A framework for the evaluation of living labs as boundary spanners in innovation

Marina van Geenhuizen
Faculty of Technology, Policy and Management, Delft University of Technology, the Netherlands

Abstract
Living labs, as a methodology to enhance user-centric innovation, have large potentials in bringing inventions to the marketplace, but their performance can benefit more from evaluation. This article develops a novel framework for evaluation of living labs, including (1) a system approach providing an analytical view on living labs’ performance and results; (2) a focus on actor-complexity and boundary-spanning needs; (3) a set of questions concerning, e.g. absorption of user-feedback, satisfaction among actors, and openness and connecting with larger networks; (4) a list of key performance factors; and (5) a focus on participatory evaluation. The design of this evaluation framework rests on a comprehensive literature search and case studies representing different actor complexity, namely home-solutions in healthcare, reconstruction of large (multi) functional buildings, and multiple combinations of activity (university campuses). Key performance factors are found to be: an early involvement of adequately skilled users in multiple learning processes, including absorption of feedback, and a broader but balanced set of actors connecting with upscaling and acceptance in the market. Also, boundaries need to be better bridged by learning how to handle conflicts and deal with intermediation, while respecting shared goals and interests. Specifically, university living labs call for maintaining a solid relation with cities and their actors. Overall, an explicitly designed evaluation framework is a key part of the working plan of living labs. The results also indicate a need for stronger attention for boundary-spanning in evaluation, because living labs are increasingly applied in comprehensive multi-activity settings.

Keywords
Living labs, user-involvement, performance, evaluation, actor-complexity

Corresponding author:
Marina van Geenhuizen, Faculty of Technology, Policy and Management, Delft University of Technology, PO Box 5015, Delft 2600 GA, the Netherlands.
Email: m.s.vangeenhuizen@tudelft.nl
Setting the scene

Many countries today devote strong attention to the knowledge economy and an appropriate application of newly created knowledge. The Triple Helix – university, business and government, or Quadruple Helix – including user groups – are important networks to bridge different ‘worlds’ (Etzkowitz, 2008; Etzkowitz and Leydesdorff, 1998). Despite improvements, boundaries and divisions between these ‘worlds’ can still be observed (e.g. Bruneel et al., 2010; D’Este and Patel, 2007; Kaufmann and Tödtling, 2001; Leydesdorff and Meyer, 2007; Ranga and Etzkowitz, 2013; Taheri and van Geenhuizen, 2016; Van Geenhuizen, 2013).

Living labs can be described as a methodology of innovation that enables collaborative learning by users, producers and researchers in a real-life environment, in which user-needs are central (Almirall et al., 2012; Del’Erra and Landoni, 2014; Leminen et al., 2014, 2015; Schuurman, 2015). With regard to use of terminology, aside from innovation methodology, the term living labs often also refers to the (temporary) organizational structure in which the methodology is implemented, and this is adopted in the remaining article.

Living labs have been positioned and categorized in different ways, e.g. on the basis of involvement of users – as co-creators versus subjects or passive actors – and the environment – as real-life versus lab-like environments (Almirall et al., 2012). Another positioning is concerned with user involvement, in terms of open versus closed, and exploiting versus exploring opportunities (Dell’Era and Landoni, 2014), and most recently type of innovation process – predefined versus iterative – and type of tools used – different standardization level (Leminen and Westerlund, 2017). The previous types of positioning not only indicate different categories of living labs, but also that living labs are ‘akin’ with ‘adjacent’ innovation methodologies. It is important to note that the different categories of living labs may require differences in approach to evaluation.

There is also an important variation in the setting of involved actors, running from relatively simple to complex (Hakkarainen and Hyysalo, 2013; Nyström et al., 2014), e.g. dependent on maturity of the technology, involvement of researchers, types of users and utilizers, enablers, etc. and driving roles, and number of different functions, the last stretching in scale from, for example, single living houses to museums, hospitals, shopping malls and parts of universities, airports and cities. Large-scale living labs, typically, contain smaller ones and connect a larger number of actors with different interests which increases complexity and needs for boundary-spanning (Harvey et al., 2014; Katzy et al., 2012; Rittel and Webber, 1973; Williams, 2002). Actor complexity is selected in this article as the leading dimension in developing an evaluation framework.

Living labs have been in place since the early 2000s and are mushrooming today, calling for more attention for evaluation. The aim of this article is to identify the characteristics of a framework for evaluation of living labs, in particular, its approaches, focus and key performing factors. While in knowledge transfer policies, attention has been paid to university transfer offices, incubators and knowledge brokers, etc. living labs have largely remained out of this range (Howells, 2006; Meyer and Kearns, 2013; Todeva, 2013). Rather, studies on living labs have often arisen from application of new information and communication technology, for example in healthcare services (ICT) (e.g. Eriksson et al., 2005; Følstad, 2008; Ståhlbröst, 2008) and have gradually spread to other domains, like rehabilitation of buildings and improving transport nodes (e.g. Kehaya et al., 2014; Salter and White, 2013; Ståhlbröst, 2012).

