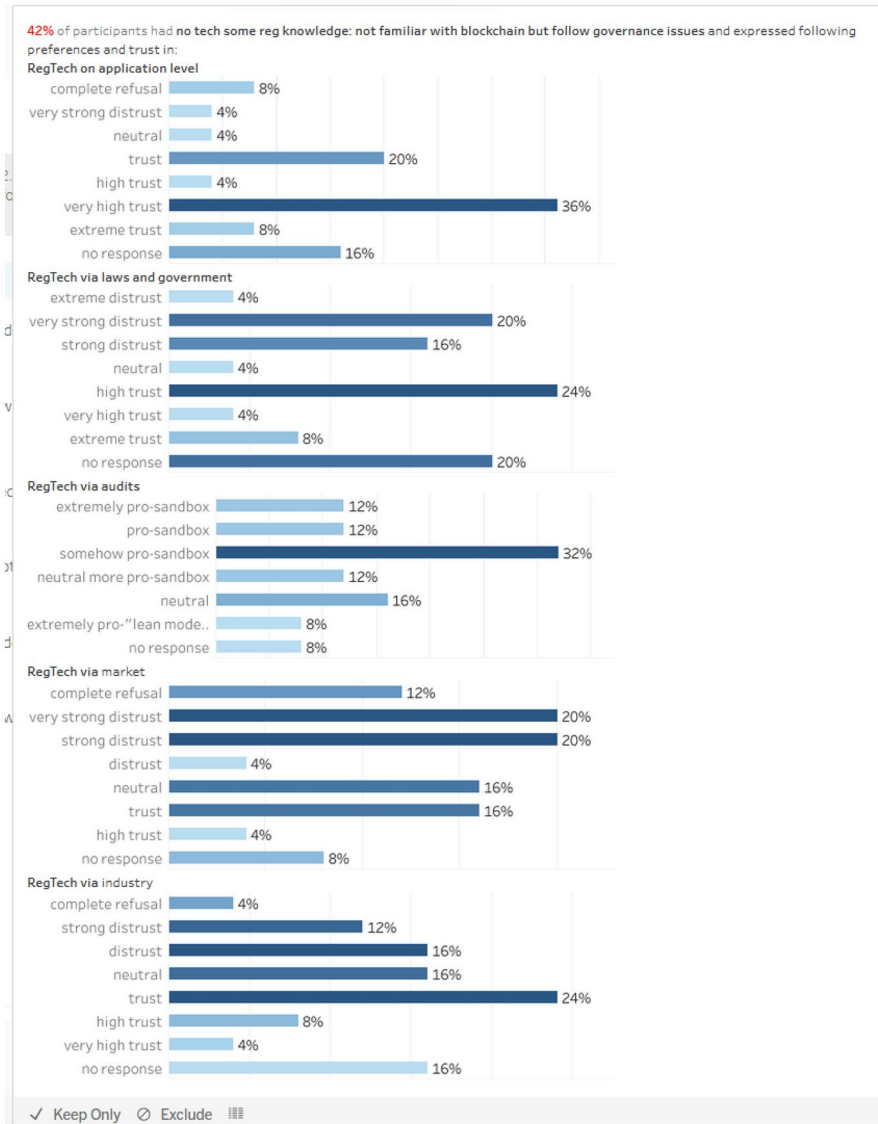
Fig. 4 Self-assessed majority of participants with no knowledge of blockchain but awareness of governance and regulatory issues describing their attitudes toward different regulatory interventions

Only one participant in that group of "experts" was curious to experiment with the algorithmic rule in Lithopy, while the general support among the 59 participants was 25%.

