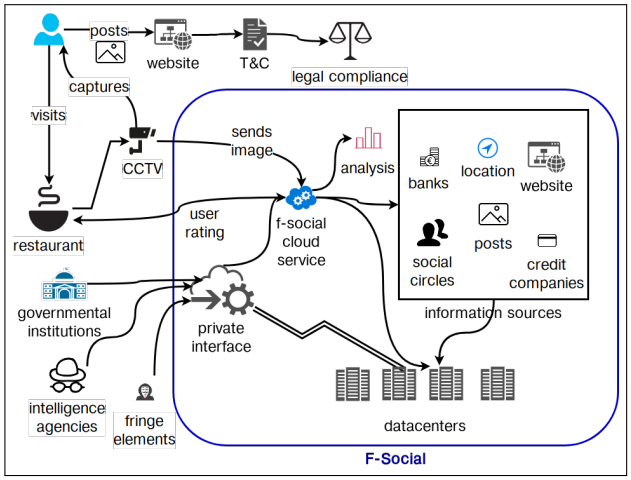

Figure 2: F-Social: Utilising the web to provide user ratings as a service

4 RESPONSIBLE INNOVATION

The world wide web, also known commonly as the web or the internet, provides a foundation for large-scale sharing of information across the globe and is the basis of a vast number of commercial entities that depend on it for their business. The basis of the web is protocols or standards that are used as a mutually understandable form of interoperability for exchanging information. This can be done via websites or abstract interfaces (also known as APIs) that let people interact with data in a semantically richer way. While advances in technology increasingly depend on this connected nature of the web, we see cases where the very foundations of the web are used to subvert privacy considerations in various ways. One way this is done is by controlling and modifying packets in the connection by the service provider[7] that allows injecting ads as well as selectively throttling traffic. While such practices may or may not have legal repercussions, they are certainly a cause of alarm regarding privacy. The control over internet packets effectively allows an identification tracker to be inserted in to the web-packet itself, which may allow any website or service to identify the user (at some abstract level) thereby further advancing the use of non-explicit information in user profiling as outlined in the previous section. While such activities can be subverted using laws and public pressures, both of these can be very difficult or time consuming to take effect. Meanwhile, this cannot be used as a means to establish control over the web or to ban it or to restrict its usage in the sense of usage.

A practically viable solution is to increase the level of awareness of the general user to better prepare them for the choice they make when handing over consent or agreeing to use a service. Terms and conditions are a legally mandated way of doing this, but they have been proven to not be effective at all due to the density of the text. GDPR mandates explicit consent that must be obtained by making the user aware of all the ways their data will be used. This approach is certainly progressive, but it is not going to stop user-profiling across the web. Trying to paint the entirety of activities related to user-profiling in a negative light would be to restrict progress in technological advancements. Therefore, we must try and come to a middle ground where progress can take place alongside addressing any practical issues the society may have at large regarding the use of such technologies. This is where the field of responsible innovation can be adopted as a good practice alongside existing approaches such as secure websites and privacy policies.

The basis of responsible innovation is that of building conceptions of responsibility on the understanding that science and technology are not only technically but also socially and politically constituted [5]. In commercial interests, the considerations of privacy and ethics have begun to take shape but are still in their nascent stages and have no effective methodology that can be integrated with business practices. The current norm seems to be to consider the practical implications of an technology after its widespread usage and in most cases only when someone else raises objections based on their perceived risks. There also exist challenges in addressing the uncertainty of technology being used and its rapid change and proliferation.

Instead of looking for a paradigm change in the way privacy and ethics are handled on the web and by organisations, we look towards the core issue underlying the lack of discussions on these topics by technologists - which is that no method for practising ethics integrates into the methodology for work readily enough to adopt it. Another challenge is the perceived authoritarian requirements for discussing ethics, which are not only not true, but also hinder discussions about these topics in open spaces such as those on the web. While privacy policies aim to discern concerns related to privacy, no such tools or practices are in usage for ethics which remains a topic of discussions that are either closed-door or absent from public view.

One way to address this to is to better enable users to tackle the responsibility of understanding technology and its effects on privacy and ethics. While it may not be practical to engage all users in an discussion either one-on-one or as a community, it is certainly practical and possible to present a discussion to the users about the ethics of technology. This will empower the users to ask questions such as "*what am I providing and what am I getting in return?*", and more importantly - "*what are the risks? are they worth it?*". These are inherently personal questions whose answers change from person to person. This is analogous to privacy policies that aim to describe the considerations about privacy but without a proper structure and a methodology, lack the necessary depth as well as guidance on how to approach the issue.