Systematic studies on conditions that enhance the reaching of aims of living labs are strongly increasing in number today. The following ones can be mentioned: Veeckman et al. (2013), Nyström et al. (2014), Schuurman (2015), Leminen (2015), Leminen et al. (2015,
2016), Logghe and Schuurman (2016), and Ståhlbröst and Host (2016). These studies focus on actors’ roles, innovation tools or learning processes. The current article takes a system approach to living lab performance and proposes an ‘overarching’ list of key performance factors, which is novel. Further, more comprehensive and multi-activity living labs in the study are exemplified by universities (campuses). Attention for university living labs with their potential to ‘harness’ large amounts of knowledge in solving urban sustainability problems, is relatively new (Evans, 2015, 2017; Evans et al., 2016; Forbes, 2012; König, 2013) and conforms a stronger societal role of universities (e.g. Bretznitz and Feldman, 2012; Goddard and Valence, 2013; Trencher et al., 2013).

Living labs require boundary-spanning between the participating organizations which have different aims, interests and cultures, but intend to learn collaboratively (Bjerregaard, 2010; Harvey et al., 2014; Nooteboom, 2009; Williams, 2002). Living labs per definition create collaborative learning between users, produces, researchers and utilizers, etc. and, in a broader network, with financial investors, regulators and (local) policy makers, citizen groups, etc. Some of the learning processes, however, reveal badly known mechanisms and results related to divides or borders between actors involved (e.g. Ernst and Chrobot-Mason, 2011; Hakkarainen and Hyysalo, 2013; Logghe and Schuurman, 2016; Ståhlbröst and Holst, 2016). Against this backdrop, the research question is as follows: What are characteristics of an evaluation framework for living labs; in particular, what are the key performance factors and needs for boundary-spanning to be addressed?

First, the context of living labs and boundary-spanning are discussed and this is followed by the design of the study. In the section hereafter, the evaluation framework is introduced with attention to five characteristics. Various key performance factors and their implications are presented next. The article closes with a summary, some critical remarks and future research lines.

Context and needs for boundary-spanning

Living lab methodology is rooted in ideas emerging in the early 1990s about potentials of city neighbourhoods as learning environments for students that are engaged with solving real-world problems (Bajgier et al., 1991; Leminen, 2015). Living labs started to be elaborated in the early 2000s when research was moved from laboratories to in-vivo settings enabling observation of interaction of users with innovations in real-life, as particularly undertaken at Massachusetts Institute of Technology in the US by William Mitchell (Eriksson et al., 2005). Much early experience has been gained in the European Network of Living Labs (ENoLL), a platform established in 2006 to foster ICT-based innovations (ENoLL, 2014). Further, the increased attention for living labs as real-life applications fits two related waves of changes in science and knowledge production, one in the early 1990s when knowledge production tended to become more socially distributed, application-oriented and trans-disciplinary (Gibbons et al., 1994) and one in recent years bringing a stronger emphasis on openness and user-driven character (public sector and civic society) and data-intensive nature of knowledge creation in solving societal challenges in Europe (EC, 2012, 2014).

In the meantime, users had become recognized as an important source of innovation, and the customer-active paradigm subsequently fostered various models of co-creation (Prahalad and Ramaswamy, 2004; Von Hippel, 1986, 2005). Models of open innovation became also more popular in those years (Chesbrough, 2003; Chesbrough et al., 2006). And finally, new ideas and solutions on urban sustainability arose alongside new forms of urban governance and here the urban living labs made their appearance (Bulkeley et al., 2016; Voytenko et al., 2015). Universities joined with plans to establish living labs on their campus
while connecting with challenges of urban sustainability (Evans, 2015, 2016; Forbes, 2012; König, 2013).

Co-creation as the intended learning process in living labs, ideally, encompasses joint problem-definition and problem-solving using improvisation and experimentation – this in designing, implementation and testing of solutions in an iterative way, while interacting and co-constructing personalized experience (Prahalad and Ramaswamy, 2004). User-values are central, e.g. encompassing ergonomic, medical, socio-cognitive and socio-economic and cultural criteria. Learning in co-creation takes multiple approaches, like testing an improvement using interviews with users, designing future scenario’s, acting in focus groups, observing users’ behavior and requiring feedback from them, etc. (e.g. Dutilleul et al., 2010; Leminen and Westerlund, 2016; Ståhlbröst and Holst, 2016). However, tension and conflict between interests may arise, e.g. if power positions are different and eventually shifting. This may cause boundaries and borders hindering co-creation, while process managers help to avoid the rise of such situations and enable users to have an actual impact. De Moor et al. (2010), Hakkarainen and Hyysalo (2013, 2016), Logghe and Schuurman (2016), Ståhlbröst (2008) and Ståhlbröst and Host (2016) are among the first researchers who address such processes and mechanisms on the operational level of living labs.

When actors from different organizations develop co-creation, as is the aim in living labs, organizational boundaries enter the scene, like originating from differences in expertise, function, culture, ideology, power-position, markets, value-chains, etc. Today, many useful results of process analysis and leadership models, practices and tactics in boundary-spanning activity are available, e.g. Ernst and Chrobot-Mason (2011) and Harvey et al. (2014) and may support in boundary-spanning to create common language, trust, common ground and interests, and commitment (community) between actors involved (De Moor et al., 2010; Mohr et al., 2010; Williams, 2002).

**Design of the study**

The study draws on the extended literature on living labs, general literature on evaluation and boundary-spanning or intermediation, and on six case studies of living labs. The leading conceptual dimension is actor-complexity, in previous literature also addressed by e.g. Leminen (2013) on participation, coordination and driving actors. Actor-complexity and the connected boundary-spanning are preferred in this study, because sharing common aims, language and trust act as *sine qua non* for intended learning processes. Accordingly, case-studies that represent contrasting actor-complexity are used: There are two case studies on *person-oriented* living labs (elderly in ambient-assisted living, less complex), two in more complex *organisation-oriented* living labs (refurbishment of a hospital and of a shopping mall) and two in comprehensive *multi-activity* living labs (university campus). These types of living labs differ in number of distinct user types and other actors, and also in power position between participating actors. They illustrate, e.g., having elderly as a single user-group and having combinations of user groups and utilizers, e.g. patients, medical staff, visitors and hospital managers.

