Abstract
The thesis develops methods for realizing trustworthy privacy assistants to help people preserve their privacy online. The thesis makes two major contributions.

First, it develops the principles and algorithms for an uncertainty-aware personal privacy assistant PURE. Using only the visual characteristics of images, PURE learns users' privacy preferences and creates a predictive model. When PURE makes a prediction (i.e. private or public) for each image, it also captures the ambiguity of privacy by calculating a level of uncertainty for its prediction using the Evidential Deep Learning model. When PURE is uncertain about an image, it delegates the privacy decision to the user. Otherwise, PURE uses its own prediction results (i.e. share or not share). Moreover, it takes into account the user's persona encompassing their risk perception, personally labeled data, and preferences for consultation. This personalized approach ensures that PURE adapts its behavior for each user, minimizing their perceived risk of privacy violations. Another crucial characteristic of PURE is that it does not require access to any additional private user information, such as personal details.

Second, this thesis designs an explainable privacy assistant PEAK that generates explanations of why a given image is considered as public or private. PEAK automatically generates tags to describe images. Afterwards, PEAK employs topic modeling techniques to extract hidden topics from the descriptive tags of images. PEAK captures explanation templates that are based on the relationship between images and their associated topics. Using a unique categorization of prediction classes, it generates explanations automatically. The user study that is conducted to evaluate the generated explanations by PEAK shows that the users find the explanations sufficient, satisfying, and understandable. Then, PEAK’s performance is also evaluated in terms of enhancing the privacy assistant PURE by helping it make predictions about difficult-to-classify images through the use of image explanation categories. PEAK is able to reduce the number of delegated images to the user without compromising the accuracy of privacy decisions.

Publications

Journals


Conferences, Workshop, Doctoral Consortium, and arxiv


Defense Jury Members

1. Assoc. Prof. Arzucañ Özgür  
   Bogazici University
2. Prof. Pınar Yolum  
   Utrecht University
3. Prof. Lale Akarun  
   Bogazici University
4. Assist. Prof. Fatma Başak Aydemir  
   Bogazici University
5. Prof. Şule Gündüz Özgüdücü  
   Istanbul Technical University
6. Assoc. Prof. Reyhan Aydoğan  
   Ozyegin University

**Defense Date:** 21.08.2023