While in this green/no-tech group, the majority refused algorithmic services in Lithopy, we can see that they expressed strong support (68%) and trust in the possibility of regulating the smart contracts on the level of code, such as "privacy by
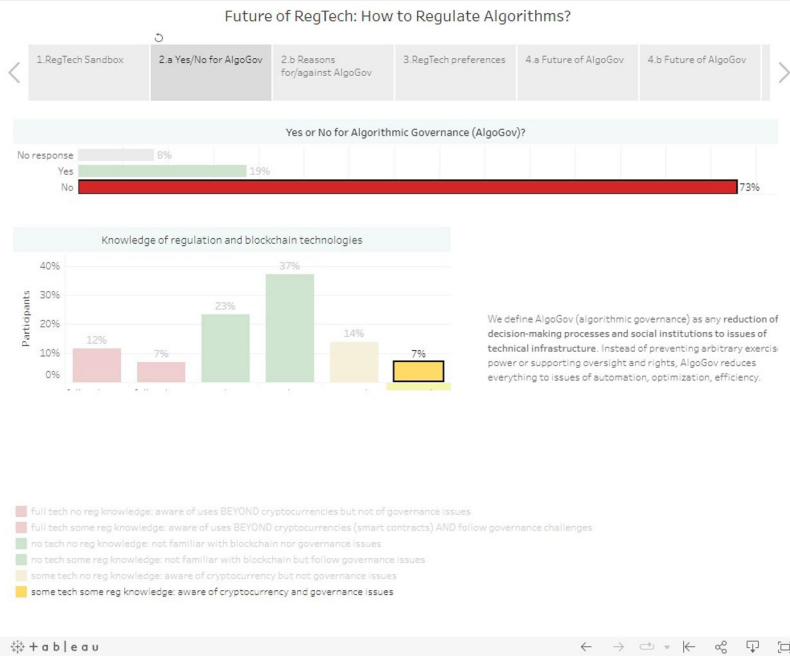
Fig. 5 Self-assessed participants with knowledge of cryptocurrency but no awareness of regulatory and governance issues rejecting algorithmic services

design" solutions (tableau slide 3 "RegTech preferences," and "RegTech via code/ applications" part of the survey—tableau slide 5). The support for "governance by design (code)" solutions in this group contrasted with the more evenly distributed feedback on the various modes of regulation (industry, market, government) in Fig. 4 (tableau slide 3 "RegTech preferences").

We can notice a similar trust in technical interventions also in the adopters of contact tracing (reports I and II). Participants that installed the app had a higher median in all technical categories except the regulation via the Ministry of Health that represented governance regulations and interventions. In the report I, we can notice that they ranked as high interventions, such as "data saved on the device" (median 8), "open-source solutions" (median 7), "anonymized data" (median 7), "automatic deletion after 30 days" (median 6.5). Despite the differences between a real application for monitoring contacts during a pandemic and a hypothetical infrastructure for selling and buying property, the 2020 reactions share many similar themes. In summary, about two-thirds of the participants refused the algorithmic services (even if they claimed public good) and expressed similar distrust of government oversight.
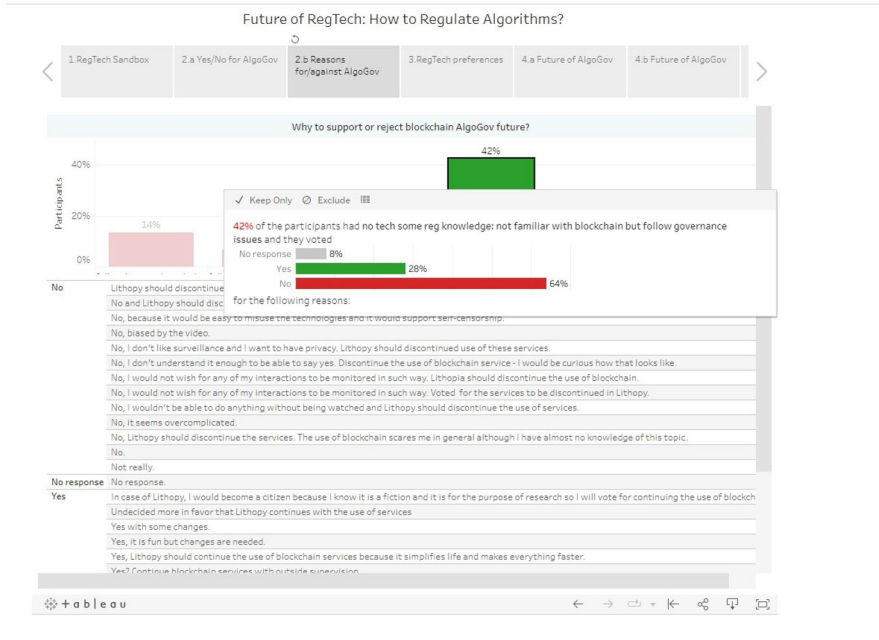
Fig. 6 Self-assessed participants with no knowledge of technology but some awareness of regulatory and governance issues accepting algorithmic services

7 Algorithmic "Polite" Surveillance

In the pre-pandemic January workshops, participants provided feedback on the pervasive surveillance performed by functional prototypes of blockchain services that use satellite and drone data in the smart contracts. We were searching for a provocative, near-future scenario that shows surveillance and automation as part of everyday life to confront the participants with a sudden and unexpected bias in the code. What we did not expect in January 2020 was that this seemingly "polite" and everyday surveillance would become a reality with contact tracing. The pandemic created an opportunity to compare the reactions to a hypothetical service with real-life decisions about contact tracing. In May 2020, we initiated the pilot longitudinal study on the same cohort to follow the decisions about a contact tracing app in the Czech Republic during the first and second COVID-19 waves in June 2020 and April 2021.

