Testing & Experimentation Facilities: Exploring the link with AI Regulatory Sandboxes, Living Labs & AI Testbeds

Authors

Dimitri Schuurman (presenter & contact)[1], Laure De Cock[2], Martin Canter[3], Thomas De Meester[2], Olivia Willems[2] & Tamara De Swert[2].

1: imec-MICT Ghent University, Belgium. dimitri.schuurman@imec.be
2: imec, Belgium.
3: VUB – FARI – AI for the Common Good Institute, Belgium.

Abstract

The current surge in AI innovations is gaining increased momentum and attention from both enthusiastic early adopters and concerned stakeholders. The European Commission aims to lead in this rapidly evolving field by implementing legal constraints to prevent potential negative consequences while also helping European AI innovators compete globally. To balance strict legislation with the encouragement of innovation, four sector-specific Testing-and-Experimentation Facilities (TEFs) were launched in 2023. These TEFs serve as specialized, large-scale sites where technology providers across Europe can test and experiment with advanced AI solutions in real-world environments. In the literature, these new innovation facilities are linked to three established concepts: regulatory sandboxes, Living Labs, and testbeds.

This paper focuses on CitCom.ai, one of the four operational TEFs, which aims to bridge the gap between AI innovators and the development of smart, sustainable cities and communities. Through a three-part triangulation study, we explore the connection between TEFs and other innovation concepts, the alignment between AI innovators' needs and the TEF service offerings, and the actual usage of these services in running or planned experiments at various TEF sites.

Our findings reveal that TEFs incorporate elements of regulatory sandboxes, Living Labs, and testbeds. Currently, services related to testbeds are the most frequently utilized in experiments. However, AI innovators have expressed the greatest need for regulatory sandbox services,

indicating a mismatch between what is offered and what is needed. This suggests that initiatives are required to address this imbalance, particularly in the provision of AI regulatory sandbox services, to ensure TEFs fulfil their intended roles and meet their initial promises.

Key words (4 to 6 keywords)

Artificial Intelligence, TEFs, AI Regulatory Sandboxes, Living Labs, Testbeds, Experimentation

1. Introduction

The European Union has been leading the way towards regulating the current wave of Artificial Intelligence (AI) research and development (Truby et al., 2022). The rationale behind these initiatives is to provide AI innovators with clear instructions and requirements on the usage, implementation, and deployment of AI applications. As an outcome, a minimal degree of trustworthiness and transparency is desired, as well as mitigating the risks of undesirable outcomes from these AI applications. Concurrently, the EU aims to minimize administrative and financial barriers for AI innovators, particularly small and medium-sized enterprises (SMEs), recognizing that AI innovation in Europe is predominantly driven by these SMEs.
As these are two seemingly opposite goals, the EU also decided to establish several concurrent actions and initiatives to achieve both. One of the most prominent instruments are the so-called Testing and Experimentation Facilities (TEFs). These TEFs are defined as specialized large-scale reference sites open to all technology providers across Europe to test and experiment at scale state-of-the art AI solutions, including both soft- and hardware products and services, e.g. robots, in real-world environments.
In this paper, we focus on CitCom.ai, one of four TEFs in operation, that aims to bridge the gap between smart, sustainable cities, and communities on one hand and the AI innovators on the other. The goal of CitCom.ai is to create TEF sites that run commercial AI experimentation, testing and validation across different locations in Europe. These sites can take the shape of a private company, a public entity, or a mix of both. The main goal is validation in real conditions of novel, next-generation, AI-powered robotics and AI-based automation, decision-support, and decision-making tools. This is done by offering services to AI innovators that provide solutions for (smart) cities and their communities.

To position the role of these TEFs, three innovation concepts have been mentioned in the literature to fill this gap: AI regulatory sandboxes, Living Labs, and AI testbeds (Buocz et al., 2023). However, as these TEFs are still very much in a start-up phase looking for their own positioning, service offering and target customer definition, research is needed to assess whether these TEFs are able to help realizing the EU's double goal: increasing the trustworthiness,

transparency and security of AI solutions, and lowering the legal and administrative barriers for AI innovators in general and SMEs in particular.
Therefore, in this paper we want to explore the relation between TEFs and the three linked innovation concepts: AI regulatory sandboxes, Living Labs, and AI testbeds.
In order to investigate the potential role and positioning of TEFs in the European AI landscape, we adopted a qualitative triangulation design consisting of three data collection methods: an interview study among AI innovators, an analysis of the current service catalogue in the light of the three innovation concepts, and an analysis of the current usage of these services in planned or executing experiments. This design allows us to converge towards an initial understanding of TEFs by combining data from multiple methods and data sources.

2. TEFs in the context of the AI act & AI innovation

The EU wants to be a frontrunner in terms of the innovative development and application of Artificial Intelligence. To achieve this, the European Commission (2024) launched its AI strategy. This strategy consisted of a package, launched in April 2021, which included communication on fostering a European approach to AI, a review of the Coordinated Plan on Artificial Intelligence (with EU Member States), and a regulatory framework proposal on artificial intelligence and relevant impact assessment.
This strategy has four main goals:

- enabling the development and uptake of AI in the EU;
- becoming the place where AI thrives from the lab to the market;
- ensuring that AI works for people and is a force for good in society;
- building strategic leadership in high-impact sectors.

