Julien Achrafi, Adela Chelminski, Yuanyuan Gong, Vlad Serban
The Great Transition
19 December 2019

Note of Intent: *Kilp*

Problem

For about ten years now, the economy is experiencing what some call the 4th Industrial Revolution: The Artificial Intelligence Revolution. This revolution is based on the new capabilities of computers, thanks in particular to machine learning techniques, to process huge amounts of data. These new capabilities, mainly used in the digital world, can be used to have a much better understanding of the behavior of people and therefore to predict their consumption. Thus, the challenge for companies is to recover as much data as possible from Internet Users. In this context, new actors that have emerged in recent years, have become essential in this quest for data: Google, Amazon, Facebook, Tencent, Baidu, Alibaba. These large platforms have in common that they provide their services for free. Hence, a Facebook user, for instance, won't have to pay a penny to use the services of the social network. As a counterpart, Facebook can use the data generated by the user, to commercialize it, for example targeting him with highly personalized ads. By this new economic process based on free access, these large platforms have become giants of the global economy, with a capacity for influence that few companies have had in the past. Indeed, these companies have made themselves unavoidable: During the first semester of 2019, nearly 2.6 billion people all over the world used Facebook services every month.

This extremely important presence of these platforms on the Internet pushes users not to pay much attention to the detailed conditions of use of the platforms, and more particularly on the treatment that will be made of the data that they generate. Thus, these platforms derive their revenues from data exploitation of their not very observant users. The platforms have a policy of collecting as much data as possible, even if a large part of this data is not necessary for the direct operation of the platform. This policy of collecting "unnecessary" data is not only applied by these major platforms but also by minor websites. Hence, according to Forrester Consulting and its study Big data in Western Europe Today (2015), 70% of European companies acknowledge that they store too much unnecessary information. This unnecessary data collection is exactly the problem we want to address.[1]

What's mainly at stake is obviously the respect of internet users' privacy. With the increased importance that the internet has taken in our lives, this issue has been identified as a major challenge. This is the case in the European Union where a major regulatory text, the General Data Protection Regulation (GDPR), has been adopted to manage data processing in the EU, in particular on this issue of unnecessary collection of data: "Controllers should also implement mechanisms to ensure that personal data is not processed unless necessary for each specific purpose"[2] - Article 25, EU GDPR. However, this text is not enough to stop this data collection mainly due to the lack of means of control of the EU and because of the vagueness surrounding the detailed operation of the platforms. The sanctions provided by the

[1] "Big Data in Western Europe Today."
[2] "Article 25 EU General Data Protection Regulation (EU-GDPR)."

GDPR against platforms that are not in compliance are reduced and consist mainly of exclusion or significant fines, as shown by the threat of a $ 1.6 billion hanging over Facebook.[3] The difficulties encountered by the EU recently in ensuring that these major platforms pay their taxes in Europe without cheating emphasize how these solutions of a potential exclusion or significant fines are not realistic.

Consequently, there is a need for a tool that will truly ensure user privacy on the internet, and more specifically when they use the services of large platforms. The constraints will be to allow access to these platforms by minimizing the data transferred to them.

Proposed Concept

In order to better help individuals take control of their digital data, we shall design a desktop browser extension that blocks non-essential data collection requests from websites. This innovative solution will be called *Kilp,* meaning the Estonian word for “shield”, symbolizing the protective manner in which this app will help users. "Customer data can be collected in three ways – by directly asking customers, by indirectly tracking customers, and by appending other sources of customer data to your own."[4] The service will work as outlined in the graphic found on the Powerpoint: once a user accesses a website, the data request that is sent to the device by the website will be intercepted by *Kilp,* which will act as a filter on the data request. *Kilp* will analyze which data requests (include direct and indirect ones) are essential for the website to properly load, and only send that data request through to the user’s device. In that way, individuals will only send back the most essential data to the website to allow it’s proper functioning. Moreover, the tracking of users across their browser by websites once they have left that website will also be minimized by blocking domains or cookies that will not compromise the functioning of the site. Our system won't collect any personal data, it only acts as an intermediary filter to receive the data collection request from a website, analyze this request, and only transmit the requests for essential data to the individual’s device.