While the Lithopy sandbox survey had provided valuable insight into the regulatory expectations concerning a biased smart contract, the next two contact tracing surveys followed whether participants installed contact tracing and how they ranked the oversight of the service. We compared the pre-pandemic responses to satellite-supported smart contracts with the decisions about Bluetooth-enabled contact tracing during a crisis. Both algorithmic services show a possibility of what we call "polite surveillance." It is an intrusive, pervasive, and everyday surveillance that citizens accept in order to signal a civic duty or belongingness to a community.

In Lithopy, the citizens expressed their belonging by engaging in a transparent life under the "eye" of the satellites and drones with an aesthetic appeal of land art interventions, dance, and pantomime. In contact tracing, the surveillance supported the performance of civic duty during a crisis when the citizens shared information on their infection status.

The initial feedback on the hypothetical blockchain services was similar to the later responses to contact tracing during a crisis. In 2020, almost two-thirds of the participants refused any "polite" surveillance and expressed distrust in the government oversight of algorithmic services. The participants perceived the smart contracts in Lithopy or contact tracing in the Czech Republic as more of extraterritorial and lawless zones than future infrastructures allowing regulations. In both periods, they also valued independent oversight above the oversight by government institutions that presented an important factor in the decisions to install or not install. While the participants who installed the intrusive services assumed that independent oversight was possible in the present, the ones who did not install rejected the technical intervention as sufficient. They refused to trust any existing technical intervention supporting oversight, such as open-source code or mobile ownership of the data.

While the first contract tracing survey from June 2020, with 35 responses confirmed the findings of the Lithopy sandbox, the second survey, with only 17 out of the initial 59 responses (April 2021), rejected the findings of polite surveillance. In the last survey, we can see growing support for algorithmic services and trust in the government oversight that problematizes the 2020 findings. We suspect that the small size suggests a self-selection bias, and those who decided to participate were favorably inclined toward contact tracing or were deeply invested in the highly polarizing issue of that period (public health NPIs, i.e., nonpharmaceutical interventions). For this reason, we mention the results in the discussion, but we decided not to include them in the final conclusions. They nevertheless provide important feedback on conducting panels and monitoring the participants and their feedback over longer periods. Panel-based research into algorithmic governance and comparison of different services provide an important corrective to the results from the experimental environment, such as sandbox, exploring the different regulatory and technical interventions.

In the hypothetical Lithopy services, surveillance was an aesthetic experience of living in a community where gestures and land art interventions are social rituals creating bonds in the village. In the case of contact tracing, surveillance became a matter of social responsibility. Both show "polite" surveillance as an intrinsic part of everyday life and interactions, and like the "polite society," they conceal the true power relations and conflicts to serve the status quo (Carter, 2000). We define as "status quo" any situation in which one group enjoys an unchallenged power to define rules, values, and code (algorithms) by simply claiming to have good intentions. The good intentions and "politeness" of algorithms will never compensate for their missing legitimacy, which is possible only if the decisions about the code become part of a political process and negotiations.

The most important result from the pilot study was that the participants expressed a low trust in the existing government interventions and preferred independent

audits and oversight of the services that bring together various stakeholders. We are using the political slogan of the new colonies demanding "no algorithmization without representation" to interpret this feedback in the next chapter as a call for hybrid governance of near-future algorithmic services. The "representation" includes ordinary citizens but also IT experts and representatives of the industry who collaborate on pragmatic solutions rather than reduce it to regulation or code.

The situation of citizens exposed to the new algorithmic services is reminiscent of colonists in new lands. They face unique challenges in balancing the power of existing institutions with the freedom and dangers of the new territories. Citizens resist the new "tyrants" (the owners of the data, algorithms, and platforms) as much as the lawlessness and slavery of their new condition and other excesses of "colonization." The present emphasis on ethical guidelines and frameworks supporting ex ante and ex post interventions or the over-promissory, technocratic "governance by design" solutions do not solve the issue of representation and sovereignty. They only delegate it to the old or new "tyrants." Instead of reducing the democratic values and regulations to code—such as privacy-by-design (Cavoukian, 2009) and society-in-the-loop (Awad et al., 2018; Rahwan, 2018)—or insisting on oversight by a public body outside the infrastructure, we think our research shows strong support for hybrid, tactical, and situated engagements with automation and infrastructure (Hee-jeong Choi et al., 2020; Shilton, 2018; Sloane et al., 2020) similar to calls for hybrid "adversarial" public AI system solutions (Elkin-Koren, 2020).