The selection of case studies representing the first two living lab categories is motivated by a sufficiently long existence of the living labs allowing for reflection on performance, and by views in literature and expert opinion. Regarding university living labs, representativeness for trends could only be assumed. Data and other information on the six living labs were derived by using a multi-source strategy, including practical reports (e.g. Guldemond et al., 2012; Kehaja et al., 2014; Kop, 2011; Vloed and Sadowski, 2013), communication and interviews with experts, output of workshops and conference meetings on healthcare living
labs in the Netherlands,\(^1\) and coverage of living labs’ websites, of which the workshops specifically enabled triangulation of results. In data collection and subsequent analysis of case material, a preliminary list of evaluation characteristics and performance factors was used mainly derived from literature, and this list was continuously checked and modified using case studies and additional data sources. The analysis has produced so-called ‘typical material’ of living labs – this as part of the iterative research in frames of ‘grounded theory’ (Eisenhardt and Graebner, 2007; Mayring, 2007). It thus provides insights representative for different actor-complexity and boundary-spanning needs in a theoretical sense. In the next section, characteristics of an evaluation framework will be discussed.

**An evaluation framework**

**Influences on living labs’ processes and results**

Perceiving living labs through the lens of system thinking is helpful in the identification of important influences on performance and results, as experienced in previous analysis of tools and policies in literature (e.g. van Geenhuizen and Thissen, 2002; Walker et al., 2001) (Figure 1). The system approach is discussed below by distinguishing between inputs, influences beyond control, learning and networking processes, and outcomes, the last as intended outcomes and unintended outcomes.

**Inputs** to the living lab create evolving learning processes, and these include the present motivation and capabilities of actors, particularly of users; sets of learning tools and models; specific expertise; financial budgets and other resources, as well as the real-life environment (e.g. Leminen, 2012; Veeckman et al., 2013). There are also inputs that serve a broader networking like in building commercial and regulatory relations. Methods and tools of living lab processes are somewhat under-searched, but more recently taken-up, among others by Hakkarainen and Hyssalo (2013) on conflicts, Nyström et al. (2014) on management of actor roles, Rits et al. (2015) on integration of business model methods, and Leminen and Westerlund (2017) on type of innovation processes (predefined versus iterative) and different tools (level of standardization). In addition, various influences affect living lab processes that are beyond control of managers of living labs (*exogenous influences*), like a shift in overall policy priority leading to budget cuts among stakeholders involved (Evans, 2017; Mohr et al., 2010).

![Figure 1. Simplified system approach to living lab processes (learning and networking).](image-url)
The outcomes of living labs are also influenced by the quality of the micro-level learning (co-creation) processes and the broader networking (D’Hauwers et al., 2016; De Moor et al., 2010; Leminen et al., 2015; Nyström et al., 2014). Outcomes in the sense of intended outcomes include a better knowledge of user wishes, a better user-quality of inventions or user-driven inventions, and speeding up and bringing more inventions to market (e.g. Schuurman et al., 2016; Veeckman et al., 2013). A better understanding of the necessary learning processes and feedbacks are often also seen as part of results of living labs (e.g. Kehayia et al., 2014; Logghe and Schuurman, 2016; Ståhlbröst and Holst, 2016). In the current approach, the term key performance factors is adopted for those inputs and (learning) processes without which the living lab would fail to reach intended outcomes.

Unintended outcomes may also happen. These include the absence of desired outcomes as well as adverse effects, and both are to a certain extent connected with bad predictability of influences and processes. Adverse effects may happen where trust vanishes, new boundaries and conflicts arise and collaboration is terminated, eventually causing some ‘disturbance’ in relationships beyond the living lab. To the author’s knowledge, adverse effects are documented in literature only in a few cases (e.g. Hakkarainen and Hyysalo, 2013; Ståhlbröst and Holst, 2016). Risks of failure, however, have been clearly addressed in the case studies in the current analysis, like short in respect for needs for self-determination among ageing people in healthcare, and the lack of networking of university living labs with regulatory organizations and city officials thereby delaying (a timely) upscaling of inventions. Also mentioned is the risk of access to the living lab by actors who disturb the original values of the living lab.

Levels, questions and participation

Evaluation is seen in the article as an inherent part of the living lab methodology while dealing with two levels, the micro-level of evolving learning processes (co-creation) and the meso-level of the connected networks (Leminen and Westerlund, 2015; Logghe and Schuurman, 2016; Ståhlbröst and Host, 2016). Attention is required to these levels as actor-complexity and its management tend to be different. The first deals with intensive co-creation of a few core-actors, while the second level provides useful input to co-creation from outside partners from different organizations, who sometimes act as insiders if necessary. Also, in large-scale and multi-activity living labs, the broader network that is interacting with the co-creation network gets large and more complex and, therefore, more difficult to manage.