After we have a large number of users, we plan to allow users to take control of their data and use it as a form of currency. We will partner with websites (or information and data companies) to set up exclusive ways to exchange data for goods/services so that individuals can sell their data directly to these companies in return for something. For example, if a user hits a paywall for a news article, under this system they would be able to pay for a subscription to the newspaper with their data, that is allowing the newspaper access to "non-essential" data in exchange for access to articles behind the paywall. However, this function will not be implemented at the beginning because it requires a high quantity of users because data is currently extremely abundant and easily collected by websites, so it is worth very little by individual. In order to increase the worth of individual’s data and actually incentivize companies to partner with our system, data will need to become more scarce and therefore worth more. This can only happen if data protection becomes more robust and if we have millions of users signed up for our service. In terms of measurement of effectiveness, we could visionalise the effect by showing customers the records of blocked requests in their personal page.

---

[3] Koch, “The GDPR Meets Its First Challenge.”

[4] Uzialko, “How and Why Businesses Collect Consumer Data.”

In order to get people to know our service, we will carry out an educational marketing campaign that will also raise people's awareness of the importance of data privacy. This is how we plan to launch our app, through short videos that explain the dangers of your personal data being collected and sold by websites, finishing the videos with the link to our website so that people can download our extension and join.

Today a user can protect themselves from tracking and data collection by meticulously adjusting their privacy settings and employing different ad blockers, VPNs, and tracking blockers. Notable examples are AdBlocker, Privacy Badger, and a variety of VPNs. However, none of these services provide a fully integrated platform that performs multiple functions in an effort to protect user data. While the proposed concept might seem highly complicated at first glance, in reality it is easily replicable across websites once it is implemented. Most websites request the exact same data and deposit similar trackers, therefore once our code is successfully able to analyze the data requests and filter out the non-essential requests, it will be able to do so for most websites out there. The software will be AI, which relates to the "Artificial intelligence, natural stupidity" theme of The Great Transition.

Expected Positive Impacts

Kilp helps enhance cybersecurity and the right to privacy for all individuals. In the EU, data protection is a fundamental right, and the GDPR is the new framework for protecting that right. Our new product is to safeguard the individual right to privacy corresponding to the GDPR. *Kilp* as a data privacy solution could decrease potential threats of surveillance by governments (domestic or foreign) as well as decrease risks of identify fraud or hacks into your accounts. Moreover, *Kilp* would allow users to take back control of their own data, ensuring it isn't being (legally) used and sold by corporations and data broker firms, whether merely for targeted advertising or for more serious means that can have negative impacts. For example, companies have been known to share sensitive user health data with other companies, presenting huge medical privacy risks and even the possibility that such information can be sold to health insurance providers, adversely affecting the insurance plans people would be offered.[5] The protection of personal data will promote the fair treatment and equality in the society.

Beyond, Kilp could bring positive impacts on sustainable development as it increases accreditation and confidence to entire cyber environment and would foster economic growth as a whole. In addition, the educational aspect of Kilp is important to raising awareness of how individual's data is used and processed, thereby increasing people's understanding of the digital world and simultaneously empowering them to take control of it.

Major risks and actions to reduce them

Beyond benefits, we have identified several existing and possible risks of our project. They come from many different sources, for instance, they would be technical and external business environment etc. In order to manage these risks and take preventative actions, we list relevant threats, evaluate the likelihood of these risks being realised, and their possible impact.

[5] Neuman and Domonoske, "Grindr Admits It Shared HIV Status Of Users."

Risk: Value matching between customers and Kilp
Summary: Our product is designed to satisfy individuals' needs on data privacy. However, if people have not been aware of the necessity of data protection, and even without knowing their data has been collected and may be used in a detrimental way, people would tolerate the invasion of privacy.
Likelihood: 4/5
Impact: As a consequence, less expected number of subscribers would be hard for us to make profits and less attractive for investors. And without enough quantity of subscribers, the second phase of our project, commodification of data can not realise.
Countermeasures: We therefore advocate the necessity of data protection and educate our customers through various educational marketing campaigns. An educational video is planned to be made to show how data collection impacts our private life and can have detrimental negative effects. Meanwhile, if possible, it would be nice to cooperate with relevant institutions for data privacy education which could also be regarded as a part of our CSR policy.