8 Algorithmic Services as New Algo-Colonies

All three surveys (January 2020, June 2020, and April 2021) showed strong support for independent oversight as an alternative to the government-based or technological interventions, including the governance by design or RegTech proposals (Mulligan & Bamberger, 2018). Based on the data from the first Lithopy sandbox survey, we interpreted the independent oversight as a search for hybrid and experimental forms of regulation that insist on the participation and representation of the citizens in the whole design and policy process.

In the original survey (Fig. 7, tableau slide 4d "Future of RegTech Preferences" and report III.), "independent audits" became the preferred mode of regulation in the sandbox (extremely important, very important, and important category covered 65% of the responses, while 29% remained neutral and only 7% refused the audits).[10] Even the second-highest valued intervention in Lithopy was "hybrid," neither a technology nor institution, but a mix of both—industry standards as a way of preventing discriminatory code (the extremely important, very important, and important categories covered 61% of the responses, 31% were neutral, and 8% rejected). In the case of contact tracing, independent oversight had the highest median value of almost 9 (out of 10) in the adopters and over 8 in the rejectors (Fig. 1).

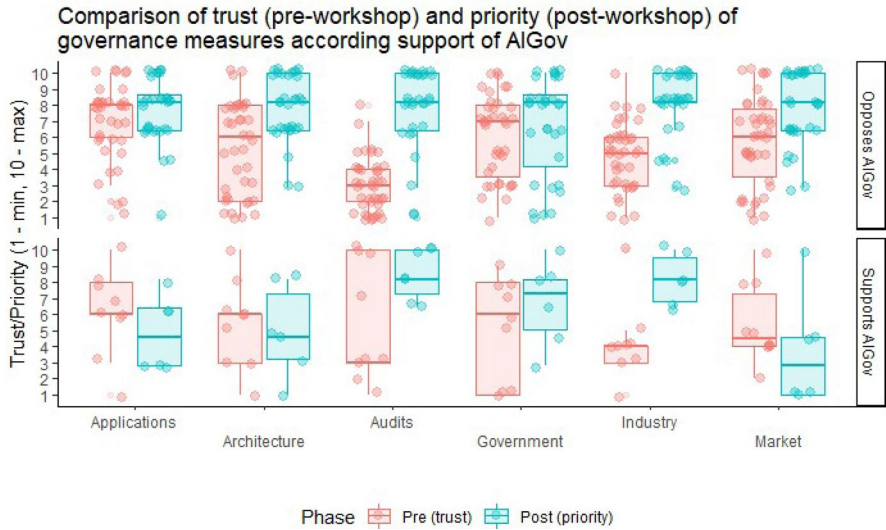[10] The sum of all 7 categories made 101% because of rounding.

Fig. 7 Comparison of attitudes at the start of the workshop in Lithopy sandbox (after seeing a design fiction movie) and end (after experiencing all the regulatory, technical, and nontechnical interventions)

While Lithopy tableau's story visualized the responses that we discussed in a previous article (Kera, 2021), here we are using a box plot (Fig. 7) to show how the attitudes toward different interventions changed at the start and end of the workshops (before and after experiencing the sandbox environment). The experimental and hybrid solutions (audits and industry) that are neither government nor market or technology-based dominate the chart. Support for audits as an example of an independent oversight grew from the median value of 2–3 to 8–9 at the end of the workshop (maximum was 10). The audits also defined something of a consensus on regulation between the adopters and rejectors of the algorithmic and intrusive surveillance later confirmed by the responses on contact tracing.