A main instrument, which is mentioned before as part of the regulatory framework proposal, is the so-called AI Act. This is the first-ever legal framework on AI, which addresses the risks of AI and positions Europe to play a leading role globally. The AI Act aims to provide AI developers and deployers with clear requirements and obligations regarding specific uses of AI. At the same time, the regulation seeks to reduce administrative and financial burdens for business, in particular for small and medium-sized enterprises (SMEs).

The AI Act is part of a wider package of policy measures to support the development of trustworthy AI, which also includes the AI Innovation Package and the Coordinated Plan on AI. Together, these measures will guarantee the safety and fundamental rights of people and businesses when it comes to AI. This is intended to strengthen uptake, investment, and innovation in AI across the EU.

The AI Act ensures that Europeans can trust what AI has to offer. While most AI systems pose limited to no risk and can contribute to solving many societal challenges, certain AI systems

create risks that we must address to avoid undesirable outcomes. Examples of such outcomes range from a chatbot from an airline company causing damage due to lies, the overflow of AI-generated information causing disinformation, an LLM hallucinating court cases, or the misappropriation of personalities' image and voice for commercial use. This European approach responds to the current state of AI regulation, which has mostly been managed on an ad hoc basis through individual member state laws and resolutions from various European Parliament committees (Truby et al., 2022). These actions are deemed necessary as there is more and more debate regarding AI and potential high-risk activities because of its nature. Currently, the discussion is mostly dealing with types of liability to be imposed under different conditions and circumstances.

The European Commission Proposal includes a risk-based approach to AI to determine how AI should be regulated in different cases. The proposal is for four categories of risk (Truby et al., 2022):

(1) Unacceptable risk, which is banned in the EU.

(2) High risk, which means that human health and safety or fundamental rights are endangered – mandatory requirements will be imposed on these types of AI, and they will be assessed to make sure that they comply.

(3) Limited risk, which imposes requirements for transparency in certain circumstances so users know they are interacting with a machine.

(4) Minimal risk, which allows other types of applications to be legally developed.

To make these large and complex ambitions a reality, the EC, together with the Member States, is co-funding TEFs to support AI developers in bringing trustworthy AI to the market more efficiently and facilitate its uptake in Europe. This is done via the Digital Europe Programme 2023-2024 via a Coordination and Support action (CSA) with the application of a cross-sector perspective to all existing sectorial Testing and Experimentation Facilities (TEFs). The selected TEFs started on January 1st, 2023, and focus on the following high-impact sectors:

- Agri-Food: "agrifoodTEF"
- Healthcare: "TEF-Health"
- Manufacturing: "AI-MATTERS"
- Smart Cities & Communities: "Citcom.AI"

Co-funding between the European Commission (through the Digital Europe Programme) and the Member States will support the TEFs for five years with budgets between EUR 40-60 million per project. TEFs are specialized large-scale reference sites open to all technology providers across Europe to test and experiment at scale state-of-the art AI solutions, including both soft-and hardware products and services, e.g. robots, in real-world environments. These large-scale reference testing and experimentation facilities will offer a combination of physical and virtual facilities in which technology providers can get support to test their latest AI-based soft- and hardware technologies in real-world environments. This will include support for full integration,

testing and experimentation of latest AI-based technologies to solve issues/improve solutions in a given application sector, including validation and demonstration. TEFs can also support market surveillance authorities and national competent authorities for both controlled testing of AI solutions, as well as direct collaboration with AI regulatory sandboxes. TEFs will be an important part of building the AI ecosystem of excellence and trust to support Europe's strategic leadership in AI. Limited literature is available to describe the nature and outlook of these TEFs, but they are stated to share common ground with concepts such as Living Labs, testbeds, and AI Regulatory Sandboxes (Buocz et al., 2023).

3. Regulatory sandboxes, Testbeds & Living Labs

AI Regulatory sandboxes

A regulatory sandbox is a tool allowing businesses to explore and experiment with new and innovative products, services, or business models under a regulator's supervision, providing innovators with incentives to test their innovations in a controlled environment. This in turn allows regulators to better understand the technology and fosters consumer choice in the long run. Establishing AI regulatory sandboxes on the EU member state level is an explicit part of the AI Act with as objective to reach a balance between innovation and regulation (Yordanova, 2019; Truby et al., 2022; Buocz et al., 2023). An AI regulatory sandbox creates an environment in which AI solutions can be tested and evaluated in order to increase reliability, trust, and acceptable risk assessment, which is facilitated by two main elements: specific procedures, involving support from experts and relevant authorities, and two-way communication between the AI innovators and the legal authority in order to reduce the burden on innovation while imposing relevant legal constraints to improve the adequacy and minimize the potential risks of AI innovation (Yordanova, 2019).

However, according to Buocz et al. (2023), to date only limited guidelines are available on how to implement, operate, or even describe the procedures of these AI regulatory sandboxes, while there remain several legal issues that are unresolved.

One of the main issues is that the current text blurs jurisdictional boundaries between the EU and member states which in turn raises concerns of legality and equal treatment for AI innovators and creates liability risks.

