

# ECCOdata: A proposal of an Empirical Co-created Canvas for Opening up Data of Public Interest

Dimitris Gouscos<sup>1</sup>, Alexandros Melidis<sup>1</sup> and Efthimios Tambouris<sup>3</sup>

<sup>1</sup>National and Kapodistrian University of Athens, Panepistimiou 3, 106 79 Athens, Greece

<sup>3</sup>University of Macedonia, Egnatia 156, 546 36 Thessaloniki, Greece

## Abstract

This paper presents the rationale, making effort and research agenda for co-creating an empirical fit-for-purpose canvas model that can guide non-expert quadruple-helix stakeholders within a Civil Society Organization (CSO) working group for Open Government and Open Data to better structure their proposals for opening up data of public interest. The authors provide the real-world case framing this work, findings from a review of relevant approaches, an imperfect seed version of the canvas model built, design choices behind its making, a research agenda for co-creation and co-assessment with end-users, a value proposition, and limitations of this effort.

## Keywords

open data, open government, canvas models, co-creation, stakeholder engagement

## 1. Introduction

Nowadays, opening up of public-interest data is becoming an established good governance practice for institutions worldwide, and an indispensable component of global policy agendas such as the UN SDGs. The multitude of open data efforts that are currently underway on these grounds gives rise to the need for guidelines that can help these efforts properly consider all present aspects and future impacts of opening up public-interest data, as well as document their plans and projects with specificity and conciseness. This need also applies to civil society initiatives for opening data. One such initiative is the Greek Open Technologies Alliance Working Group for Open Government and Open Data, which is asking its subgroups and members to co-create documented proposals for opening up data of public interest. The research community has acknowledged this need and come up with proposals for open data canvas models, which are well positioned to serve as the guidelines sought. Still, the diversity of field contexts does not always allow existing models to cover all needs, especially felt needs to make a model more fit-for-purpose by actively participating in its own making.

The objective of this paper is to report work resulting from such a need that has been encountered in practice and from the understanding that the best way to meet this need would be to embark on a twofold making and research effort for building a seed solution to be further developed through co-creation. The more specific needs that led in this direction come

---


*EGOV-CeDEM-ePart 2022, September 06–08, 2022, Linköping University, Sweden (Hybrid)*

✉ gouscos@media.uoa.gr (D. Gouscos); almel01@gmail.com (A. Melidis); tambouris@uom.edu.gr (E. Tambouris)

🆔 0000-0001-9252-8686 (D. Gouscos); 0000-0001-9255-8770 (A. Melidis); 0000-0001-8036-9788 (E. Tambouris)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

from the following inputs : 1) An understanding of the encountered need for an opening up public-interest data guideline. In the context of coordinating a CSO-hosted working group for open government and open data, comprising non-expert members, and from reviewing these members' proposals for opening up data of public interest, we have understood that group members best need some guideline to follow for structuring their proposals with more coherence and completeness. 2) The canvas modeling paradigm. Reflecting on possible formats for such a guideline, we have turned for a number of reasons to the canvas modeling paradigm, showcased by the business canvas model. 3) Offerings and limitations of available open data-related canvas models. Reflecting on whether available open data-related canvas models are fit for our purpose, we engaged in a literature review of such models for aspects covered and potential fit to our own needs and context. This review has identified some gaps that led to the decision to embark on an effort to develop a new canvas model. 4) Pragmatic open science and co-creation practices. Reflecting on how to best foster end-user engagement in and acceptance of the new model sought, we turned to pragmatic open science and co-creation practices and reviewed their potential to improve end-user engagement in the making process and end-user acceptance of the made artifact. This review has led to their adoption, and to the final decision to embark on a making and research plan for building an intentionally imperfect seed version of the sought model, to be further co-created with end-users. 5) Desirable aspects of open public interest data. Reflecting on how to best make this seed model version inclusive of aspects desirable for open data that serve the public interest, a need which stems from the context of our work overall, we reviewed selected relevant literature and identified some aspects and principles that the seed model sought needs to integrate.

In this respect, the rest of paper is structured as follows: section 2 frames the effort reported by presenting the case and needs encountered in practice; section 3 presents the review of current approaches performed for choosing for choosing the canvas model family, and then visiting currently available open data-related canvas models for identifying common aspects and any gaps, reviewing pragmatic open science and co-creation practices for their end-user engagement and acceptance potential, and reviewing aspects and principles that open data of public interest should meet; then sections 4 and 5 present the model built in a first imperfect seed version to spark co-creation/co-assessment, and the methods followed and design choices made in this effort; and section 6 introduces the future work agenda that opens up in the quest of co-creation and co-assessment with stakeholders, value claims and limitations of the work reported.