In the first 2020 survey on contact tracing, the preference for an independent oversight was also in sharp contrast to the government oversight as the least popular regulatory intervention for both the adopters and rejectors. The trust in the government's ability to regulate contact tracing improved only in the last April 2021 survey. The low trust in the government regulation could be cultural (post-communist history of state control abuse), but the data from the first survey in Lithopy show a more complex issue. Since we included questions on how participants imagine the independent oversight, we can notice the expectations of participation and representation of citizens in the new algorithmic services. On tableau 6a (RegTech via Audits overview), we can see strong support for the sandbox model of developing and regulating algorithmic services (47% support and 44% consider it a priority). The majority also ranked the "security of the services" and "transparency of the data" as more important categories in auditing than the "compliance with existing regulations" (only 31% consider this as S priority). In terms of who should audit (tableau 6c and 6d), there was a preference

for independent bodies (41% consider it as priority) while government oversight was a priority only for 22%, which is less than the industry defined certificates ranked as a priority for 29% of the participants. The ranking of government regulation and oversight (tableau 9b) shows limited support for all categories except one: "communication and improvement of literacy."

The exploratory sandbox helped the participants to envision a more active role in the design and regulation of the new services, which we believe influenced their final responses and preference for independent oversight. By experiencing hybrid and iterative engagements that crossed the divisions between the code and values, new infrastructures, and old institutions, the participants realized their limits on every level. Instead of resigning upon their agency or expecting the existing institutions of governance to save them, they became something of data "colonists." They were trying to figure out their sovereignty in the new territories between the old and new power structures, the code, and norms. Instead of a matter of pre-defined values and norms, the regulatory expectations became a question of representation and participation in the process of "algorithmization."

To summarize the experiences and feedback from the sandbox together with the later responses on contact tracing, we are using the credo of the American colonists demanding "no taxation without representation" and claim that the algo-colonists are demanding "no algorithmization without representation." The new data or algo-subjects under surveillance in the new (digital) territories are trying to balance the different forces, opportunities, and threats. On one side, they are still citizens from the old "world" with rights and expectations that the new territories do not always respect. On the other, they are new algo-subjects or what Rouvroy described as "infra-individual data and supra-individual patterns" (Rouvroy, 2013) under the control of the owners of the code. In the new land, they struggle to define their sovereignty and agency on new grounds.

The "no algorithmization without representation" credo summarizes the search for hybrid and experimental procedures that would guarantee the agency of the new "colonists" and legitimacy of the regulatory processes. We are gradually learning how to represent the various stakes in the emerging digital infrastructures where we are not only citizens but also users and stakeholders in some visions of the future. The fundamental challenge seems to be the participation in the design and policy processes that will decide on the future of the new environment. We can describe this challenge as a question of sovereignty of the new inhabitants on the emerging platforms or "algorithmic sovereignty" (Roio, 2018).

The attempts to regulate the new data and "algo" territories often ignore the issues of participation and representation of citizens. They insist on aspirational and normative frameworks that rely upon the institutional actors claiming to represent the public interests. These solutions remind us of the problematic legacy of the so-called "parastatals," the post-colonial attempts at sovereignty through corrupt public–private partnerships (Khanna, 2012). To control the algorithmic "colonization" by emerging services, we need to insist on the direct participation of the citizens and their representation at every step of the design and regulation processes.

The regulatory and exploratory sandboxes create a space between the regulation and code where the users can become stakeholders and explore their sovereignty,

representation, and participation in governance. We can see from the report III (Fig. 7) that after experiencing such a hybrid environment, participants increased their support for all technical and non-technical interventions, which indicates a stronger sense of agency. The hybrid and experimental sandboxes thus offer an alternative to the ethical guidelines or ex ante and ex post interventions that are difficult to implement. They engage the participants directly in the attempts to embed the legal but also cultural and social norms into the corporate-owned and machine-readable code and infrastructures.

The experience in the sandbox enabled the participants to define the level of representation they expect in the future algorithmic services and prepare grounds for their sovereignty. To support the adoption but also the regulation of emerging and disruptive technology, we need to represent the interest of the new stakeholders and data/algo colonists and make the clashes between the different interests and stakes transparent and visible. We think the sandboxes offer such an environment to experience the conflicts and uncertainties without insisting on final decisions and unity. They are "trading zones" that support the interactions between the participants without insisting on any unity that would guarantee clear sovereignty (Reshef Kera, 2020b).

9 Summary

The key challenge in regulating algorithmic services is to engage citizens not only as test subjects or users of new services but as actual stakeholders in the future as something of a new territory with unclear sovereignty, political representation and participation. We summarized this as an issue of participation and representation in the process of "algorithmization." The emerging algorithmic services present a similar challenge as any extrajudicial territory or transnational, intergovernmental, and supranational organization where "there is no overarching sovereign with the authority to set common goals even in theory, and where the diversity of local conditions and practices makes the adoption and enforcement of uniform fixed rules even less feasible than in domestic settings" (Zeitlin, 2017).