Moreover, in the current proposal the AI regulatory sandboxes are not necessarily uniform among the member states which creates confusion and uncertainty in the market, which is the opposite of the intention (Truby et al., 2022). Yordanova & Bertels (2024) add that the current guidelines explicitly emphasize the possibility of multi-jurisdictional regulatory sandboxes. The fact that the service lacks the standardization associated with regulation makes regulatory sandbox activities unfit for cross-border provision of services (Truby et al., 2022). Therefore, as a conclusion there are currently more questions than answers on how to regulate AI, and AI regulatory sandboxes could be a cornerstone to build uniform regulation, but in the current guidelines this is not the case yet (Yordanova & Bertels, 2024). Summarizing, multiple questions remain unanswered

(Truby et al., 2022): How to deal with testing of AI innovations in multiple member states? How to transfer the lessons learned from one national testing site to another national testing site? How to easily replicate tests and experiments in different European countries?

Living Labs & AI

Living Labs, recognized by the European Network of Living Labs (ENoLL) as 'Open Innovation Ecosystems' have emerged as dynamic innovation intermediaries. Based on iterative feedback processes, Living Labs aim at creating sustainable impact and provide real-life environments for testing and co-creating innovations (Leminen, Westerlund, and Nyström, 2012). To this end, they orchestrate stakeholder networks across the Quadruple Helix, involving government, research institutes, companies, and citizens. This orchestration occurs at multiple levels, with a specific focus on the organizational level, where they manage, monitor, and coordinate different LL projects (Schuurman, 2015). In their early days, LLs were mostly linked to ICT innovation, with a heavy emphasis on the European context of these evolutions (Eriksson, Niitamo, and Kulkki, 2005).

In the light of these evolutions, the current Testing & Experimentation Facilities share a lot of common ground with these early Living Labs, where 'ICT innovation' can be replaced by 'AI innovation.' As AI is a hot topic in innovation research, policy making and beyond, it is maybe surprising that in the context of Living Labs, not a lot of research attention has been dedicated to the combination of both. In the literature, we find examples of AI among healthy ageing Living Labs (Rauschenberg et al., 2021) and AI applications in Urban Living Labs (see e.g. Nguyen et al., 2022 and Frey et al., 2022).

The closest account of the role of Living Labs for AI innovation can be found in the work of Vilarino (2022), but this is mostly an exploration and plea for the pivotal role Living Labs should play in the context of Digital Transformation and AI innovation, focusing on the social and societal impact of AI, rather than a study that analyses the actual impact of Living Labs on AI.

When we look into practical examples where AI and Living Labs are combined, we find a majority of AI applications in the context of university research groups that want to combine and extend their technological capabilities in a more societal context (e.g. the AI Living Lab at the Brunel University London) or AI as enabler in specific domains, without Living Labs having a specific focus on AI itself (see www.openlivinglabs.eu).

However, all these research and practitioner examples deal with Living Labs linked to a single location or organization. This differs quite significantly with the concept of TEFs where the facility consists of multiple TEF sites spread across Europe with different service offerings and even different legal entities. In this regard, there are a few studies that explored the multifaceted roles of LLs within innovation ecosystems or networks. For instance, Gamidullaeva (2018) envisions LLs as crucial innovation intermediaries, fostering extensive networks and ensuring continuous integration. They adopt a critical role in coordinating innovation activities among multiple actors at the systemic level, eliminating barriers, and harmonizing the efforts of ecosystem participation. Within TEFs, this ecosystem should be regarded as the whole of Europe.

Testbeds & AI

A testbed is a platform for conducting rigorous, transparent, and replicable testing of scientific theories, computing tools, and new technologies. The term is used across many disciplines to describe experimental research and new product development platforms and environments. A typical testbed could include software, hardware, and networking components. In the context of AI, there is a long tradition of setting-up and using testbeds. For example, a report from 1995 describes the development and demonstration of the Advanced Artificial Intelligence Technology Testbed (AAITT), a structured development paradigm and associated toolkit supporting the design, analysis, integration, evaluation, and execution of large-scale, complex, distributed systems, composed of knowledge-based and conventional components, in the context of various United States Air Force domains. In this report, a testbed is defined as a facility that provides tools for experimenting with software system configurations in order to optimize performance and solutions (Zapriala et al., 1995). There is an abundance of studies describing testbed settings that make use of AI, for example in the context of 5G and 6G connectivity (e.g. Nahum et al., 2020 & Wang et al., 2023). There are fewer studies that report on testbeds for AI innovation in the context of cities. One example is a study by Meta et al. (2021) where the camp Nou stadium in Barcelona is used as a testbed for an Urban Digital Twin by using AI methods and techniques. However, the study mainly described the set-up and ambitions of this setting without a lot of data on outcomes or learnings yet.

4. Method

To investigate the potential role and positioning of TEFs in the European AI landscape, linked to the three discussed innovation concepts, we adopted a qualitative triangulation design. We chose the CitCom.ai TEF on smart cities and communities as case study and used three data collection methods: an expert interview study, an analysis of the current service catalogue in the light of regulatory sandboxes, Living Labs and testbeds, and an analysis of the current usage of these services in planned or executing experiments. This design allows to converge towards an initial understanding of TEFs in general and the CitCom.ai TEF in particular by combining data from multiple methods and data sources (Flick, 2004).

