User Profile Acquisition: A Comprehensive Framework to Support Personal Information Agents

Gerrit Kasper, Diego de Siqueira Braga, Denis Mayr Lima Martins and Bernd Hellingrath
Westfälische Wilhelms-Universitt Münster, ERCIS
Leonardo-Campus 3, Münster, 48149
Email: {gerrit.kasper, diego.siqueira, denis.martins, bernd.hellingrath}@uni-muenster.de

Abstract—**Personal information agents play an important role in the acquisition, processing, and distribution of information and are of utmost importance in many areas. However, in order to meet user's information needs, these agents must direct their tasks by reasoning upon the user-related information. In this sense, user profiling plays a crucial part. User profiling is an attempt to deal with the information about Internet users and its diversity and complexity through categorization. In this paper, an overview of the user profile acquisition is provided. For this purpose, a framework is presented which describes the most important aspects and frame conditions of user profile acquisition. The three user profile types Explicit, Implicit, and Hybrid are discussed, and user profile acquisition and enrichment techniques are elaborated for each of them. Furthermore, the challenges regarding user profile acquisition and the legal requirements are presented.**

I. INTRODUCTION AND MOTIVATION

In today's information society, modern systems and technologies produce a huge amount of data and information which volume, velocity, variety, and complexity are constantly increasing [1], [2]. This leads to the challenge of handling the information overload [3]. An essential part to reduce the omnipresent information overload is the selection and anticipation of information that provides the most value. Estimating and preparing the most valuable information and selecting the most appropriate communication, distribution and operation methods rely on the details and their accuracy that are known about the information recipients (i.e. users). A user could either search for information by oneself or be a designated information recipient as well as a subject of an action performed by a certain party.

For these reasons, personal information agents [4] have been introduced to help users in tasks such as information filtering, content retrieval, news recommendation, and searching. However, in order to better attend user's information needs and consumption style, these agents must rely their processes and activities on models that capture (at some extent) user's interests and behavior. In this sense, profiling is an attempt to support agent decision-making by the identification and understanding of underlying patterns, in particular, situations [5] involving the target user. Profiling techniques are applied in many cases, for instance, to provide personalized content in adaptive web-pages [6] or to segment customer groups for targeted advertising campaigns [7].

In the context of user profiling, the information that describes and characterizes the user is constituted in user profiles which are ultimately the basis for the application of actions and policies. In this paper, it is referred to a combination of related user profile definitions from Brusilovsky, Kobsa & Nejdl [6], Gauch et. al. [3], Gao et. al. [8] and Eirinaki & Vazirgiannis [9] in which a user profile is described as "*a set of user-centric data models representing appropriate approximations of the interests and activities of either a group of users or a single person. It contains demographic information as well as information about users behavior, intentions, interests and preferences*". User profiles or rather useful information about the users need to be extracted from different sources which are referred to as the user profile acquisition.

The goal of this work is to provide an overview of user profile acquisition as part of the user profiling process by answering four coherent research questions: (I) *How can user profiles be acquired?* (II) *What are user profiles and their characteristics?* (III) *Which techniques could be used to acquire user profiles?* (IV) *What are possible implications and challenges of user profile acquisition?*

To answer these questions a user profile acquisition framework is created (see Figure 1) and three different user profile types are identified. Further, for each possible data extraction methods are elaborated and their advantages and disadvantages discussed from a companys or researchers perspective.

The profile acquisition framework provides an overview of the topic and describes the user profile acquisition in two phases: User Profile Extraction and Enrichment. Besides, two further aspects: the User Profile Dimensions, which are related to the three user profile types Explicit, Implicit and Hybrid, and the Legal Requirements which describe the legal frame conditions for the techniques of the two phases. The first phase User Profile Extraction, which is primarily the in focus of this paper is subdivided into the three user profile types and provides an insight into the three different user profile acquisition categories. In the second phase Enrichment (5.), further techniques are presented that provide additional information and details on the acquired user profiles.