5 ETHICS CANVAS

We developed Ethics Canvas as a novel method to address these challenges using a tool that encourages discussions pertaining to practising ethics in research and innovation. We evaluated existing approaches of responsible innovation [4] which are focused on the design of business but not on technologies involved in the innovation process. To integrate a discussion of ethics into existing methods of discussions, we used the Business Model Canvas (BMC) [3] which allows collaborative discussions about the business and

[7] Also known as Internet Service Provider or ISP

encourages a common understanding of how the business can create, deliver, and capture value.

The Ethics Canvas helps structure discussions on how technology affects stakeholders and the potential it has for ethical considerations. Currently at version 1.9, the Ethics Canvas has evolved over time to better capture and reflect discussions. It consists of nine thematic blocks (see Fig. 3) that are grouped together in four stages of completion. The first stage (blocks 1, 2) requires identifying the stakeholders involved based on the technology under consideration. These are then used to identify potential ethical impacts for the identified stakeholders in stage two (blocks 3-6) and non-stakeholder specific ethical impacts in stage three (blocks 7, 8). Stage four (block 9) consists of discussions structured around overcoming the ethical impacts identified in the previous stages. The ethics canvas can be printed or used as an web application that can be used without an account and can be downloaded. Certain features such as collaborative editing, comments, tagging, and persistence are made available through an account. The source of the application is hosted online and is available under the CC-by-SA 3.0 license. We are working on the next iteration of the canvas and intend to exploit web technologies to provide a cohesive experience around discussing ethics. We welcome ideas, suggestions, and collaboration regarding the same.

We take the example of Nosedive, and the scenario presented in this paper of achieving such a situation through aggressive user-profiling, and use the Ethics Canvas to discuss its ethical implications. The canvas itself used for this example is available online and is provided hereby under the CC-by-SA 4.0 license.

The first stage involves identifying the types or categories of individuals affected by F-social and its services for providing metrics or ratings. Alongside users of F-social, this also includes any user (or individual affected by) of the organisations using the service for obtaining ratings. In the hypothetical scenario, this would include customers of the restaurant. This would also include any individual that is not on F-social but who is present in a picture or mentioned in a post. Extending this to all information sources used by F-social, if it uses any dataset of individuals such as from credit companies, then any individual in that dataset should also be included in this stage. For groups affected, these would be people averse to being tracked such as journalists who might wish for 'safe' places for meetings. This would also include people in positions of power such as politicians or bureaucrats where information about who may inadvertently be present at the restaurant could lead to misuse. Any category of minorities who might inadvertently be aggressively profiled are also at risk.

In the second stage, we explore how these stakeholders might be affected by discussing the potential ethical impacts of the technology. With respect to behaviour, users might find more incentive to post things that positively affect their ratings and refrain from posting negatively affecting things. They are also more likely to provide information if it helps them achieve monetary or other forms of benefits from services that use the ratings to vet customers. This behaviour might encourage an acceptance of invasion of privacy as the legally the users would be willing to provide the information for intended use by F-social. In terms of relations, if ratings take into account the social circle, then users are more likely to want to have their social circle made up of people who would have an positive effect on their rating. This is seen in Nosedive as well where people try to be nicer to others who have a higher rating than them in the hopes of increasing their rating whereas the inverse of this invites people who have higher ratings to treat those with lower ratings with contempt. This will lead to a change in the general perception of individuals and places based on the ratings and metrics they cater or reflect. For example, places that only cater to people with higher ratings automatically are seen as 'exclusive' whereas places that readily accept people with lower ratings might be seen as not being 'classy'. To a certain extent, this phenomenon is observable today with regards to monetary spending capacity. This leads to a natural resentment of each group towards the other which might be open for exploitation by fringe elements of the society for their benefit. A layer of people in a position of power might exploit this opportunity to their advantage such as between organisations and employees where accessing ratings might be considered to be 'in-contract', thus making it unavoidable to prevent the information to be used to influence things such as salary and promotion. As this effect would be very subtle, it cannot be guarded against, but can be mitigated through an open approach and awareness in general.

We consider the potential impacts of these in stage 3, starting with the impact on the services offered by F-social. Since the metric or rating would be seen as an important factor, its algorithm could be open to being gamed. When this information is exposed to the general public, it could lead to a huge backlash and negative repercussions. This effect might still take place even if the information might be partially true or completely false. Additionally, with the primary medium of such services being the web, any issue affecting the security and integrity of communication at this level also affects the service itself. Thus, attack vectors such as man-in-the-middle and DDoS would lead to disabling the service, which might lead to users being denied services. Based on the region and specitivity, this attack can be used to reject a particular subset of regions for political purposes. Methods such as ad-blockers and ISP injections (the practice of an ISP injecting or modifying packets) could be used to subvert the functioning of the service which might lead to unintended consequences for the users.