For the expert interview study, we selected 14 representatives from relevant organizations. Both the public and private sector were targeted, as both sectors can participate in the TEF as customers, adopters, users, researchers, etc. Once a participant agreed to be interviewed, they were requested to sign an informed consent form that provided further details on the data anonymized data processing. Table 1 gives an overview of the participants. The goal of the interviews was twofold:

- Get a good understanding of the hurdles AI innovators experience to reach commercialization by focusing on their needs and challenges for AI deployment in the smart city context.
- Discover, define, and validate predefined assumptions on the optimal set of services that should be offered through the TEF to tackle the above needs and challenges.

Table 1 Participants in the interviews and their main activities

Participant	Category	Main activities
Agoria	Federation	Agoria includes over 2,000 member companies from the manufacturing, digital, and telecom industries, with 70% being small and medium-sized enterprises (SMEs).
BeMobile	Large enterprise	BeMobile has evolved into a comprehensive smart mobility company, providing solutions for every sector of the mobility industry (e.g., logistics firms, tolling operators, original equipment manufacturers (OEMs), government bodies, port authorities, and parking operators).
Cevi	Large enterprise	Cevi is a major player in the Flemish IT market, specializing in building modern and reliable computer systems, primarily for public institutions (provincial governments, municipalities, Public Centres for Social Welfare, libraries, etc.), but also for other companies.
Curaevia	Start-up	CuraeVia helps organizations move towards a better, more sustainable customer journey with the aid of their Mobility as a Service (MaaS) software.
Digitaal Vlaanderen	Flemish government	Digital Flanders is the digital partner for Flemish and local governments. They implement digital transformation projects and guide governments through every step of the process.
Eurocities	Federation	Eurocities is the premier network of major European cities, encompassing over 200 large cities as members from both inside and outside the European Union. It serves as a collaborative hub where local governments convene to exchange ideas and collaborate on initiatives that drive positive change.
Hub.brussels	Brussels goverment	hub.brussels orchestrates and energizes various thematic networks of companies, working collaboratively to transform Brussels into the most vibrant, inspiring, and promising region for entrepreneurship.
IDLab	Academic research group	IDLab centers its research on internet technologies and data science. They create technologies that surpass current solutions in communication subsystems, high-speed and low-power networking, distributed computing and multimedia processing, as well as machine learning, artificial intelligence, and web semantics.
Mpact	Non-profit	Mpact strives to enhance mobility systems to be more efficient and accessible by adopting the principle of 'doing more with less'.
Paradigm.brussels	Brussels government	Paradigm is the go-to IT partner in the Brussels-Capital Region, capable of handling any task related to computer development, IT support, telematics, and cartography.
Roadeo	Start-up	RoadEO is creating a worldwide platform for monitoring and predicting road quality, which advises public road authorities and commercial construction companies on the optimal timing and location for road maintenance.
Smartends	Start-up	SmartEnds combines IoT software and hardware to transform the waste management industry. Their leading product, BrighterBins, utilizes advanced technology and highly optimized sensors to facilitate smart waste management.
NMBS/SNCB	National railway company of Belgium	NMBS/SNCB is the public railway operator and manages the freight (B-Cargo) and passenger services.
Univrses	Start-up	Univrses is committed to enhancing urban infrastructure through the application of computer vision and machine learning technologies.

A comprehensive topic guide was developed to ensure uniformity in the interviews while providing enough flexibility for detailed discussions. The approach of semi-structured interviews by use of a topic guide is supported by literature (Bernard, 2006). The topic guide is added as addendum to this paper.
The responses from the interviews were transcribed, coded, and categorized to identify distinct statements, themes, and keywords. Subsequently, the frequency of these statements was calculated by counting their occurrences. This frequency serves as an indicator of the level of attention given to specific topics, practices, or challenges. This method of coding and counting is commonly employed as a systematic technique in qualitative research (Elliott, 2018).

As a second method, we analysed the current service catalogue of CitCom.ai and linked them to the characteristics of the three discussed innovation concepts: regulatory sandboxes, Living Labs, and testbeds. We did this by comparing the elements mentioned on the CitCom.ai website where the service catalogue is presented to the characteristics that were discussed in our literature review. This deductive approach allowed us to link the seven service categories to the three innovation concepts.

Finally, we also collected data from all 16 CitCom.ai TEF sites on the experiments they were planning or executing. This was collected via a template that questioned the usage of (technical) infrastructure, the (planned) service(s) to be delivered to the customer, the type of customer, the TRL-level of the AI solution, the stakeholders involved, etc. The content of the template was discussed and iterated during three workshops that took place in the context of the project together with all site owners and other stakeholders involved in the CitCom.ai project. Unclarities were discussed and resolved, and feedback was considered to clarify the different elements from the template.
These templates will be used during the entire duration of the project and will be updated when the status of the experiment changes. For this analysis, we have used the data collected from the 26 first experiment templates. For this study, we counted the different services that were mentioned in the experiment templates and looked at the popularity of the seven service categories. Coupled to the previous exercise, this enabled us to identify the popularity of the service categories linked to the three innovation concepts.