At least five questions need to be addressed in the evaluation of which the first is ‘traditional’, while the other ones are typical for living labs: (1) are the product/service development and design process sufficiently on schedule (working plan and budgets); (2) are learning results from users (user feedback) sufficiently integrated into the design process (De Moor et al., 2010; Sauer, 2013); (3) do the designing actors remain sufficiently aligned with each other, with a common vision and common interests (e.g. Hakkarainen and Hyysalo (2013); (4) partially overlapping previous points, what is the satisfaction of the participant actors with the results and processes so far, and (5) is the living lab sufficiently open to attract partners in a broader network enabling support in upscaling and implementation (e.g. Evans, 2017)? Finding answers to these questions can benefit from action-research (Logghe and Schuurman, 2016; Reason and Bradbury, 2008; Ståhlbröst and Host, 2016). Action-research captures specific user-needs, problem perceptions, amount of satisfaction and perceptions on potential solutions in developing feedback, but also trust and integration among learning partners.
Evaluation is quite common in project or program management (Rossi et al., 2003; Wholey et al., 2010), and accordingly also in living labs. Preparing evaluation encompasses, e.g. developing an evaluation design, including ideas about the (causal) relation(s) that are subject to evaluation (key performance factors); data gathering and analysis; checking outcomes etc. Typical for the living lab methodology, however, is a participatory approach in which the actors are prominently present by elaborating questions, providing data and perceptions, and by justifying conclusions (Fawcett et al., 2003). Also, the living lab actors provide feedback to the evaluation results and need to agree on use of the results in improving the living lab. In problematic situations, evaluation results may serve as a basis for negotiation between actors in search for solutions (Kuhlmann, 2003). Overall, a model of participatory evaluation is an important input to the living lab methodology, though the amount of participation per actor group, as well as input from outsiders may differ, e.g. according to power positions of actors.

Another, complementary, evaluation method is more traditional and deals with base-line measurement at start of the living lab and compares this with situations later-on (ex-post), or alternatively, compares the living lab with a sufficiently similar ‘control’ situation without a living lab intervention (quasi-experimental) (Hyysalo and Hakkarainen, 2014). Dependent on needs for information on effectiveness, this type of evaluation may be included.

**Case study analysis**

**Introduction**

Four case studies are discussed, representing person-oriented and organization-oriented living labs, with an emphasis on co-creation and the broader networking. In addition, university living labs are discussed but these are relatively new; this is the reason why their analysis is different and more preliminary. Discussion of each of the types of living labs is concluded with an indication of the used evaluation.

**Person-oriented living labs**

The first project, in the region of Eindhoven (the Netherlands), targeted elderly people of Turkish origin (Table 1). The aim of the living lab was threefold in providing ICT tools, namely to enhance home care (low threshold Skype interface with care-providers), home fitness training with health improvement, and home safety (sneak-thief detection and emergency button), while adapting the technology solutions to the specific users. The actor complexity and boundary-spanning were cultural in nature following from barriers with the user group (Kop, 2011). The preparation on dealing with specific user-needs started already prior to project design as a specific input factor, supported by coaches from Turkish community, in some cases grand-children. Considering a broader network, business actors were not active, while relations with a university were limited to an ex-post evaluation study (Van der Vloed and Sadowski, 2013). In terms of outcomes, the target group became more involved in home fitness and indeed improved health conditions. They also accepted some ICT tools and suggested certain new home safety measures; however, their willingness to pay for ICT solutions appeared to be low.

Living Lab Amsterdam is a more extended (ICT solutions) and long-term version of the first living lab. The approach to collaborative learning was a mix of more and less active user involvement, including user interviews on tested applications, acting in focus groups and in collaborative experiments of specific applications. Experiments concerning observation of
elderly at home served the picturing of activities of daily life in terms of (in)dependence and need for support, and the identification of useful combinations with other services such as alarm systems, a mood button, etc. The main challenge in boundary-spanning was getting and keeping the elderly involved. Accordingly, trust was a specific input that could be created by using already established personal relationships between elderly and care professionals and by demonstrating working versions of the ICT solution before project start. Potential boundaries were related to privacy in ‘being observed’, lack of self-determination in switching-off the observation system, and not being aware of passing the point of no return in accepting in-house ICT (ARC, 2013). The main outcomes of this living lab are an increased acceptance of ICT tools for ambient-assisted living and home care, and a better insight into the wishes and values of elderly people. Business partners were almost absent in the networks; however, universities were strongly involved in learning processes, by providing management knowledge and also domain knowledge on the ICT solutions.

With regard to evaluation, the first living lab was subject to a formal evaluation by an outside institution after ending, while the second living lab was evaluated more implicitly on the way, e.g. by guiding the learning processes.

Table 1. Person-oriented living labs in senior houses (ambient-assisted living).

<table>
<thead>
<tr>
<th>Case study 1</th>
<th>Case study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Doornakkers: Eindhoven (NL)</td>
</tr>
<tr>
<td>Aim</td>
<td>Adoption of healthcare and illness prevention, using ICT for home care, fitness and home safety</td>
</tr>
<tr>
<td>User involvement in learning</td>
<td>Elderly of Turkish origin; passive role but could switch to active</td>
</tr>
<tr>
<td>Learning partners</td>
<td>Care provider; social housing provider; security services</td>
</tr>
<tr>
<td>Broader network partners (examples)</td>
<td>Municipality Eindhoven; Brainport (region); university</td>
</tr>
<tr>
<td>University involvement</td>
<td>External evaluator (ex-post)</td>
</tr>
<tr>
<td>Key boundary-spanning to enable learning</td>
<td>Overcome culture and age barriers, and create trust among users</td>
</tr>
<tr>
<td>Key learning factor(s) as inputs</td>
<td>(1) Preparation: study of user needs prior to project design; (2) Specific coaches to develop trust</td>
</tr>
<tr>
<td>Outcomes of learning</td>
<td>Increased use of ICT with better physical health condition of users</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Formal evaluation after ending</td>
</tr>
</tbody>
</table>