## **2. Framing of the research case**

Open Technologies Alliance (GFOSS, <https://gfoss.eu/>) is a Greece-based CSO promoting openness in all technology domains. In 2021Q2 GFOSS launched a Working Group for Open Government and Open Data (O2gd WG) to be co-coordinated by the first and third authors for the Sep 2021 – Aug 2023 period. The second author, currently GFOSS General Director, regularly contributes in policy decisions for this working group, whereas one more GFOSS Board Member undertakes a supervisory role. O2gd working group bylaws and planning (published at <https://opengov.ellak.gr/2021/09/20/opengov/>) aim at fostering inclusive [1] and multilateral

[2] public influence, alongside social innovation [3] by populating the working group with quadruple-helix [4] stakeholders, coming from the academic, public, private, civil society and journalism sectors. Through an ongoing open call, the O2gd WG has currently integrated some 50 stakeholders, staffing 4 transversal and 8 topical working subgroups. Topical subgroups refer to the domains of climate change; water resources; energy transition; public health; public-interest budgets; local administration; reception of refugees; and public-interest services and public administration, and their agreed mission is to propose to open up some public-interest data and/or government processes that are not currently available in their reference domains in open format or at all, and subject their proposals to public influence at IAP2/Consult [5] or higher level. This mandate, therefore, also covers proposals for establishing new open datasets/dataspaces and new open government processes.

The works of each O2gd subgroup constitute a thick participation effort in the sense articulated by Nabatchi and Leighninger [6], i.e. an effort asking much from a few people. At the same time, the O2gd WG as a whole has a relation to O2gd subgroups which constitutes a public influence relation in its own right, and so much so at the IAP2/Empower [5] level: O2gd WG's ultimate goal is to ideally place final decision making in the hands of O2gd subgroups, which embody the public, and deliver them the promise that it will adopt and promote (as it does not have itself an implementation mission) what they decide. For this ideal situation to be achieved, a number of conditions are required which encompass, among other, attitude shifts and methodological guidance. With regards to attitude shifts, the O2gd quadruple-helix WG has managed to avoid resistances to effective multistakeholder collaboration due to low cross-sectoral trust, as well as barriers to creativity due to discontent and pessimism. However, once creativity within subgroups is sparked, the need has been observed to work together subgroup members for moving from more closed making attitudes (working on things alone, releasing them only upon perfection, keeping them ours) to more open making or co-making attitudes (working on things together, releasing them with accepted imperfections, making them common property). Gaining understanding of this needed attitude shift has helped the authors to adopt themselves a more open making mindset in the development of the open data canvas model proposed in this paper. With regards to methodological guidance, on the other side, the thinking and proposals for opening up data of public interest have been observed to show imbalances like a bias of inspiration towards data supply (what data we know to be available already) rather than data demand (what data we would like to see available), forward-running emphasis to 'how' technical aspects (data formats, digital delivery) before fully working out 'why' aspects (value, public interest), or absence of focus on the dynamics of data opening, such as actors affected, positive/negative impacts, new possible synergies, or concepts of success.

These observations have led to infer that a sort of guideline is best needed in order to help subgroup members arrive at more coherent and complete proposals for opening up data of public interest, built through a more open making mindset. Still, in the O2gd WG context, which is critically impacted by stakeholder motivation and collaboration dynamics, the form and contents of this guideline, as well as the process itself to follow for arriving at it, need to be chosen with the overarching goal of maximizing members' engagement and acceptance.

### 3. Literature review findings

Trying to see how current approaches can provide findings and/or gaps that can help to arrive at a guideline for opening up public-interest data needed in our case, we have explored a number of such approaches from the literature, on a) canvas models in general, b) open-data related canvas models, c) pragmatic open science and co-creation practices, and d) aspects to consider for public-interest data. For this exercise we have used a non-systematic review lens, trying to identify for each topic some sources to further study that best combined up-to-datedness and citation impact. We have also tried to take stock of published literature reviews, where available. Out of the different modeling choices that can be imagined to specify how a proposal for opening up public-interest data might look like, we have intuitively set out to explore canvas models. From approaching Osterwalder & Pigneur's work on Business Model Canvas [7], also part of NESTA's toolkits [8], and Maurya's work on Lean Canvas [9], we find canvas models to fit our purpose well since: 1) they are brief in template and textual size and focused on thinking only about essentials; 2) they abstain from technical terminology, which renders them accessible to lay audience, although perceived linguistic ambiguities and misconceptions may always need to be resolved; 3) they accommodate the 'what' as well as 'how' aspects of a vision, plan or proposal; 4) they strike a good balance between structure and freedom, by non-hierarchical bird's eye view presentation, leaving interdependencies emerge organically from reflection, and by open-ended questions, leaving room to creativity; and 5) importantly, they are modular, allowing to separate concerns and manage work accordingly.