5. Results

Interview study
The expert interview study, that was carried out with the participants having no prior knowledge of the CitCom.ai TEF, revealed the following barriers and enablers, ranked first, second, and third:

1. **Barrier: Innovation aversion**

Cities and communities often lack expertise in data, technology, and AI, leading to a slow and unclear decision-making process for adopting innovations.

Enabler: Innovation support

Cities and communities would benefit from enhanced support in innovation, which would improve their risk mitigation and reduce their 'fear of innovation.'

2. **Barrier: Data silos**

Today, data is often scattered, unstructured, and fragmented across silos, which restricts the potential for innovation.

Enabler: Common data(sharing) standardizations and frameworks

Common standards and frameworks would enhance data interoperability and maturity, which are crucial for advancing AI innovation.

3. **Barrier: Complex regulations**

A lot of stakeholders struggle to understand, comply, and keep up with regulations, such as GDPR (General Data Protection Regulation), the data act, the AI act, etc.

Enabler: Regulatory support

AI innovators require assistance to navigate through national and EU legislation, particularly as new directives have emerged in response to advancements in AI technology.

Based on the feedback of the interviewees, the following initial CitCom.ai services were ranked, with the most desired service category ranked first:

1. Guidance to understand EU legal framework
2. Providing a regulatory sandbox
3. Access to real-life TEF data
4. Algorithm training using real life TEF data
5. Algorithm validation in a real environment
6. Data requirements mapping

The 'data hunting for required data sets' service was not ranked as this was not desired by the participants. Additionally, they mentioned three more services that were missing from the initial predefined set of services:

- Matchmaking process between AI innovators, cities, and communities
- Community building to foster data sharing and trust
- Improving data and algorithm maturity, quality, and interoperability

Moreover, the participants also proposed as a condition to use TEF services that the datasets on offer are integrated, maintained and monitored.

Service category segmentation

Based on the interview study and on input from the operators of the 16 TEF sites, a new service catalogue was created. These 16 TEF sites are spread across Europe (Denmark, Finland, Sweden, Belgium, the Netherlands, Luxembourg, France, Spain, Italy, and Poland) and have a different focus, set-up, and characteristics. The differences between the TEF sites are intended to increase the possibilities for real-life validation and maximize the value proposition which in return will increase customer satisfaction for the CitCom.ai TEF overall. Some of the TEF sites promote themselves as active Living Labs (e.g. DOLL - the Danish Living Lab for smart and sustainable urban innovation) and already deliver services, whereas others are research institutes or RTO's, of which some also already offer services, and finally there are organizations linked to a municipality, which mostly do not offer paying services themselves. Multiple organizations can offer services on a TEF site. However, a service provider is not necessarily a TEF site. A TEF site consists of at least one testing zone, being a physical location where experiments take place or that is studied or affected by the experiment, usually located in a city, with experiments taking place, affecting, or studying the real environment. For more low-level TRL experiments, this testing zone can be a more lab-like environment (e.g. the dynamic vision lab of DTI, Denmark) or in specifically constructed testbeds that are closed for the general public (e.g. the autonomous vehicle test track at UTAC, France). Usually, there will be specific datasets available linked to the testing zone. A TEF site is usually linked to and situated in a city or municipality. The local ecosystem of a TEF site includes all the actors and their interactions linked to the specific topic or focus of that site. At least one of these actors owns datasets that can be used in experiments. Depending on the type of site, the TEF site can offer different services based on the AI Innovators' needs. Some of the TEF sites support the AI Innovators going from TRL 6-7, whereas others offer services that are more fitting from TRL level 7-8 or even TRL 8-9. Officially, TEFs are established to offer services only from TRL level 6 or higher (European Commission, 2023).

The inputs of the interview study were used to categorize the long list of services that were proposed by the TEF sites. Based on a deductive segmentation exercise, the following seven service categories have been proposed to reflect the capabilities and service offering of the 16 TEF sites:

Table 2 TEF service categories – see also www.citcom.ai

1. PHYSICAL FACILITY SERVICES Physical services enable AI innovators to test in city infrastructure with guidance, street closure, and installation support, reducing deployment time and advancing Technology Readiness Levels (TRL). A select group of nodes facilitates faster collaboration by streamlining the interaction between AI innovators and cities. Access to real infrastructure enhances TRL by validating solutions against real-life scenarios. Citizen engagement involves ecosystem participation for tailored testing and validation.	**2. VIRTUAL FACILITY SERVICES** Virtual facility services provide easy access to computing resources, supporting remote experiments and real-time data analysis. This accelerates solution validation, advancing Technology Readiness Levels.
3. ALGORITHM CREATION & VALIDATION Algorithm creation and validation services can be offered in collaboration with AI innovators. The co-created algorithms can validate others in a feedback loop, involving data collection, storage, preprocessing, machine learning frameworks, model training, evaluation, hyperparameter tuning, deployment, scalability, and monitoring. The process prioritizes security, privacy, and continuous improvement through a feedback loop, ensuring effective and efficient AI algorithm development aligned with specific project requirements and preferences. The comprehensive approach addresses data quality, infrastructure, and performance metrics, promoting seamless integration, encryption, and access control, enhancing the deployment and performance of AI algorithms.	**4. COMPLIANCE & ETHICS ASSISTANCE** Compliance assistance involves audits, program development, and training to mitigate risks, with experts adapting practices to regulatory changes. Ethics assistance promotes ethical behavior, guiding data use and developing codes of conduct. These efforts play a key role in proactive risk management, limiting legal and ethical risks, regulatory violations, and financial losses. These services are crucial for understanding and complying with evolving EU landscape requirements. In virtual facility services, security and compliance prioritize data protection through robust measures like encryption, access controls, and authentication.
5. IMPACT ASSESSMENT This service category focuses on the desirability and viability of AI innovation. While most TEF services assess feasibility, there is a crucial need to evaluate the actual impact on the environment, stakeholders, and end-users. Desirability relates to meeting stakeholder needs and solving problems, while viability considers the business model, revenue generation, costs, and benefits. TEF sites excel in answering these questions by tapping into local ecosystems, facilitating contact with relevant stakeholders, and addressing parameters linked to impact and outcomes.	**6. OPPORTUNITY ASSESSMENT & SCOPING** Services that help identify opportunities, define innovation scope, and align stakeholders, while also mitigating risks by evaluating financial viability, technical feasibility, and cities' needs. Namely: understanding target audiences, assessing organizational viability, and evaluating technical feasibility. Various activities are conducted based on idea maturity and customer requirements, including exploring, co-creating, and validating solutions to prepare for experiments. These services - focusing on customer needs and maturity, range from (i) idea exploration, understanding problems, assessing resources (ii) co-creation to characterize solutions and (iii) preparatory activities to select the best technical approach, including data, infrastructure, and expertise needs.
7. ECOSYSTEM ENGAGEMENT They focus on creating a collaborative platform dedicated to AI innovation that brings together researchers, academics, institutions, industry partners, and other stakeholders. The goal is to facilitate knowledge exchange, collaboration on research & development projects, and the development of a vibrant R&D community. Suppliers can showcase their technology and AI expertise, backed by capital investors looking for opportunities beyond state of the art. Cities and other potential buyers can share their interests by presenting real-life use cases for which they seek technology innovations.	

When we look back at the three innovation concepts, the first three service categories can be linked to the AI testbed. Within an AI testbed, the focus is put on rigorous, transparent, and replicable testing of new technology with a typical testbed consisting of software, hardware, and networking components. This clearly applies to services related to physical facilities such as sensors, city infrastructure, etc. The same goes for services related to virtual facilities such as High-Performance Computers, local Digital Twins... Thirdly, algorithm creation and validation is a service category that can also be associated with testbeds as the emphasis is on quality assessment, performance metrics... etc.

The fourth service category, compliance & ethics assistance, is clearly in line with the AI regulatory sandbox concept. This service category is all about (legal) compliance, risk assessment and codes of conduct for ethical AI that complies with the EU regulation. However, in the

literature review we already found out that the current legislation is not ready yet for effective implementation all over Europe and that some critical barriers still exist.
The remaining three service categories clearly go beyond a purely technical testing of the feasibility of the AI innovation or the compliance with regulatory requirements. Impact assessment looks at the actual impact of the innovation for stakeholders, whereas ecosystem engagement-services are intended to involve these stakeholders in the innovation process, be it in earlier stages to discover opportunities or in later stages to assess market potential. Finally, opportunity assessment and scoping refer to co-creating new AI innovations, possible after ecosystem engagement, and scoping the actual experiment that will run in the TEF. As already appears from the terminology, these final three categories link strongly to Living Labs with their emphasis on active stakeholder engagement and co-creation with the intention to create AI innovation that is desirable with a viable business case.

As a final step, we used the coded services from the experiment templates filled out by the TEF site owners. Based on the first 26 experiments that were logged in the respective templates we get a first impression on the nature and outlook of these experiments, and of the popularity of the service categories and the link with the concepts related to TEFs.
In the 26 first logged experiments, we already discover some interesting tendencies. First, in terms of TRL of the tested solution, we label the majority (N=15) as 'experiment' (TRL 5-6). The category 'test' (TRL 6-7) is mentioned 4 times and the category 'validation' (TRL 7-9) 7 times. This is an indication that the current TEF cases are still situated mostly in the experimental TRL-levels that are facilitated by TEFs. In terms of the main 'customer' of the experiments, half mentions SME's (N=13), followed by public sector (N=6), academia (N=5) and large enterprises (N=2). This is already in line with the assumed target group of the TEFs, which are mostly aimed at SMEs. However, other customer groups also seem to have potential.
In total, 53 services are listed within these 26 experiments, which makes a mean of just over 2 services per experiment. Eight experiments only list one service, and two experiments contain 6 services.