Source: Kop (2011); Amsterdam Region Care & ICT (2013); Van der Vloed and Sadowski (2013).
Organization-oriented living labs

Health innovation lab (HIL) is part of a larger initiative in the Copenhagen area, in Denmark (Table 2). HIL had the aim to design a methodology in hospital renovation/refurbishment drawing on input from user-driven methods and simulation, thereby reaching solutions that are scalable and transferable to connected hospitals in the region. The envisaged solutions were concerned with operation theatres and waiting rooms, but also with e-Health use at patients’ home. The complexity in boundary-spanning was two-fold requiring specific inputs. First, the multi-user situation, including hospital managers and medical professionals aside from patients, and secondly, the need for a good match with user capabilities/skills regarding application of simulation tools. Like in the previous case studies, the business world remained mainly off-side due to limited aims. In contrast, the university was

Table 2. Organization-oriented living labs in refurbishing of hospital and shopping mall.

<table>
<thead>
<tr>
<th>Case study 3</th>
<th>Case study 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Healthcare Innovation Lab, Copenhagen, Denmark</td>
</tr>
<tr>
<td>Working years</td>
<td>February 2010–2012</td>
</tr>
<tr>
<td>Aim</td>
<td>Design of refurbishing tools for hospitals and design rules for eHealth</td>
</tr>
<tr>
<td>User involvement in learning</td>
<td>Patients, clinicians and hospital managers (University Hospital Herlev) (strong patient involvement)</td>
</tr>
<tr>
<td>Learning partners</td>
<td>Regional hospitals</td>
</tr>
<tr>
<td>Broader network (examples)</td>
<td>Capital Region of Denmark and Danish Business Authority (both financial investors)</td>
</tr>
<tr>
<td>University involvement</td>
<td>Provider of domain knowledge and management knowledge (through university hospital)</td>
</tr>
<tr>
<td>Key boundary-spanning in learning</td>
<td>Training of user capabilities and creating strong team-building</td>
</tr>
<tr>
<td>Key learning factors as inputs</td>
<td>(1) Conscious user selection and training; (2) Management back-up across sectors; (3) Trust creation; (4) Multi-disciplinary input</td>
</tr>
<tr>
<td>Outcomes of learning</td>
<td>Sets of rules to which innovations need to respond; insight into management of new innovation tools</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Mainly implicit</td>
</tr>
</tbody>
</table>

strongly involved through its hospital, with input of domain knowledge and training in use of simulation tools, and team building, all with the aim to encourage all parties to interact and accelerate design. In more detail, boundary-spanning was enhanced by trust creation and open dialogue, emphasizing ‘human values’ like passion among managers. The relevant outcomes included sets of rules to which innovations in hospital design and e-Health need to respond, aside from understandings in innovation management (Ruff and Jacobson, 2012).

The main aim of reconstruction of Alexis Nihon shopping mall (Montreal, Canada) was allowing disabled people to resume their life and social integration, in particular shopping (Kehayia et al., 2014) (Table 2). Using multiple projects, first, obstacles and facilitators of participation in shopping by disabled persons were identified. Second, technology and interventions were developed in-vivo, and finally, these were implemented and their impacts evaluated. The focus was on design of better wheel-chair navigation and way-finding technology, in combination with novel refurbishment of the mall. The two main user-groups, disabled persons and rehabilitation services providers, thanks to specific inputs, could adopt different roles in mutual learning, like in joint experiments, focus groups, etc. In addition, commercial partners were closely involved because they had to bring the co-created solutions to the pilot stage. Enhancing a strong commitment and partnership of the core actors was also crucial as an input, because of diversity between the actors (Kehayia et al., 2014).

Unlike the previous case studies, this living lab was strongly supported by broader network activities, like in communities of practice, other participatory methods (Mazer et al., 2015) and business-related networks. Other key inputs, partly related to boundary-spanning, were a multi-disciplinary and multi-sector approach, ranging from construction technology to transport behavior, health and psychology. The outcomes can be summarized as co-created innovations in the envisaged technologies, alongside better understanding of the multi-disciplinary and multi-sector aspects.

Finally, regarding evaluation, HIL has adopted implicit ways of evaluation. This also holds for Alexis Nihon, but the last was also more formally evaluated, two years after start. This included comparing expected results with actual results through measuring satisfaction among the participating actors and their perceived importance of the results in daily life. This process supported a sound continuation of living lab activities (Mab-Mackay Rehabilitation Center, 2013).