Risk: Reputation and effectiveness
Summary: As a data protection product, Kilp needs to build trust with customers and convince them that our product will not collect personal data. Besides, unlike ad-blocker that customers could easily measure its effectiveness by seeing less targeted advertisements, the effectiveness of Kilp is not that apparent.
Impact: A reputation damage would not be compensated by other aspects and largely affect future economic performance. Similarly, an effective and trustworthy product can gain more customers and build positive reputation.
Likelihood: 2/5
Countermeasures: Kilp needs to demonstrate transparency, professionalism, and that it will not collect nor trace individual information by demonstrating how Kilp works in a simple way. In terms of measurability, users of Kilp are able to see rejected data collection requests from their personal profile page.

Risk: Technological issue
Summary: As Kilp is a new innovative product, we are suffering from several technical difficulties (code smell and block customization) to finalise the product
Impact: The technological issue will ultimately determine if we could officially launch the product, Kilp and break into the market.
Likelihood: 2/5
Countermeasures: In order to overcome technical issue, we would invite external consultants to contribute to it, but it may increase the cost.

Risk: External environment
Summary: The appearance of Kilp would definitely damage the interest for personal data collecting and selling companies, the commercialisation of Kilp would be opposed by them. Besides, there are threats of other similar competitors in the market and Google chrome, Firefox are also intended to develop this kind of browser.

Impact: Opponents are significant and retaliation would be possible. If Google or Firefox enters the same market, we shall loose part of the market share.
Likelihood: 3/5
Countermeasures: In this situation, we have to take the preemptive opportunities in the market and build customer stickiness. Even though there are threats from other competitors *Kilp* still has its competitive advantage. Compared with existing and conceptual products, *Kilp* would be capable to protect individual data privacy in a more multifaceted way. Apart from that, *Kilp* would start with a free version and then plan to launch a paid premium version.

Deployment Strategy and Major Milestones

Our solution could be implemented directly following a development phase, which would include its conception (through coding) and testing, followed by a marketing phase (with an educational focus) which would attract users. There do not seem to be any technological or regulatory barriers to its implementation.

After developing *Kilp*, it would be further enriched through AI learning which would expand its usefulness to different types of websites. After its launch, an open-source platform would further enhance its scope and help overcome any shortcomings or flaws. This is in line with our user-focused and participative approach to the project.

Following its development into a functioning tool, we would deploy a marketing campaign focused on describing the issue at hand (data privacy), and educating our prospective users on how our solution would empower them in the transmission of their private data to websites whilst protecting their privacy.

ROI Analysis

Initial investment is low, development costs include a small team (2 developers in addition to the 4 creators of the project) and a workspace. For the creation of a minimum viable product, we estimate a maximum cost of 50 000€ based on other AI and machine-learning products.

During the first two years, costs would mainly be focused on the development team for our solution and its AI learning phase, whilst trying to minimise the impact of the Operations Team's (the four creators of the project) salaries. Our other cost driver, marketing, would be mainly financed by crowdfunding.

After the first two years, we expect to have achieved enough market penetration so that we can introduce a premium version of our product, which would monetize user's data for their personal use. Even with a relatively low price (3,6€ per year, equating to 0,3€ per month), the sheer volume would allow us to achieve a profit as soon as the third year of functioning.

Organizations

At first, sponsors of the project would be the users themselves, through two annual crowdfunding campaigns, as well and qualitatively through feedback and the open-source platform. Partners could range from technology and IT publication to internet-based companies who could partner with us in order to "get ahead of the curve" on the issue of data privacy.

Works Cited

Privacy Regulation. “Article 25 EU General Data Protection Regulation (EU-GDPR),” September 5, 2018. https://www.privacy-regulation.eu/en/article-25-data-protection-by-design-and-by-default-GDPR.htm.

“Big Data in Western Europe Today.” Xerox, 2015. https://www.xerox.com/downloads/usa/en/b/Big_Data_in_Western_Europe_today.pdf.

Koch, Richie. “The GDPR Meets Its First Challenge: Facebook.” GDPR.eu, January 4, 2019. https://gdpr.eu/the-gdpr-meets-its-first-challenge-facebook/.

Neuman, Scott, and Camila Domonoske. “Grindr Admits It Shared HIV Status Of Users.” NPR, April 3, 2018. https://www.npr.org/sections/thetwo-way/2018/04/03/599069424/grindr-admits-it-shared-hiv-status-of-users.

Uzialko, Adam. “How and Why Businesses Collect Consumer Data.” Business News Daily, August 3, 2018. https://www.businessnewsdaily.com/10625-businesses-collecting-data.html.