Table 3 Number of services per service category

Link to concept	*Service category*	*Number of times mentioned*
TESTBED SERVICES	1A. VIRTUAL FACILITY SERVICES	17
	1B. ALGORITHM CREATION & VALIDATION	10
	1C. PHYSICAL FACILITY SERVICES	8
LIVING LAB SERVICES	2A. OPPORTUNITY ASSESSMENT & SCOPING	6
	2B. ECOSYSTEM ENGAGEMENT	5
	2C. IMPACT ASSESSMENT	5
REG SB SERVICES	3A. COMPLIANCE & ETHICS ASSISTANCE	2

In terms of frequency, 'VIRTUAL FACILITY SERVICES' is clearly the most offered service with 17 experiments. 'ALGORITHM CREATION & VALIDATION' and 'PHYSICAL FACILITY SERVICES' are the second and third most offered service categories, with respectively 10 and 8 experiments. As we discussed in the previous paragraphs, these three service categories are in line with the AI testbed concept, and in the context of what the TEFs want to achieve, they seem a core offering towards AI innovators. Based on these numbers, there is demand as well as an offering for these service categories.

The next three service categories have very similar numbers: OPPORTUNITY ASSESSMENT & SCOPING (6), ECOSYSTEM ENGAGEMENT (5) and IMPACT ASSESSMENT (5). Again, this is perfectly in line with the innovation concept which they are all three linked to: Living Labs. These three service categories can be regarded as an attempt to link AI solutions with actual needs and (wicked) problems of cities and communities, and assessing whether they actually deliver, whereas the first three service categories could be linked mainly to the technical feasibility of the AI solution and apply more to the AI innovator only.

The final service category, 'COMPLIANCE & ETHICS ASSISTANCE,' has the lowest number of occurrences with only two experiments. This category is clearly linked to the final concept: AI regulatory sandboxes. In a way, it does not come as a surprise that this category is the least developed right now as the different member states are still in the process of setting-up regulatory sandboxes. The AI act specific regulation passed only recently. This will prompt many AI providers to seek assistance in understanding and applying this regulation. In the future we foresee that these types of services, and specifically the 'regulatory learning' (Gonzalez Torres & Sawhney, 2023), will become a very important part of the value proposition of the TEFs, especially as our expert interview study revealed that this type of services was one of the most mentioned and needed.

6. Conclusion

Within our research, we discovered that three more established innovation concepts can be linked to the newly established TEFs: regulatory sandboxes, Living Labs, and testbeds. Based on the actual service offering in the 16 different TEF sites of CitCom.ai and on the current popularity of these service categories in the first 26 defined experiments, the TEF is mainly operating as a testbed with 'technical' services making up the majority of the current experiments in preparation or executing mode. However, the results of the expert interviews clearly indicated the need for 'regulatory' services, which stresses the potential role of TEFs for regulatory support and sandboxes, although these services are almost absent in the current experiments.

Additionally, the results showed that several services were missing from the predefined set: community building and matchmaking, which should tackle the 'innovation aversion' barrier; and improving maturity, quality, and interoperability, which should tackle the 'data silos' barrier. Finally, we observe a clear need for any potential combination of the available services in the TEFs, showing the added value of grouping those services into a modular single point of access.

Therefore, our research demonstrates that TEFs can play a vital role for AI innovators as trusted intermediaries and innovation incubators, but that getting the regulatory aspects clear and implemented in their service offering is crucial. This also positions TEFs in a driving seat for the definition and implementation of the AI regulatory sandboxes. The fact that different sites in different member states work together and in sync in these TEFs holds a lot of potential to overcome some of the current hurdles within the AI act surrounding the establishment of these regulatory sandboxes. However, the fact that these sandboxes are not yet in place also puts a large risk on the potential success and scaling of the current TEF service offering. Without these sandboxes, international scaling and cross-border experiments are a lot harder to accomplish, so we suggest dedicating time and resources to figure out the establishment and service offering of the AI regulatory sandboxes via the 16 TEF sites in their respective countries. However, we believe that the upcoming Coordination Support Action on Regulatory Sandboxes (see European Commission, 2024b), as well as other similar initiatives within the TEF's and interested working groups, will allow interesting approaches and development on AI regulatory sandboxes novel implementations and operationalization.

Future research should look on the evolution of this TEF but should also consider the establishment and evolution of the other sectorial TEFs. More input is needed regarding market needs, the success of the different service categories and the legal implementations of AI regulatory sandboxes. We believe that more of this kind of research, which is very closely following the current European developments and evolutions, is needed to increase the chances of a successful implementation and operation of TEFs in the European AI ecosystem.

Notes & references

This paper was written in the context of the CitCom.ai project which has received funding from European Union's Digital Europe programme under the Grant Agreement No 101100728. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union Neither the European Union nor the granting authority can be held responsible for them.

Buocz, T., Pfotenhauer, S., & Eisenberger, I. (2023). Regulatory sandboxes in the AI Act: reconciling innovation and safety? Law, Innovation and Technology, 15(2), 357-389.

Elliott, V. (2018). Thinking about the coding process in qualitative data analysis. The qualitative report, 23(11), 2850-2861.

Eriksson, M., Niitamo, V. P., & Kulkki, S. (2005). State-of-the-art in utilizing Living Labs approach to user-centric ICT innovation-a European approach. Lulea: Center for Distance-spanning Technology. Lulea University of Technology Sweden: Lulea, 12.

European Commission (2023). https://digital-strategy.ec.europa.eu/en/activities/testing-and-experimentation-facilities. Accessed 02/05/2024.

European Commission (2024a). https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence. Accessed 15/05/2024.

European Commission (2024b). https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/digital-2024-ai-act-06-sandbox. Accessed 30/05/2024.