Some clients of the service may try to use the service for purposes that might not be acceptable from legal, business, ethical, or moral viewpoints. In such cases, it might be in F-social's interest to try and vet their clients and access to the service which can further complicate consequences as they would effectively be denying information and business based on some agenda. This might or might not be acceptable based on how it is implemented, but usually, this leads to complicated terms and conditions that get more complex with time. The government or a state-level entity might want automatic-access to the data which might be an issue on several fronts. Based on the region and the political standing where this occurs, it may or may not lead to problems involving privacy. For example, in a more democratically open region, the potential backlash or political checks and balances present might mitigate the issue or subvert it using legislation to either deny access or grant it based on public opinion. In cases where the government is more isolated and practices authoritative governance, it is less likely to back down from its stance and might threaten to ban the service entirely unless it accepts the set terms.

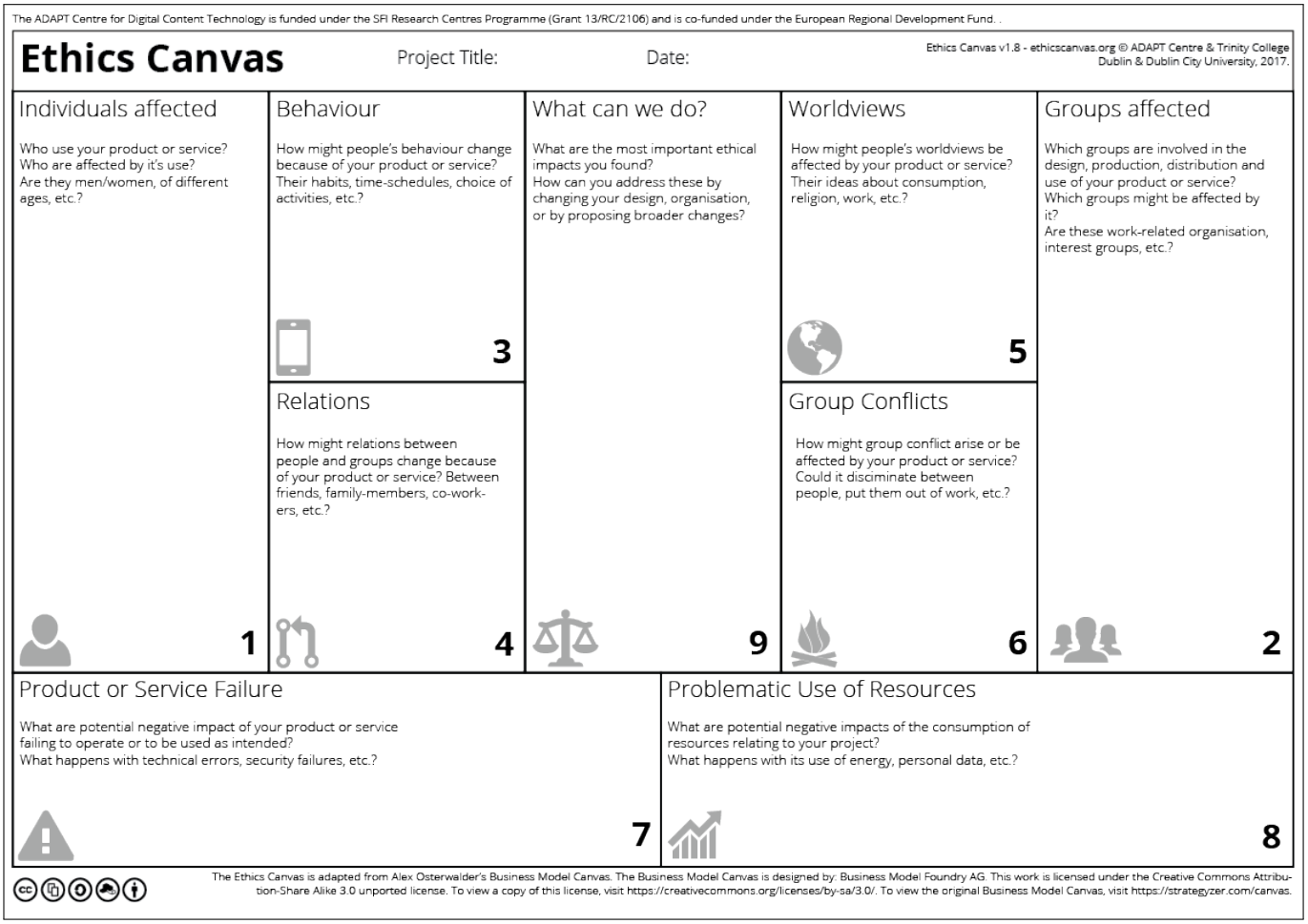

Figure 3: The Ethics Canvas showing 9 blocks that are divided into 4 stages for encouraging structured discussion on ethical issues related to research innovation.

