CAP-A: A community-driven approach to privacy awareness

Alexandru Stan[1], Konstantina Geramani[1], George Ioannidis[1] ,*Ioannis Chrysakis[2,4], Giorgos Flouris[2], Maria Makridaki[3], Theodore Patkos[2], Yannis Roussakis[2], Georgios Samaritakis[2], Nikoleta Tsampanaki[2], Elias Tzortzakakis[2] and Elisjana Ymeralli[2].*

[1] IN2 Digital Innovations GmbH, Germany
[2] FORTH-ICS, Greece
[3] FORTH, PRAXI Network, Greece
[4] Ghent University, IDLab, imec, Belgium

In an increasingly instrumented and inter-connected digital world, citizens generate vast amounts of data, much of it being valuable and a significant part of it being personal. However, controlling who can collect it, limiting what they can do with it, and determining how best to protect it, remain deeply undecided issues [1].

CAP-A[1] deploys a socio-technical solution based on collective awareness and informed consent, whereby data collection and use by digital products are driven by the expectations and needs of the consumers themselves, through a collaborative participatory process and the configuration of collective privacy norms (see Fig.1). The proposed solution creates a new innovation model that complements existing top-down approaches to data protection, which mainly rely on technical or legal provisions.

Fig.1 An overview of CAP-A project's approach

The ongoing project aims to deliver a global repository of consumer and developer generated content about the privacy behaviour of mobile apps, along with tools that will help consumers

[1] https://cap-a.eu/

understand the Terms of Service, Privacy Policies and other legal documents (ToS) and their implications via crowdsourced approaches and visual cues.

The objective is to foster collective intelligence and co-creation of solutions, and to facilitate the participation of all involved stakeholders through an open architecture, thereby allowing novel uses of the privacy-related content. Ultimately, the CAP-A ecosystem will strengthen the trust bond between service developers and users, encouraging innovation and empowering the individuals to promote their privacy expectations as a quantifiable, community-generated request.

Novel Approach

One of the novel concepts of the projects is the formulation of collective privacy norms, as explicitly declared through privacy expectations. These are connected with measurable features, offering a wealth of data for the different stakeholders (developers, social scientists, policy makers) to conduct analyses and interpret the behaviour and mindset of users. This is particularly relevant for industrial stakeholders, who will be able to better understand the privacy needs of their (potential) customers and to turn them into a competitive advantage for their products.

Thus at the heart of CAP-A is a *Semantic Privacy Repository* which contains the privacy-related and ToS information regarding digital products. The repository combines the benefits of semantic technologies with collaborative editing capabilities that enables the user to express privacy preferences about each product or category of products, while at the same time permitting developers to explain their policies and automatically access the underlying data. Additionally, it offers a public place for experts to post findings about products. Thus the Semantic Privacy Repository will collect information on concepts like:

- *Privacy Policy Practice (PPP)*: a specific privacy-related process associated with the app, such as “access to camera”, “minimisation of data collected” etc. Each PPP is characterized by a set of attributes, which may be requests, expectations or evidences
- *Privacy Policy (PP)*: a set of Privacy Policy Practices that a given application implements.
- *Request*: a triple <PPP, request_attribute, value> issued by a developer, e.g., <access_to_camera, purpose, unspecified>, <access_to_camera, duration, session>. Requests specify the privacy characteristics of an application, as specified by the developer.
- *Expectation*: a triple <PPP, request_attribute, value> provided by an end user to declare her preference towards an app behaviour she considers reasonable. Expectations are identical in form to requests, but they are issued by users.
- *Evidence*: a triple <PPP, evidence_attribute, value> issued by a developer or an end user, where the evidence_attribute can specify either a text segment found in the ToS documents or a URL.

CAP-A end-users can access the CAP-A Portal and the CAP-A Mobile application (see Fig.2).