Multi-activity living labs: Universities

Universities started to apply a living lab methodology since around 2010. Early examples are the University of British Columbia (UBC) in Vancouver (Canada) and the University of Manchester in Manchester (UK) (Evans et al., 2015). In university living labs, applied research and education are fostered by using the campus and its facilities to develop and test real-time solutions in energy sustainability, transport, food production, healthcare, etc., thereby offering opportunities to students and staff and interested actors in the city to learn in daily-life (Evans, 2015; Forbes, 2012; Graczyk, 2015). The variety in on campus experimentation is large, witness the following projects (e.g. Evans, 2015; Evans et al., 2015; Salter and White, 2013; UBC, 2016; University of Manchester, 2017): UBC’s FARM aimed at designing new solutions in food production and land-use in the frame of bridging urban and rural environments; Save, 2014; UBC’s Centre for Integrative Research on Sustainability (CIRS) aimed at experimental building where different utility systems are integrated; Manchester’s Ferranti Building where experimental integration of photovoltaic (PV) on-roof systems enables simulation and monitoring of use in domestic settings, and Manchester’s on-campus experimentation with cycling lanes.
University living labs are faced with many challenges in boundary-spanning. On the input-side, there is a practical bottleneck in human resource management (HRM) and responding to regulation (Evans, 2017). Real estate personnel gets different working tasks if living lab methodology is set up in existing buildings, causing new duties and legal responsibilities. Also, university living labs are often technologically complex, thereby increasing the chance of technical failure (UBC Insiders, 2015) and hindering replication of buildings and merits in the city (Evans, 2017). Replication may also fail due to the absence of active broader networks with local policy and with rules and regulatory organizations (standardization). A risk of a different type is the entry of an actor in the living lab or broader network that represents values which are not compatible with the original values of the living lab. As a final point concerning boundaries: some universities have on the input-side very favorable on-campus conditions. Such conditions, like self-financed projects, basic control over utilities and ownership of spacious lands (campus) at the urban fringe, are however difficult to match with real-life conditions in cities, and this makes learning on replication and upscaling that facilitates application and adoption, more complicated.

With regard to evaluation, it seems that some on-campus projects have been evaluated in an implicit way and somewhat fragmentary. More comprehensive and systematic evaluation tends to be under way, but is complicated due to the many different actors involved (Fawcett et al., 2003).

Key performance factors

As an overall result of literature and case study analysis, the conclusion is derived that designing an evaluation plan within a working plan that enables a continuous evaluation is a key performance factor in itself. Further, the remaining list of key performance factors can be divided into evolving learning processes and underlying values, broader networking, and the real-life environment (Table 3). The factors that relate to ‘pronounced’ actor-complexity are highlighted in the table (italic).

First, evolving learning processes (co-creation): there should be an early involvement of users, a timely preparation of dealing with vulnerable users, and a sufficient match of motivation and capabilities/skills among the living lab actors with the aim and methods (tools) of learning (Almirall et al., 2012; Leminen et al., 2012; Schuurman et al., 2015; Ståhlbröst and Holst, 2016). The use of multiple approaches and tools (eventually multi-disciplinary and multi-sector) is also important, as is flexibility in actor roles. Roles may be temporary and change with adaptation in the network and with new understandings (Nyström et al., 2014). What needs to be accepted is a certain unpredictability in learning processes and their outcomes which is inherent to experimentation and creative design and calls for some flexibility in actor roles. Specific attention is needed to the ability of participating actors to handle conflicts and work with intermediaries in solving them (e.g. Hakkarainen and Hyssalo, 2013). All previous factors serve to contribute to gain sufficient user feed-back and to integrate it into the development and design process (De Moor et al., 2010; Sauer, 2013; Ståhlbröst and Host, 2016). Furthermore, learning processes are preferably open, however, when practical proposals are supported and taken on to attract financial investment, a transparent selection model is needed (Guldemond et al., 2012; Veeckman et al., 2013). This feature is less important if novelty of the solutions is high or priority to involve the business world is low.

As a second subset of factors: Collaborative learning and design cannot work without respecting important social values and these values are preferably considered prior to the start of living labs to prevent the rise of practical obstacles, e.g. connected to legal authorization, privacy and identity (Dutilleul et al., 2010; Kop, 2011). Not sufficiently respecting of
values may cause loss of trust and eventually enhance conflict and delay. Important values are connected to users and managers, while often broader societal values are incorporated in the living lab methodology, like sustainability. A third subset of factors is connected to the broader networking. The composition of actors needs to be balanced while including all relevant ones; however, at the same time, there is a need to avoid involvement of many diverse actors, having one dominating the other(s) and causing the rise of strong interdependency between part of the actors (Guldemond et al., 2012). Though the need for openness and neutrality deserves due attention, in particular, connecting to external actors that cast serious doubt on maintaining the living labs’ core values should be avoided. At the same time, to