Based on an understanding of the stakeholders involved, the functioning of the service, and how its potential effects on the behaviour, relations, and the service itself, we try to mitigate the impact of these through stage 4. Since F-social's main issues of concern lie in its using the data to calculate the metric or rating, it is difficult if not impossible to ask it to stop providing the service as it this would constitute asking it to stop a legally acceptable practice. Instead, users or organisations might ask F-social to be more open about its usage of personal data, and take into consideration the ethical aspects of its usage. The algorithm used to calculate the rating could be requested to be openly evaluated to ensure that no bias is present and that it's practices are legal. This can be done at several levels, from an internal committee tasked by F-social to a governmental enquiry. This would provide a level of authenticity and oversight to the usage of the data as well as prevent false information from spreading regarding the service.

On the aspect of preventing such usage of personal data altogether, the best possible way would be to disseminate a better and simpler understanding of the issues involved in the hopes to raise a public outcry about the practice and to get legislators involved to draft better laws protecting the concerns of its users. However, this approach has its weakness in the form of being bound by regions where political powers may not work in a cooperative manner. Thus, F-social may end up receiving patronage from one government while being completely banned by another. The medium of the web can be used extensively to share information towards the service and its ethical concerns, similar to existing organisations that concern privacy.

6 CONCLUSIONS

Users may not necessarily have total or absolute control over their data being used or collected. While progressive laws, particularly GDPR, specify several constraints and obligations over the use of persona data, the ultimate control lies with the data subject or the user. Even though consent is a mandatory affair, the responsibility of providing it correctly requires the user to first understand all the implications of the technology which is a difficult task. Instead, this responsibility can be shared by all parties involved, including

the general community. Through this paper, we tried to discuss the implications of user-profiling and how it can be readily provided over the web as a service. We took the example of the episode Nosedive from Black Mirror to consider the ethical implications of such a service and developed a hypothetical scenario which tried to replicate the episode using existing technologies. To structure the discussion with a methodology, we used our tool, the Ethics Canvas, to understand the stakeholders involved, how the effects of the service on the behaviour and relations of stakeholders, and how the service may be used for unforeseen purposes. We concluded the discussion with a few action points addressing the issue of mitigating the ethical issues surrounding user-profiling.

Through this paper, we have hoped to present the argument that the technology itself should not form the basis of judgement over issues related to privacy, ethics, and morality. Rather, we need an open discussion involving the people who design and provide this technology, to identify such issues before the technology affects the general public. Such pragmatic discussions will lead to a better code of conduct, which may not be adopted into legislation immediately, but may be helpful in shaping the course of acceptable practices in the near future. One way to achieve this is through having an ethics-policy or an open approach to practising ethics towards responsible innovation. The ethics canvas is one such methodology and tool which encourages discussions in a manner that fit existing business practices. We envision such tools will help practitioners of responsible innovation communicate their good intentions to their intended users, for example, by publishing the ethics canvas for their service.

7 ACKNOWLEDGMENTS

This paper is supported by the ADAPT Centre for Digital Content Technology, which is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

The authors thank Wessel Reijers and other researchers involved in the development of the Ethics Canvas methodology.

REFERENCES

[1] K. Boda, Á. M. Földes, G. G. Gulyás, and S. Imre. User tracking on the web via cross-browser fingerprinting. In *Nordic Conference on Secure IT Systems*, pages 31–46. Springer, 2011.

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, L119:1–88, May 2016.

[3] A. Osterwalder and Y. Pigneur. *Business model generation: a handbook for visionaries, game changers, and challengers*. John Wiley & Sons, 2010.

[4] T. J. Pinch and W. E. Bijker. The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social studies of science*, 14(3):399–441, 1984.

[5] J. Stilgoe, R. Owen, and P. Macnaghten. Developing a framework for responsible innovation. *Research Policy*, 42(9):1568–1580, 2013.

[6] K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. In *Proceedings of the 13th international conference on World Wide Web*, pages 675–684. ACM, 2004.

[7] F. J. Zareen and S. Jabin. A comparative study of the recent trends in biometric signature verification. In *Contemporary Computing (IC3), 2013 Sixth International Conference on*, pages 354–358. IEEE, 2013.