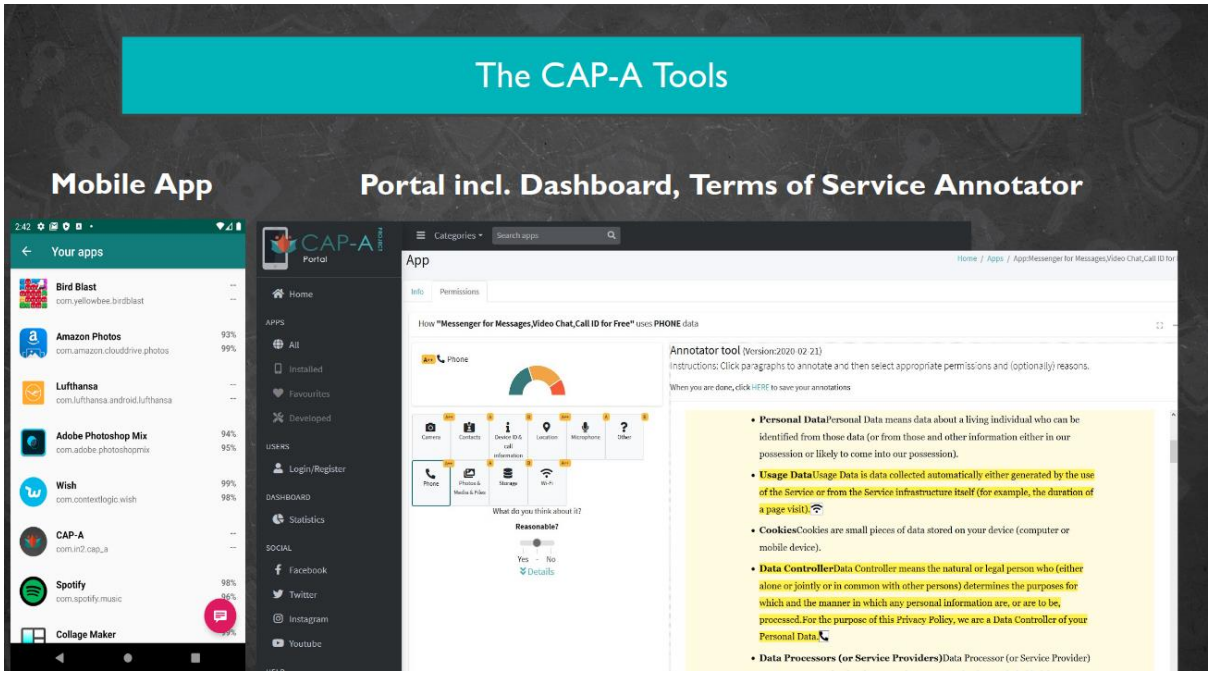


Fig.2 The CAP-A tools

The *CAP-A Portal* is a web-based application that enables anonymous users to see basic information about the data collected in the *Semantic Privacy Repository* while at the same time allowing registered users to participate in the CAP-A community, actively expressing their PPPs about mobile applications (they use), annotating ToS documents or exploring the collected data through a *Dashboard*. A rich user profile allows for the specification of preferences. Users who are app developers can also use dedicated functionality on the portal to "claim" the applications they developed and explain in more detail, through the *Terms of Service (ToS) Annotator,* why certain data policies were chosen.

Each mobile application receives two scores:

- The community score, which is calculated based on how close the community's expectations are to what the app is requesting.
- The privacy friendliness score, which is based on how honest the app is with regards to its privacy-related requests (e.g., if there are evidences that it is operating differently than what is specified in the ToS documents), how clear the ToS documents are, how frequently the ToS documents change etc.

Users of the portal can search or browse the list of Android mobile applications. Once the details of an app are explored, the user gets more information about it, its CAP-A scores and is invited to express her expectations about the PPPs of that app, i.e. for each device permission requested by the app the user can indicate, using an opinion slider, if she finds reasonable or not the request to grant access to that type of data. Moreover, experienced users can also share interesting articles about the app, also rating their credibility.

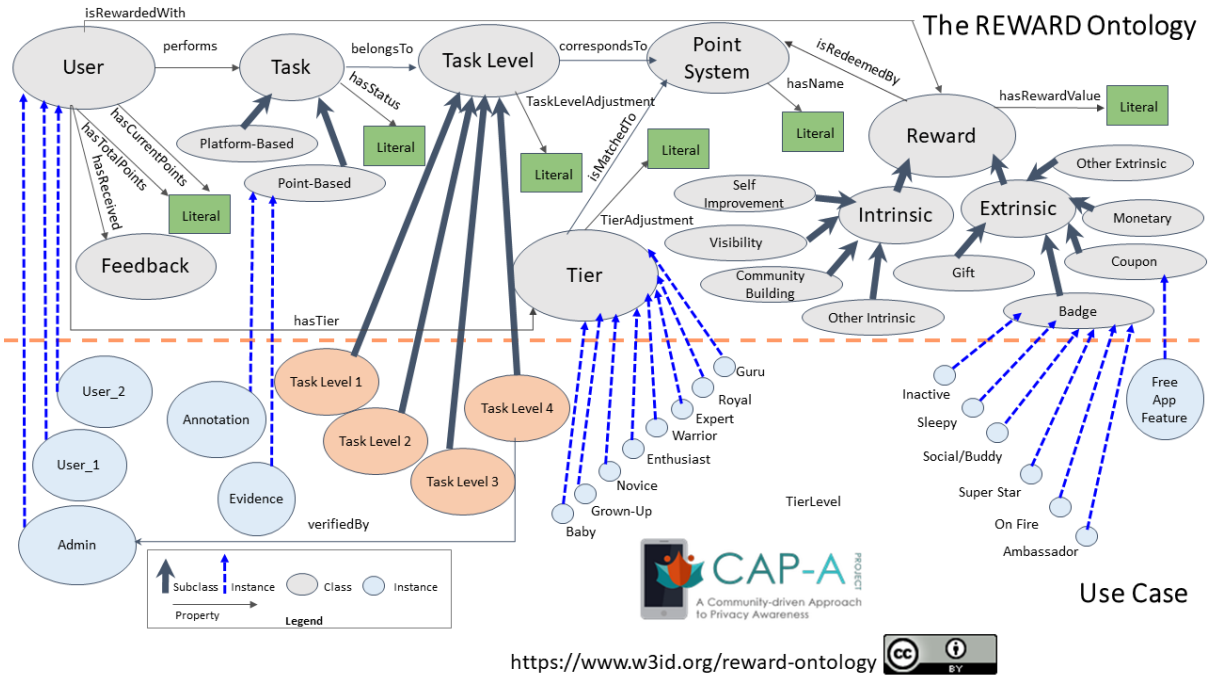
CAP-A users can also choose to install a dedicated CAP-A native Android application. *The mobile CAP-A app* allows users to easily get an overview of how privacy conscious the apps installed on their phone is. It does this by conducting an "audit" of the user's apps, and providing as a result an overview of the each app installed and their CAP-A scores. Users can