Table 3. Key factors in performance of living labs.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evolving learning processes (co-creation)</td>
<td>• Develop a working plan (intended results, budgets) and plan for continuous evaluation at start, incl. participatory nature, major boundary spanning issues, etc.</td>
</tr>
<tr>
<td></td>
<td>• Early involvement of users and timely preparation in dealing with ‘vulnerable’ ones</td>
</tr>
<tr>
<td></td>
<td>• Sufficient motivation of actors to participate</td>
</tr>
<tr>
<td></td>
<td>• Adequate capabilities/skills of actors to perform roles and interact, dependent upon openness and flexibility in models/tools and exploitation/exploration</td>
</tr>
<tr>
<td></td>
<td>• Multiple approaches and collaboration tools, and flexibility in actor roles, dependent upon openness and exploration/exploitation (eventually, multi-sector and multidisciplinary)</td>
</tr>
<tr>
<td></td>
<td>• Ability to deal with unpredictability, dependent on openness and exploration/exploitation</td>
</tr>
<tr>
<td></td>
<td>• Ability to handle conflicts and work with intermediaries</td>
</tr>
<tr>
<td></td>
<td>• Sufficient gaining and absorption of end-user feed back</td>
</tr>
<tr>
<td></td>
<td>• Transparent project selection and decision, eventually, design of business models</td>
</tr>
<tr>
<td>Evolving learning processes: values</td>
<td>• Legal issues, like liability, IP issues, data ownership and access</td>
</tr>
<tr>
<td></td>
<td>• User-values: trust, privacy, cultural identity, wishes of self-determination, cultural ‘distance’ to ICT, transparency in decision-making</td>
</tr>
<tr>
<td></td>
<td>• Values among managers: commitment, passion, risk-taking</td>
</tr>
<tr>
<td></td>
<td>• Societal values, like sustainability and responsibility</td>
</tr>
<tr>
<td>Broader networking</td>
<td>• Avoid large numbers of actors, powerful/dominant actors and strong interdependency; avoid actors that do not comply with living lab values</td>
</tr>
<tr>
<td></td>
<td>• Give attention to upscaling, financial investment, regulators, contractors, etc.</td>
</tr>
<tr>
<td></td>
<td>• Develop embeddedness with focus groups, community of practice, etc. and supportive policies (region, sector)</td>
</tr>
<tr>
<td></td>
<td>Indirectly related to boundary-spanning</td>
</tr>
<tr>
<td>Real-life environment</td>
<td>• Shape an ‘inviting’ arena where improvisations and tacit knowledge are shared and inventions created and validated</td>
</tr>
<tr>
<td></td>
<td>• Settle issues concerning access to places, and implementation of new infrastructure (timing, responding to willingness-to-pay)</td>
</tr>
</tbody>
</table>
enable to bring the co-created solutions to application, or more specific market-place, a sufficient inter-organizational learning needs to be developed, like with regulators, financial investors, policy makers, civil society, market experts, etc. Further, community building is important with citizens and market actors to prevent low willingness to be involved (Evans, 2017; Mazer et al., 2015; Ståhlbröst and Bergvall-Kåreborn, 2011; Veeckman et al., 2013).

A final subset of key input factors, though less clearly connected to boundary-spanning, is the real-life environment (Almirall et al., 2012; Liedtke et al., 2012). This environment is often forwarded as a given fact, but real-life settings need to enable participating actors to go beyond just a realistic scenario, by constituting arenas where different meanings arise, results from improvisation get captured and novel solutions are created and validated (Almirall et al., 2012).

Most key input factors identified for person- and organization-oriented initiatives, also apply to university living labs, but some important additional issues in boundary-spanning can be mentioned. University living labs tend to be more complex because of the variety of activity (domains) and additional tasks in teaching students, the pronounced involvement of university real estate and facility services, and continuous building of on-campus and off-campus communities, all requiring strong management and maintaining common interests (Evans et al., 2015). For example, there is the challenge to remain focusing on sustainability if powerful actors get involved in real-estate development tending to use buildings as showcases and profit-making objects. A further main challenge is maintaining a fine-tuned relationship with the city and civil groups to provide a pipeline of applied research projects and to guarantee continuity in up-scaling and testing opportunities in the complex city itself (Evans et al., 2015; Mazer et al., 2015). A related issue is preserving the relationship with urban reality; too advanced technical solutions need to be avoided as these prevent learning by using (in living, working) due to constraints from regulation and potential technical failure (Evans, 2017; UBC Insiders, 2015). The same holds for using models of financial funding on-campus that cannot be replicated outside the campus.

Overall, a different strength of actor-complexity calls for different emphasis on the above key performance factors, given differences between single versus multiple involvement of user groups, single versus diverse disciplines/sectors and activities, and single versus multiple external networks. However, living labs may be different on other dimensions than direct actor-complexity. Important diversity is concerned with openness and pre-defined character of models and tools, type of drivers of the living lab, and strategic focus on exploitation or exploration, calling for a different emphasis in use of the evaluation framework and key performance factors (Table 3). Accordingly, if innovation models and tools are loosely defined and relatively open (Leminen and Westlund, 2017), special attention is required for matching capabilities/skills among learning partners, including dealing with low predictability and alignment with the working plan. An open and somewhat volatile character also requires special capabilities of the management. Further, drivers of living labs may be other than users (Dell’Erra and Landoni, 2014), e.g. municipalities and hospitals versus citizens and patients as users. If other actors are drivers, there is a need to continuously evaluate the power distribution and respect for user-values, as well as the working of learning models and tools that satisfy all actors, particularly the driving actor. And finally, if the orientation of the living lab is on exploitation and capitalization of current resources, competences and knowledge (Gupta et al., 2006), focal attention in learning tends to be on willingness-to-pay, business models, upscaling, financing, regulation, standardization, etc., which causes evaluation to focus on the concomitant networking and support, and developing them in time. By contrast, exploration captures, discovers, and creates new knowledge by using variation, risks, experimentation, plays and flexibility in learning. Accordingly, evaluation needs to
focus on a rich set of models and tools, but also on capabilities and skills of learning actors to deal with creativity and low predictability.

Conclusion

This article has developed an evaluation framework for living labs and proposes a continuous evaluation integrated in the work plan (process design) of living labs. The proposed evaluation framework includes a system approach; a focus on actor-complexity and boundary-spanning; a set of basic questions to be addressed; a preliminary list of key performance factors; and a focus on participatory evaluation. The system approach, which distinguishes between a set of inputs to living labs’ learning and networking, learning processes and networking themselves, circumstances that are beyond control, and intended and non-intended outcomes, provides a novel application of system thinking in understanding the performance of living labs. There are, however, differences between living labs that matter for evaluation. This article takes strength of actor-complexity as an ‘overarching’ dimension in categorization of living labs, because this complexity and the boundary-spanning involved are basic conditions enabling the intended learning processes and networking. Accordingly, three types of living labs have been analysed, person-oriented, organization-oriented, and multi-activity living labs, the last as multi-functional, large-scale and employing multiple networks. With increasing actor-complexity, also managerial efforts get stronger.