Flick, U. (2004). Triangulation in qualitative research. A companion to qualitative research, 3, 178-183.

Frey, C., Hertweck, P., Richter, L., & Warweg, O. (2022). Bauhaus. Mobility Lab: A Living Lab for the Development and Evaluation of AI-Assisted Services. Smart Cities, 5(1), 133-145.

Gamidullaeva, L. A. (2019). Inter-organisational network structures and knowledge diffusion through innovation intermediaries. Global Business and Economics Review, 21(6), 756-776.

Gonzalez Torres, A. P., & Sawhney, N. (2023). Role of Regulatory Sandboxes and MLOps for AI-Enabled Public Sector Services. The Review of Socionetwork Strategies, 17(2), 297-318.

Leminen, S., Westerlund, M., & Nyström, A. G. (2012). Living labs as open-innovation networks.

Meta, I., Serra-Burriel, F., Carrasco-Jiménez, J. C., Cucchietti, F. M., Diví-Cuesta, C., García Calatrava, C., ... & Eguskiza Martínez, I. (2021). The camp nou stadium as a testbed for city physiology: a modular framework for urban digital twins. Complexity, 2021, 1-15.

Nahum, C. V., Pinto, L. D. N. M., Tavares, V. B., Batista, P., Lins, S., Linder, N., & Klautau, A. (2020). Testbed for 5G connected artificial intelligence on virtualized networks. IEEE Access, 8, 223202-223213.

Nguyen, H. T., Marques, P., & Benneworth, P. (2022). Living labs: Challenging and changing the smart city power relations? Technological Forecasting and Social Change, 183, 121866.

Rauschenberg, C., Schick, A., Goetzl, C., Roehr, S., Riedel-Heller, S. G., Koppe, G., ... & Reininghaus, U. (2021). Social isolation, mental health, and use of digital interventions in youth during the COVID-19 pandemic: A nationally representative survey. European Psychiatry, 64(1), e20.

Schuurman, D. (2015). Bridging the gap between Open and User Innovation? exploring the value of Living Labs as a means to structure user contribution and manage distributed innovation (Doctoral dissertation, Ghent University).

Sherwood, R., Gibb, G., Yap, K. K., Appenzeller, G., Mckeown, N., & Parulkar, G. (2010). Can the production network be the testbed? In 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10).

Truby, J., Brown, R. D., Ibrahim, I. A., & Parellada, O. C. (2022). A sandbox approach to regulating high-risk artificial intelligence applications. European Journal of Risk Regulation, 13(2), 270-294.

Vilariño, F. (2022). Unveiling the social impact of AI through living labs. In Artificial Intelligence and Innovation Management (pp. 233-257).

Wang, C. X., You, X., Gao, X., Zhu, X., Li, Z., Zhang, C., ... & Hanzo, L. (2023). On the road to 6G: Visions, requirements, key technologies, and testbeds. IEEE Communications Surveys & Tutorials.

Yordanova, K. (2019). The shifting sands of regulatory sandboxes for AI.

Yordanova, K., & Bertels, N. (2024). Regulating AI: Challenges and the Way Forward Through Regulatory Sandboxes. Multidisciplinary Perspectives on Artificial Intelligence and the Law, 441.

Zapriala, J., Geigel, T., Hollingsworth, D., Mendenhall, H., Whitebread, K., & MARTIN MARIETTA CORP MOORESTOWN NJ. (1995). Advanced Artificial Intelligence Technology Testbed (p. 0117).

Addendum – topic guide

The interview was organized into four sections, each progressively introducing the concept of TEFs:

1. Introduction

First, a short introduction of the interviewers and the CitCom.ai project was given, detailing the research objectives and the reasons for seeking the interviewee's input. The goal was to get a better understanding of the interviewee's expertise and their position within their organization.

2. Barriers and enablers

To get insight in the current practices and challenges, the following questions were asked:

a. What are your current practices to train, test & validate your AI solutions?

b. What are the main challenges today in the mobility industry from your point of view?

c. What role do you see AI play in resolving these issues?

Next, the interviewers inquired about which services should be implemented in the CitCom.ai TEF to address the previously mentioned needs, ensuring the questions were posed without interviewer suggestions to minimize bias.

3. Innovation confrontation

In this section the concept of TEFs, and the scope of the CitCom.ai TEF specifically, was introduced. The interviewer explained the role of the TEF as an orchestrator that fosters collaboration between AI innovators and cities by offering services. The following characteristics of the CitCom.ai TEF were highlighted:

a. A one-stop-shop for companies to improve their processes, products or services using digital technologies.

b. A facilitator for the validation of novel AI-driven services in real-life environments before their further massive deployment

c. Expertise centre for the design and the implementation of AI testing methodologies in real-world environments

Here again, the interviewees were asked to reflect on the solutions to their challenges, but this time focusing on the following predefined TEF services:

a. Providing a regulatory sandbox
b. Guidance to understand EU legal framework
c. Data hunting for required data sets
d. Access to real life TEF data
e. Data requirements mapping
f. Algorithm training using real life TEF data
g. Algorithm validation in a real environment

4. Wrap up and conclusions

The interview was concluded with an opportunity for the interviewees to offer additional insights or takeaways that were not yet discussed.