---

[1]Some of the used icons, or parts of them are created by various authors from thenounproject.com. A complete list is provided in the [10].
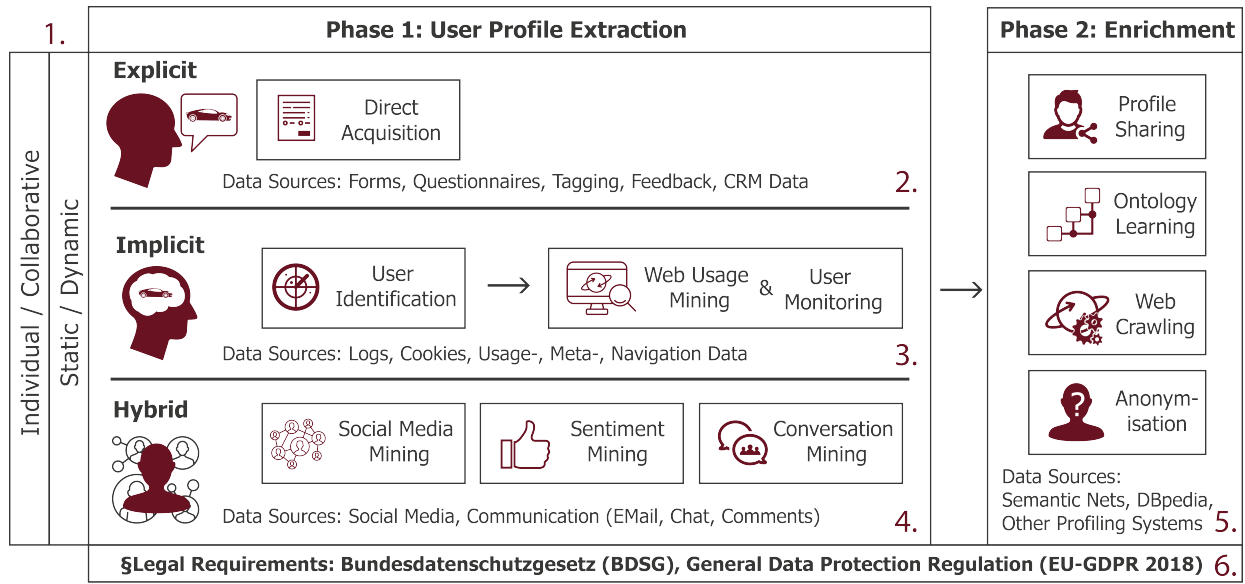
Fig. 1. User Profile Acquisition Framework[1].

II. Background

A. Literature Research Protocol

In order to answer the research questions (see Section I), the following data sources are used to provide an overview of the existing relevant academic literature in the area of user profiling and user profile acquisition from the year 2000 to 2017: Web Of Science, IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect and Google Scholar. The found literature was then analyzed regarding the potential fit of entries. Selection criteria were the domain area, number of citations, relevance scores provided by the data sources, and the authors. These data sources were used for forward and backward search in order to identify further relevant publications that are citing a certain primary study or are included in a publications bibliography.

B. User Profiling Process

The process of user profiling refers to the process of building people's profiles by means of data-driven methods [5]. The general purpose of user profiling is to collect information about the users and their interests as well as the period of validity, in order to improve the information quality and to infer the users intentions [3]. The roots of user profiling can be found in the Knowledge Data Discovery model of which many steps were adapted [11], [12]. The user profiling process comprises of six steps (i.e., preliminary grounding, profile acquisition, profile construction, profile learning, interpretation, and application) which can involve iterations and can contain loops between any two steps [11].

In the first step, Preliminary Grounding of the user profiling process, an understanding of the application domain and the relevant prior knowledge is developed, and the goals are specified. In the second step, Profile Acquisition, which is also referred to as Profile Extraction [13], the data is collected from different selected sources. In the third step, Profile Construction, the data is pre-processed to increase the data quality and to reduce its complexity. Then it is analyzed, and user profiles and models are constructed [9], [3]. In the fourth step, Profile Learning, the data is processed with different filtering and learning approaches to find out which data elements and users match [14]. The profile learning can be considered as a continuous process which ensures precise and up-to-date user profiles. In the last two steps, Interpretation and Application, the constructed profiles are evaluated on their relevance and adapted to specific domains. This includes the definition of actions and policies which are applied when a group or user match a relevant profile [11].