The proposed evaluation framework forwards five questions, a traditional one on alignment of processes (results) with working plans and budgets, and four questions typical for living labs, namely, on integration of user feedback in the design process, alignment of the different actors in goals and interests, and openness to attract and collaborate with partners in implementation of the solutions. An overall question on satisfaction of participating actors with processes and results is important as well. In addition, the proposed framework provides a set of preliminary key performance factors which can be used as a ‘check-list’ in the design of on-going evaluation. The factors that are directly connected to strong actor-complexity and prevention/mitigation include adoption of multiple approaches and tools, learning to handle conflicts and deal with intermediation, respecting basic ethical/legal issues, in particular preserving common goals and interests. In a broader networking, it is important to avoid large numbers of actors and dominant ones, and to avoid ‘newcomers’ that introduce conflicts with established values. Giving sufficient attention to upscaling and adoption in real-life is also paramount but much more complicated in the case of university living labs (Evans et al., 2015). Finally, the framework puts an emphasis on the participatory nature of evaluation, meaning that elaboration of the questions, analysis and conclusions are performed by or subject to interpretation and feedback of the participating actors. The degree of participation may differ to a certain extent, and so does influence of the management and of actors from outside. Bringing the above characteristics together in one integrated framework constitutes the main novel contribution of this article to the literature. However, not all characteristics of evaluation are directly related to actor-complexity, reason why the article also paid attention to other dimensions of living labs, like degree of openness in innovation models and tools, type of actors as drivers of the living lab and its results, and orientation of the activities involved, exploration and exploitation.

Causality remains an important issue in evaluation as the system approach does not prove causal relations and does not provide an assessment of their strength. Action-based research can serve as one part of the solution because it captures problematic processes, different problem perceptions as well as perceptions of problem solutions. This situation connects to challenges in future research. One of the limitations of the current study is the small number of
qualitative case studies, reason why the results on key performance factors call for thorough quantitative testing, e.g. using regression modelling or structured equation modelling, drawing on a large and representative sample enabling statistical generalization. In particular, the causal relationships between inputs and outcomes need to be quantitatively assessed. Fuzzy set analysis can be helpful if measurement is mainly at categorical level and emphasis is put on fuzziness of data on behaviour and perceptions, etc. (Smithson and Verkuilen, 2006). If results of quantitative analysis are combined with understandings from action-based research, evaluation and monitoring can be carried out on a more thorough basis. In addition, the case study analysis was restricted to Northwest Europe and Canada, implying influence of specific cultural traits and values (Hofstede and Hofstede, 2005; Van Geenhuizen, 2016). For example, the bottom-up and participative character of living labs and their evaluation presented in this article may not match with planning and management cultures elsewhere in the world. This calls for more research on adaptation of living lab design and evaluation framework to other cultural settings. And finally, the results of this article point to the need for more understanding of actor-complexity and needs for boundary-spanning, particularly where these are strong. For example, how can boundary-spanning prevent impacts from cultural and cognitive differences and diverse power positions between the living lab actors, like that of elderly nurses and (large) technology companies developing their prototype, and how should that be better managed (Hakkarainen and Hyysalo, 2013)? Actor-complexity tends to increase more recently with the growing application of living labs in parts of universities, cities, airports, etc. and this justifies the question on how boundary-spanning can be (better) facilitated in living lab methodology and also the follow-up question on how participatory evaluation could be designed with high levels of actor-complexity? Future research might also take-up these two important questions. And finally, the evaluation framework itself needs a further testing, along with investigation of managerial implications.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note
1. The study benefitted from various workshops and conference presentations, e.g. organized by intermediary institution Medtech West (11 June 2015), named ‘Living Labs in Medical Delta’ (NL). Furthermore, there has been extensive communication and exchange of ideas with Dr. Nick Guldemond at Delft University of Technology in 2013–2015, and interviews with Dr. Iain Evans in Vancouver, Canada, 14 July 2015 and Prof. Dr. James Evans in Manchester, UK, 14 April 2017.

References
ARC (Amsterdam Region Care&ICT) (2013) Innovating in Care in Collaboration with the End-User (in Dutch). Amsterdam: Author.


Evans I (interview, July 14th (2015) UBC, Vancouver).

Evans J (interview, April 14th (2017) University of Manchester).


Save P (2014) UBC Living Labs: Innovation in Accelerating the Adoption of Sustainable Technologies for Campus Infrastructure. Vancouver, UBC.
UBC (University of British Columbia) (2016) Campus as a living lab. Available at: https://research.ubc.ca/impact-innovation/campus-living-lab (accessed 8 January 2016).


Marina van Geenhuizen is Full Professor of Innovation and Innovation Policy in Urban Economies, in the Faculty of Technology, Policy and Management at TU Delft, since 2007. Her research focuses on theory and empirics of commercialization of new knowledge, including Triple Helix development. Her fields of application are sustainable energy solutions, medical and healthcare innovation, sustainable transport innovation and more recently, the overarching theme of responsible research and innovation. She is author of around 90 reviewed journal papers in English, and of a 100 chapters in English books. She is also the principle editor of nine Volumes in English. In addition, she acts as reviewer for the Netherlands Organization for Scientific Research and has been member of various local and national advisory boards on innovation and urban development.