From reviewing canvas models in general we are able to obtain some interesting common characteristics. More specifically, a canvas model needs to adopt simple language, be rich of concepts, and enable visual representation. On the other hand, canvas models are often designed for the business sector, can be vague, and are not developed using co-creation processes. Actually, vagueness is an interesting characteristic when considering co-creation. Indeed, a controlled degree of vagueness on the seed artifact from which the co-creation process could provide a sufficiently low threshold and a sufficiently challenging motivation for end-users to intervene with questions and answers for clarifying things, and in this way for co-creation to begin.

From reviewing open data canvas models [10, 11, 12, 13] we conclude that they are mainly inspired by the business canvas model in terms of logic and structure as well as understanding value creation by open data.

The idea for creating canvases especially for open data value creation can serve different purposes and stem from different angles. For instance, Gao & Janssen [10] combined an analysis of the business model canvas with a literature review on open data specifics in order to introduce an adapted open data canvas aiming to take into account elements conducive to value creation out of data such as data providers, infomediary, value proposition, activities, resources, channels, cost, partners, users. This canvas pays attention to the role of different stakeholders and the activities with which they earn benefits and ultimately value from open data.

In another vein, the Open Data Institute is heavily focusing on ethical matters in relation to data use. Thus, the ODI Data Ethics Canvas, ODI [14], consists of a framework aiming to assist any kind of data project to apply practical guidelines to tackle ethical concerns with regards to spotting possible negative impact as a result of an open data project at play. Furthermore,

based on the Business Model Canvas, the ODI has produced open standards for data canvas [15] in an effort to promote a shared understanding of what people and organizations need when developing an open standard for data publishing, access and use.

From reviewing some pragmatic open science [16] and co-creation practices [17, 18, 19, 20] in a number of application domains, two important findings are drawn. Firstly, these approaches include the option to start collaboration with stakeholders in a more or less tabula rasa style, either from a non-artifact, i.e. from a broad partial description of what is sought, or from an imperfect artifact, i.e. from a preliminary and sufficiently open-ended version of what is sought. Secondly, these approaches argue for achieving stakeholder process engagement and results acceptance at levels at least equivalent to, if not superior than, those achieved by more closed making approaches.

Lastly, from reviewing work on public interest data, the following desirable aspects to be met when making these data open are drawn: 1) accessibility, broadness, curation and documentation, as defined by Mc Grath-Lone et al [21]; 2) avoidance of privacy risks, as discussed by Blankertz [22] and included in the accessibility aspects of Mc Grath-Lone et al [21]; and 3) findability, accessibility, interoperability and reusability attributes, as defined in the FAIR principles [23], and also overlapping with the aspects recognised by Mc Grath-Lone et al [21].

#### **4. The seed model built**

The seed version which has been built for the canvas model for opening up public-interest data is structured along 6 sections comprising 24 open-ended questions in total. A bare-bones textual representation of this model is presented in Table 1.

This model is influenced by the business canvas model as well as open data canvas models presented above. For instance, the main elements of open data canvases in Gao & Janssen [10] literature review such as stakeholders and benefits & values are also present in the proposed canvas in the ecosystem tracing and value expectancy perspectives. Also, some specific contents under the main elements of barriers or resources and key activities such as, for instance, ICT infrastructure, data discovery and collection, data integration and cleaning, data processing and analysis, application design development and deployment, visualization or application could be traced as responses to ECCOdata canvas model questions about data to open and implementation.

#### **5. Methods of work**

To build this seed version of the model we have tried to take stock of the review findings in four distinct ways: 1) keep this model compliant to the canvas logic; 2) focus on fit to the process of opening up public-interest data, rather than to other facets already covered by models from the literature; 3) take stock of the main findings from pragmatic open science and co-creation practice review; and 4) accommodate desirable public-interest data attributes. On these grounds, we have made a number of design choices for imperfection, presented in Table 2.