then easily see which apps have been positively or negatively evaluated by the CAP-A community and can themselves choose to express their expectations. For registered users the CAP-A mobile apps is connected to the portal allowing favourites and installed apps to be marked. Finally, the mobile app also serves the purpose of anonymous data collection, informing the Privacy Repository about which apps are installed.

User engagement strategy

In CAP-A we combine an awareness campaign with a reward and engagement strategy, in order to maximize user participation and contribution. Crowdsourcing can give a solution to the direction of privacy awareness in several tasks, such as annotating Privacy Policy/Terms of Service documents, identifying GDPR concepts among them, and evaluating the privacy friendliness of apps and services. However, one of the most fundamental challenges in crowdsourcing (in general) is how to recruit and evaluate users to take advantage of their contribution. Therefore, there is a need to engage users through intrinsic (fun, autonomy, reputation) or extrinsic motives (money, learning, forcedness, implicitness, task autonomy) [2]. For this reason, a well-defined reward mechanism is required to enable user engagement and ensure the success of the total approach.

In our reward strategy, we combine both intrinsic and extrinsic types of rewards as introduced in successful reward mechanisms or loyalty programs that are being used in the industry or in several crowdsourcing rewarding papers ([3], [4], [5], [6]). Below we present our rewarding mechanism's conceptual model which is based on REWARD Ontology [7]:



Acknowledgements

This work has been supported by the CAP-A project which has received funding from the European Union's Horizon 2020 research and innovation programme under the NGI_TRUST grant agreement no 825618.

References:

[1]. Giorgos Flouris, Theodore Patkos, Ioannis Chrysakis, Ioulia Konstantinou, Nikolay Nikolov, Panagiotis Papadakos, Jeremy Pitt, Dumitru Roman, Alexantru Stan, Chrysostomos Zeginis, “Towards a Collective Awareness Platform for Privacy Concerns and Expectations”, Confederated International Conferences: CoopIS, C&TC, and ODBASE 2018, Valletta, Malta, October 22-26, 2018

[2]. Yang, D., Xue, G., Fang, X. and Tang, J. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. In Proceedings of the 18th annual international conference on Mobile computing and networking (pp. 173-184). ACM (2012)

[3]. Antikainen, M.J. and Vaataja, H.K.. Rewarding in open innovation communities–how to motivate members. International Journal of Entrepreneurship and Innovation Management, 11(4), pp.440-456. (2010).

[4]. Scekic, O., Truong, H.L. and Dustdar, S.,. Incentives and rewarding in social computing. Communications of the ACM, 56(6), pp.72-82 (2013).

[5]. Kavaliova, M., Virjee, F., Maehle, N. and Kleppe, I.A.. Crowdsourcing innovation and product development: Gamification as a motivational driver. Cogent Business & Management, 3(1), p.1128132. (2016).

[6]. Cappa, F., Rosso, F. and Hayes, D.. Monetary and Social Rewards for Crowdsourcing. Sustainability, 11(10), p.2834. (2019).

[7]. Chrysakis, Ioannis, Giorgos Flouris, Theodore Patkos, Anastasia Dimou, and Ruben Verborgh. "REWARD: Ontology for reward schemes." In 17th Extended Semantic Web Conference: Posters and Demos, pp. 1-5. (2020).