III. Overview of User Profile Dimensions, Types and Acquisition Techniques

As defined in section I a user profile is a set of user-centric data models that represent one or more users and their attributes, which are called dimensions in the following. Based on the applied user profile acquisition techniques and the users awareness, the user profiles can be differentiated into three types: The Explicit, Implicit and Hybrid User Profiles. An overview of the user profile types as the top three elements and their dimensions is provided in Figure 2.

A. User Profile Dimensions

One dimension of the user profiles is the coverage. A user profile can either be related to each user *individually* or as an aggregate user profile to a group of users (*collaborative*) [9], [5], [15]. One drawback of individual profiles is the lack of serendipity, for instance, if they are used to provide recommendations, as they are only focusing on one users previous interests [6].

Another dimension describes the timeliness. A user profile can either be *static* or *dynamic*, as well as its attributes. Static
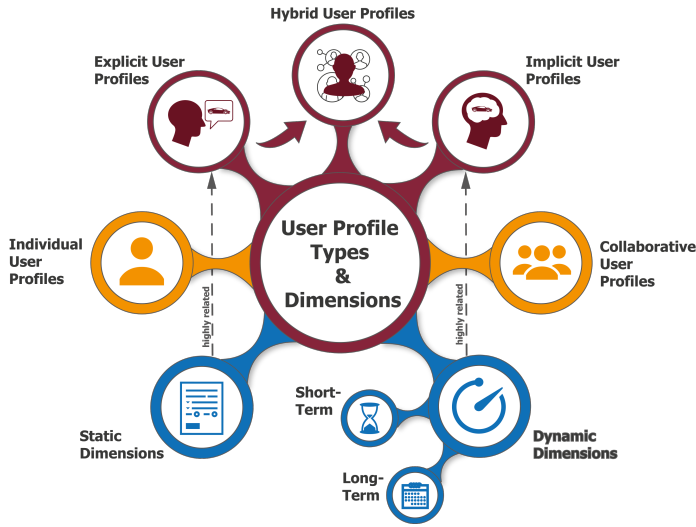
Fig. 2. Overview - User Profile Types and Dimensions[1].

user profile dimensions maintain the same information over time or are very unlikely to be altered, whereas dynamic profiles dimensions are considering a time aspect which can be differentiated between *long-* and *short-term* dimensions [9], [3]. Static user profile dimensions or attributes could represent for instance demographic information such as the name, surname, gender, marital status, relatives, and date of birth or origins.

In contrast, dynamic user profile dimensions are more likely to change at a higher rate. Long-term dimensions or attributes could comprise of geographic, socioeconomic or personal and psychographic information for example, like the current employment, revenue, interests, hobbies, and ideals. Short-term user profile dimensions cover information that is very likely to change in a short period of time. They often represent information about the users lifestyle, for instance, the current motivation, spontaneous interests, activities or attitudes. A further example of long- and short-term dimensions is provided by Gauch et al. [3], in which a musician searches the web for her daily research and a spontaneous vacation. The website provider could extract the musicians interests about certain types of music which are considered as a part of the long-term dimensions, whereas the vacation-related interests are part of short-term ones.

To maintain the dynamic user profile attributes various forgetting and updating mechanisms are used on a regular basis [16], [17]. These forgetting methods are part of the profiling process step Profile Learning.

B. User Profiles Types and Acquisition Techniques

1) Explicit User Profiles: An explicit user profile requires the users active participation and is often the initial outcome of a user profile acquisition process in which the user creates the profile on its own. This includes that the users are most probably aware of the provision of their personal information [18], [12], [15]. Explicit user profiles could be seen as a

framework with the focus on the current state of a user and less about a possible anticipation. Generally, these profiles are referred to as master data, and they are highly, but not exclusively, related to static characteristics such as the address, the name and surname, marital status, bank details, and explicitly stated interests or knowledge. They are usually generated within a registration process as a series of questions designed to acquire users data precisely, which can take place in online forms or offline surveys [6], [4].