In the seed model built we have tried to keep these imperfection choices controlled, to ensure a basic level of quality. Understandably they leave the model open to criticism, but this, from

**Table 1**

Textual representation of the seed version of the ECCOdata canvas model

Sections	Questions
data to open	<p>which data are to open?            who are the people who create these data?            where do these data now exist?            what is it that these data currently lack in terms of openness?</p>
value expectancy	<p>if these data open, will they help growth by facilitating entrepreneurship? how?            will they help society by making everyday life easier? how?            will they help informing the public by facilitating journalists work? how?            will they help education by facilitating scientific research? how?            will they help public policy making in some area? how?            will they help transparency, accountability, public participation? how?</p>
ecosystem tracing	<p>what do these data refer to? who makes this? who needs this?            if these data open, are there people who may lose? who are they, and how may they lose?            do the people concerned by these data work together? do they feel rivals? how so?            could gain-side and lose-side people work together so that nobody loses? how?            could gain-side and lose-side people work together to create new gains? how?</p>
influence reception	<p>from people concerned by the data to open, who has had an influence on this proposal?            at what stage was their influence expressed? how was it incorporated?</p>
implementation	<p>when these data open, how will they look like?            on what platform(s) will they be available?            what functions will be offered for handling them?</p>
concept of success	<p>what would this data opening effort ultimately need to achieve to be considered a success?            who could ascertain that the data opening effort has ultimately achieved its goal?            can we use any existing criteria to ascertain success? which ones?            do we need to devise any new criteria to ascertain success? which ones?</p>

an open making perspective, is an advantage: they keep the model lean, and offer the grounds for this criticism to be made by the stakeholders themselves, thus engaging the latter in the co-creation and co-assessment sought.

Alongside these design choices, we have set a number of syntactic rules for formulating the model's textual elements, drawn from the choices observed in canvas models from the literature, as follows: 1) avoid adjectives, as they tend to denote a possibly unclear division of the meaning space conveyed by the noun they refer to; 2) prefer verbs, nouns and verbal nouns, as they tend to convey a more integral meaning space; 3) avoid technical terms that may not be known to all; 4) use everyday language and the simplest words possible; 5) use plain syntax, with simple

**Table 2**

Design choices for imperfection of the seed version of the ECCOdata canvas model

Design choice	Remarks
bias towards the research case	aspects frequently present in (e.g. implementation) or missing from (e.g. concept of success) discussions with O2gd topical subgroups are included for instance, financial or organizational aspects
incompleteness	of planning data opening projects and delivery services are not included in the seed version
abstraction	more specific aspects (e.g. privacy concerns, FAIR requirements) are substituted by more generic concepts (potential losses, data locus and form)
vagueness	terms used like “help”, or “concerned” may mean different things the seed version hints at, but does not provide a visual representation,
non-visualization	in order to avoid premature design decisions and make this important task grounds for co-creation

subject-verb-object constructs; 6) formulate questions in 5W1H terms, to give to worked-out canvases story completeness; and 7) apply a “one question one aspect” principle.

For the overall model-building effort, we have followed the steps mentioned in section 2 above. Within the seed model-building step, we have employed IAP2’s participation spectrum as an internal collaboration paradigm: we have mutually exchanged roles to inform each other about the need and alternative solutions, consulted each other for feedback from the literature and provided feedback to each other’s concerns and aspirations, involved each other in the design choices made, collaborated in order to apply them, and finally empowered each other to propose specific sections or questions of the seed model.

## 6. Conclusions and future work

Future work for this model building effort comprises a number of strands, which call for open making and for research work in parallel:

1. formulate a model co-growth process for successive rounds of model use for canvas-building, co-assessment of the current model version and co-creation of a next improved version, together with internal and external stakeholders of our research case, from the O2gd group members and beyond;
2. explore how the imperfect design choices for the model’s seed version can be resolved for less bias, more completeness, less abstraction, less vagueness, visualization in next model versions, through this co-growth process;
3. explore how two important canvas-logic dimensions, namely domain agnosticity (avoidance of binding to any specific application domain of public-interest data) and closed openness (avoidance of “anything else to note”-style questions) can best be maintained or raised in future versions (leading to domain-specific models, or inclusion of free questions);

4. assess how the model can best evolve to meet goals such as • use process versatility (ability to build canvases through any more waterfall or more agile process); • tolerance to stakeholder contribution irregularities (lazy, crazy, lossy, bossy and messy contributions to canvas-building); • simplification, demystification and trivialization of the data opening effort; • progress quantification and manageability of this effort; • support for visionary thinking and two-sided dynamics (gains, losses, capabilities, impacts from data opening);
5. evaluate, in the seed version and future improvements, the model's stakeholder-side quality, encompassing conceptual quality, effectiveness, short learning curve, time efficiency, acceptance, and contribution to a positive participation experience.