The sandboxes are sites that support this experimental or experimentalist approach to governance (Sabel & Zeitlin, 2012) and emphasize the participation of stakeholders in the entire policy and design cycle from decision making to reflection and implementation. This is an iterative process with many risks and uncertainties, but it is essential that the regulation and policy include prototyping and design engagements with the stakeholders and thereby extend the discursive nature of the governance processes.

The experimental governance of emerging services in the sandboxes offers a pragmatic alternative to the hierarchical, command-and-control models of governance, but also to the aspirational ethical frameworks and recommendations that come too late or achieve too little. The experiments with near-future algorithmic life in the sandboxes thus mitigate some of the asymmetries of power, knowledge, and know-how in novel infrastructures and enables a "situated automation" (Hee-jeong Choi et al., 2020).

It is automation situated and contextualized through experiments that insist on citizen representation and engagement in the design of the algorithms exploring hybrid forms of sovereignty and agency. The citizens as data colonists in the new "transnational" territories and infrastructures are vulnerable subjects who are often responsible to everyone and no one. The lack of clear mandates, responsibility, and shared goals in the emerging infrastructures makes this situation similar to the transnational or extraterritorial regulatory challenges.

The pilot study demonstrated the feasibility of an experimental approach to regulation and governance that supports a "situated automation" via participation and representation of the new data settlers. We offered them an exploratory sandbox as a site that connects the legitimacy based on participation with the efficiency of the code to solve the issues with the democratic deficit.

The exploratory sandbox allows the participants to experience the alternative futures and make more informed decisions or negotiate the preferred interventions. Instead of normative frameworks and guidelines that often hide the stakes, power relations, and interests, the sandbox environments offer direct experiences with regulating algorithmic services. While there are many limitations in this study that we have mentioned (small and nonrandom sample, comparison of two different services, and non-responses), the results are nonetheless useful for designers and policymakers. We present a strong case for participatory and experimental approaches to algorithmic regulations that are backed by comparison with real-life decisions.

Instead of a strong and empirically tested hypothesis, the pilot study shows the need for more experimental and longitudinal approaches to identifying the regulatory expectations and experiences of the citizens. Our main reason for comparing the two services was to check how the aspirations defined in the sandbox relate to real-life decisions; at least in 2020, they were similar. The experimental approaches via sandbox allow the participants to experience how their choices affect the final design and to make iterative decisions upon the regulation and code. We claim that the hybrid and experimental approach to algorithmic services is a viable alternative to the reductionist views of regulation of algorithmic governance involving only ethical or technical issues.

Author Contribution Denisa Reshef Kera (DRK) conceived the first part of the research (Lithopy sandbox) and designed experiments used in the workshop, while Frantisek Kalvas (FK) designed the second part of the research (eRouska) and gathered the data on that. DRK worked on tableau visualization and interpretation while FK solely completed the visualizations of data in R (eRouska visualizations). They both participated equally in the analysis and interpretation of the data; DK wrote the paper, and FK was active in the revisions. Both authors read and approved the final manuscript.

Declarations

Ethics Approval and Consent to Participate University of Salamanca, approval no. 321 of the project "AnticipatoryLedgers: diseño anticipatorio y marco ético para tecnologías de contabilidad distribuida (blockchain o DAG)" in April 14, 2019.

Conflict of Interest The authors declare no competing interests.

References

Alaassar, A., Mention, A. L., & Aas, T. H. (2020). Exploring how social interactions influence regulators and innovators: The case of regulatory sandboxes. *Technological Forecasting and Social Change, 160*, 120257. https://doi.org/10.1016/j.techfore.2020.120257

Alexander, L. (1992). What makes wrongful discrimination wrong? Biases, preferences, stereotypes, and proxies. *University of Pennsylvania Law Review, 141*(1), 149. https://doi.org/10.2307/3312397

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). *The Moral Machine Experiment.* https://doi.org/10.1038/s41586-018-0637-6

Aziz. (n.d.). *Guide to forks: Everything you need to know about forks, hard fork and soft fork.* 2020. Retrieved January 21, 2020, from https://masterthecrypto.com/guide-to-forks-hard-fork-soft-fork/?lang=en

Begby, E. (2021). Automated risk assessment in the criminal justice process. *Prejudice.* https://doi.org/10.1093/OSO/9780198852834.003.0009

Bromberg, L., Godwin, A., & Ramsay, I. (2018). Cross-border cooperation in financial regulation: Crossing the Fintech bridge. *Capital Markets Law Journal, 13*(1), 59–84. https://doi.org/10.1093/cmlj/kmx041