To conclude, we believe that the work reported in this paper delivers a three-fold value proposition. Tool-wise, this work delivers an open data canvas model which is complementary and symbiotic to existing ones, explicitly focusing on the thinking required for opening up data of public interest. Research-wise, this work proposes the idea of bringing open science practices into the building of methods for open data, and lays the grounds for such a making and research effort. Finally, public practice-wise, this work delivers results that can help experts to interact with the non-expert public, and the non-expert public to actively participate, in the advancement of issues of public interest.

At the same time, the major limitations of the early work reported in this paper stem from its initial empirical departure point, and from the fact that it is only on the outset as a research effort, with results not having yet been validated.

## References

- [1] IAP2, Core values, ethics, spectrum – the 3 pillars of public participation, n.d. URL: <https://www.iap2.org/page/pillars>.
- [2] OGP, Multi-stakeholder forums – collaboration is the cornerstone, n.d. URL: <https://www.opengovpartnership.org/multistakeholder-forums/>.
- [3] SINA, Social innovation academy playbook, 2021. URL: <https://socialinnovationacademy.org/playbook/>.
- [4] E. Carayannis, D. Campbell, 'mode 3' and 'quadruple helix': toward a 21st century fractal innovation ecosystem, *International Journal of Technology Management* 46 (2009) 201–234.
- [5] IAP2, Iap2 spectrum of public participation, 2018. URL: <https://www.iap2.org/page/pillars>.
- [6] *Public Participation for 21st Century Democracy*, John Wiley & Sons, Inc, New Jersey, 2015.
- [7] A. Osterwalder, Y. Pigneur, *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers* (The Strategyzer Series, John Wiley & Sons, Inc, New Jersey, 2010).
- [8] NESTA, Business model canvas. *nesta toolkits collection*, 2018. URL: <https://www.nesta.org.uk/toolkit/business-model-canvas/>.
- [9] A. Maurya, *Running Lean: Iterate from Plan A to a Plan That Works*, 2nd ed., O'Reilly, Boston, 2012.



- [10] Y. Gao, M. Janssen, The open data canvas – analyzing value creation from open data, *Digital Government: Research and Practice* 3 (2021) 1–15.
- [11] T. Kronsbein, R. Müller, Data thinking: A canvas for data-driven ideation workshops, in: *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS)*, 2019, p. 561–570.
- [12] T. Neifer, D. Lawo, M. Esau, Data science canvas: Evaluation of a tool to manage data science projects, in: *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS)*, 2021, p. 5399–5408.
- [13] M. Kaufmann, Big data management canvas: A reference model for value creation from data, *Big Data and Cognitive Computing* 3 (2019) 19.
- [14] ODI, Data ethics canvas, 2021. URL: <https://theodi.org/article/the-data-ethics-canvas-2021/>.
- [15] ODI, Open standards for data canvas, n.d. URL: <https://standards.theodi.org/useful-tools/how-to-use-the-open-standards-for-data-canvas/>.
- [16] B. Fecher, S. Friesike, Open science: one term, five schools of thought, in: S. S.Friesike (Ed.), *Opening Science The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*, Springer, Cham, 2014, p. 17–47.
- [17] T. Eseonu, Co-creation as social innovation: including ‘hard-to-reach’ groups in public service delivery, *Public Money & Management* 42 (2022) 306–313. doi:10.1080/09540962.2021.1981057.
- [18] S. M. C. Loureiro, J. Romero, R. G. Bilro, Stakeholder engagement in co-creation processes for innovation: A systematic literature review and case study, *Journal of Business Research* 119 (2020) 388–409. doi:10.1016/j.jbusres.2019.09.038.
- [19] M. Ramírez, Co-creation and open innovation: Systematic literature review .comunicar, *Media Education Research Journal* 54, XXVI (2018).
- [20] W. Voorberg, V. Bekkers, L. Tummers, A systematic review of co-creation and co-production: Embarking on the social innovation journey, *Public Management Review* 17 (2015) 1333–1357. doi:10.1080/14719037.2014.930505.
- [21] L. Mc Grath-Lone, M. Jay, R. Blackburn, E. Gordon, A. Zylbersztejn, L. Wiljaars, R. Gilbert, What makes administrative data research-ready? a systematic review and thematic analysis of published literature, *International Journal of Population Data Science* 7 (2022). doi:10.23889/ijpds.v7i1.1718.
- [22] A. Blankertz, How to use data for the public interest, even – or especially, in: in a pandemic. *Heinrich-Böll-Stiftung Hong-Kong Asia | Global Dialogue*, 2020. URL: <https://hk.boell.org/en/2020/11/24/how-use-data-public-interest-even-or-especially-pandemic>.
- [23] GO FAIR, Fair principles. go fair initiative, n.d. URL: <https://www.go-fair.org/fair-principles/>.