Burke, A. (2019, July 1). Occluded Algorithms. *Big Data & Society, 6*(2), 2053951719858743. https://doi.org/10.1177/2053951719858743

Carter, P. (2000). *Men and the emergence of polite society, Britain, 1660–1800 | Reviews in History.* Longman. https://reviews.history.ac.uk/review/195

Cavoukian, A. (2009). Privacy by Design - The 7 foundational principles - Implementation and mapping of fair information practices. *Information and Privacy Commissioner of Ontario, Canada.* https://doi.org/10.1007/s12394-010-0062-y

Chwalisz, C. (2020). Reimagining democratic institutions: Why and how to embed public deliberation. *Innovative Citizen Participation and New Democratic Institutions.* https://doi.org/10.1787/339306DA-EN

Danaher, J., Hogan, M. J., Noone, C., Kennedy, R., Behan, A., De Paor, A., Felzmann, H., et al. (2017). Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence. *Big Data & Society, 4*(2), 2053951717726554. https://doi.org/10.1177/2053951717726554

De Filippi, P., & Hassan, S. (2018). Blockchain technology as a regulatory technology from code is law to law is code. In *arXiv.*

Elkin-Koren, N. (2020). Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence. *Big Data & Society, 7*(2), 205395172093229. https://doi.org/10.1177/2053951720932296

Fan, P. S. (2017). Singapore approach to develop and regulate FinTech. In *Handbook of Blockchain, Digital Finance, and Inclusion, Volume 1: Cryptocurrency, FinTech, InsurTech, and Regulation.* https://doi.org/10.1016/B978-0-12-810441-5.00015-4

Financial Conduct Authority, & Authority, F. C. (2015). Regulatory sandbox. *Fca, November*, 26. https://doi.org/10.1111/j.1740-8261.2011.01810.x

Fitsilis, F. (2019). Imposing regulation on advanced algorithms. *Springer International Publishing*. https://doi.org/10.1007/978-3-030-27979-0

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Goodman, B., & Flaxman, S. (2017). European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation'. *AI Magazine, 38*(3), 50–57. https://doi.org/10.1609/aimag.v38i3.2741

Gromova, E., & Ivanc, T. (2020). Regulatory sandboxes (Experimental legal regimes) for digital innovations in brics. *BRICS Law Journal, 7*(2), 10–36. https://doi.org/10.21684/2412-2343-2020-7-2-10-36

Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic bias: From discrimination discovery to fairness-aware data mining. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13–17-August-2016*, pp. 2125–2126. https://doi.org/10.1145/2939672.2945386

Hee-jeong Choi, J., Forlano, L., & Reshef Kera, D. (2020). Situated automation. *Proceedings of the 16th Participatory Design Conference 2020 - Participation(s) Otherwise - Volume 2*, 5–9. https://doi.org/10.1145/3384772.3385153

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The Weirdest People in the World?" *Behavioral and Brain Sciences, 33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Herrera, D., & Vadillo, S. (2018). *Regulatory sandboxes in Latin America and the Caribbean for the FinTech Ecosystem and the Financial System*.

Hildebrandt, M. (2018). Algorithmic Regulation and the Rule of Law." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*(2128), 20170355. https://doi.org/10.1098/rsta.2017.0355

Introna, L. D. (2016). Algorithms, governance, and governmentality: On governing academic writing. *Science Technology and Human Values*, *41*(1). https://doi.org/10.1177/0162243915587360

Johnson, G. M. (2020). Algorithmic bias: on the implicit biases of social technology. *Synthese, 198*(10), 9941–9961. https://doi.org/10.1007/S11229-020-02696-Y

Kenety, B. (2020). Over 1 million Czechs download eFacemask app, but many fear 'Big Brother' is watching | Radio Prague International. *Aktualne.Cz*.

Kera, D. R. (2021). Exploratory RegTech: Sandboxes supporting trust by balancing regulation of algorithms with automation of regulations. In M. H. ur Rehman, D. Svetinovic, K. Salah, & E. Damiani (Eds.), *Trust Models for Next-Generation Blockchain Ecosystems* (pp. 67–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-75107-4_3

Khanna, P. (2012). The rise of hybrid governance. *McKinsey & Co. Insights & Publications*.

Kroll, J., Huey, J., & Barocas, S, Felten, E., Reidenberg, J., Robinson, D., & Yu, H. (2017). Accountable Algorithms. *University of Pennsylvania Law Review, 165*(3).

Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., See, D. et al. (2019, November 7). WeBuildAI: Participatory Framework for Algorithmic Governance. *Proceedings of the ACM on Human-Computer Interaction, 3*(CSCW), 1–35. https://doi.org/10.1145/3359283. (January 1, 2017):633.

Madir, J., Lim, B., & Low, C. (2019). Regulatory sandboxes. In *FinTech* (pp. 302–325). Edward Elgar Publishing. https://doi.org/10.4337/9781788979023.00028

Mulligan, D. K., & Bamberger, K. A. (2018). Saving governance-by-design. *California Law Review, 106*(3), 697–784. https://doi.org/10.15779/Z38QN5ZB5H

Nemcova, J. (2021). eRouška nefunguje, nainstalovalo si ji málo lidí. Při trasování nepomáhá, říkají hygienici | iROZHLAS - spolehlivé zprávy. *IRozhlas*.

Rahwan, I. (2018). Society-in-the-loop: Programming the algorithmic social contract. *Ethics and Information Technology, 20*(1), 5–14. https://doi.org/10.1007/s10676-017-9430-8

Reshef Kera, D. (2020a). Experimental algorithmic citizenship in the sandboxes: An alternative to ethical frameworks and governance- by-design interventions. In C. Meza, L. Hernández-Callejo, S. Nesmachnow, Â. Ferreira, & V. Leite (Eds.), *Proceedings of the III Ibero-American Conference on Smart Cities (ICSC-CITIES2020a)* (pp. 29–43). Instituto Tecnológico de Costa Rica. https://www.researchgate.net/publication/351148723_Experimental_Algorithmic_Citizenship_in_the_Sandboxes_an_Alternative_to_Ethical_Frameworks_and_Governance-_by-Design_Interventions

Reshef Kera, D. (2020b). Sandboxes and testnets as "trading zones" for blockchain governance. In *Advances in Intelligent Systems and Computing: Vol. 1238 AISC*. https://doi.org/10.1007/978-3-030-52535-4_1

Reshef Kera, D. (2020c). *Sandboxes and testnets as "trading zones" for blockchain governance*. 3–12. https://doi.org/10.1007/978-3-030-52535-4_1

Reshef Kera, D. (2020d). Anticipatory policy as a design challenge: Experiments with stakeholders engagement in blockchain and distributed ledger technologies (bdlts). *Advances in Intelligent Systems and Computing, 1010*, 87–92. https://doi.org/10.1007/978-3-030-23813-1_11

Reshef Kera, D., Kraiński, M., Rodríguez, J. M. C., Sčourek, P., Reshef, Y., & Knoblochová, I. M. (2019). Lithopia: Prototyping blockchain futures. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3290607.3312896

Roio, D. (2018). *Algorithmic sovereignty*. University of Plymouth.

Rouvroy, A. (2013). The end(s) of critique: Data behaviourism versus due process. *Undefined*. https://doi.org/10.4324/9780203427644-16

Sabel, C. F., & Zeitlin, J. (2012). *Experimentalist Governance*. https://doi.org/10.1093/OXFORDHB/9780199560530.013.0012

Shilton, K. (2018). Engaging values despite neutrality. *Science, Technology, & Human Values, 43*(2), 247–269. https://doi.org/10.1177/0162243917714869

Shneiderman, B. (2016, November 29). The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight. *Proceedings of the National Academy of Sciences, 113*(48), 13538–13540. https://doi.org/10.1073/pnas.1618211113

Shorey, S., & Howard, P. N. (2016). Automation, big data, and politics: A research review. *International Journal of Communication, 10*, 5032–5055.

Sloane, M., Moss, E., Awomolo, O., & Forlano, L. (2020). *Participation is not a design fix for machine learning*.

Susskind, J. (2018). Future politics. living together in a world transformed by tech. *Oxford University Press*, 516.

Vili Lehdonvirta. (2016). The blockchain paradox: Why distributed ledger technologies may do little to transform the economy. *Oxford Internet Institute Blog*. https://www.oii.ox.ac.uk/blog/the-blockchain-paradox-why-distributed-ledger-technologies-may-do-little-to-transform-the-economy/

Winfield, A. F. T., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2133). https://doi.org/10.1098/rsta.2018.0085

Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation and Governance, 12*(4), 505–523. https://doi.org/10.1111/rego.12158

Zeitlin, J. (2017). *Extending experimentalist governance?: The European Union and transnational regulation*. Oxford University Press.