An overview of the explicit user profile acquisition techniques is provided in Table I. In general, considering their characteristics, it is recommended that the following explicit techniques not be solely used, instead of in combination and with the use of implicit or hybrid user profile acquisition techniques to acquire more valuable user profiles.

2) Implicit User Profiles: The implicit user profiles describe everything about the users which can be inferred by observing the users behavior and the related metadata. In contrast to the explicit user profiles, the users do not necessarily provide the information on purpose or are aware of it. Instead of describing the current state of the user, the implicit user profiles focus more on the anticipation of the users characteristics and interests, in the means of trying to learn something about the users [4], [15]. Information and data which could be acquired from the users are for instance, the device information such as the hardware and network environment details, the topics the users are interested in (e.g. consultation or purchases of items), and the users intentions, goals and behavior pattern (e.g. inferred from analyzing the mouse movement or navigation logs) [8], [18]. Since changes of this kind of information are more frequent, the acquisition of implicit user profiles is more related to the dynamic dimensions [9].

The next step after the profile acquisition is the pre-processing and the model construction. In contrast to the pre-processing of explicit user profiles e.g. by analyzing free-form surveys, the pre-processing of implicit user profiles could be executed more efficient and precise with automated methods which reduce the cost of pre-processing [35].

An overview of the seven implicit user identification and user data extraction techniques is provided in Table II.

3) Hybrid User Profiles: The hybrid user profiles are the combination of the implicit and explicit ones. Static attributes of a user are enriched and extended with the dynamic information acquired such as the implicit user profiles. This type of user profiling helps to be more efficient and maintains the accuracy since information gets updated temporally [12], [48]. By the proposed hybrid profiling approach with the timeliness dimensions static and dynamic and the objective dimensions content and collaborative from Poo et al. [15], it can be deduced that hybrid user profiles can be represented by the comprehensive form of social media platforms and the data they collect explicitly and implicitly from the users, their relationships and their actions.

Hybrid user profiles are used in many areas of application such as personalization, marketing, recommendation, police

TABLE I
EXPLICIT USER PROFILE ACQUISITION TECHNIQUES OVERVIEW.

Concept	Advantages	Disadvantages
Forms and Surveys [19], [20], [21], [22]	1. Different degrees of freedom (free text, drop-down, checkbox) 2. Users are familiar with it 3. Easy implementation, adaption 4. Easy analysis of data	1. It is necessary to provide a value to the users 2. Desired evaluations and questions must be known 3. Forms are intentionally completed inaccurately
Ratings [23], [24], [25], [26], [27], [28], [29]	1. Different degrees a) Unary (e.g. "Facebook-Like") b) Binary (e.g. "good/bad") c) Numerical/Ordinal 2. Prediction for unrated category items 3. Simple approach for users feedback	1. Not always reliable and consistent 2. Users tend to notify and rate already high rated items 3. Cold start problem
Tagging [30], [31]	1. Initial preference acquisition phase 2. Free nature of tags 3. Results are Folksonomies which relate users, items, and tags	1. Semantics are not necessarily the same for all users 2. Users might be unsure about what they are tagging 3. Different items with same tags to save time
Dialogs and Comments [32], [33], [34]	1. Adaptive Text-Based Natural Language Dialogs 2. Configurable dialog management 3. Analyze and extract values 4. Straightforward questions that are adapted based on what is already known about the user 5. User is expected to actively ask questions	1. Novice users may not be capable of formulating questions 2. Users overestimate the programs capabilities which disappoint them 3. High acquisition effort

TABLE II
IMPLICIT USER PROFILE ACQUISITION TECHNIQUES OVERVIEW.

Concept	Characteristics
Authentication [36]	1. Unambiguous identification (Login) 2. Multi-device 3. New user problem
Cookies [9], [37], [38], [39], [40], [41]	1. Stored at the client and used within subsequent requests 2. Session identification (e.g. Clickstreams) 3. Time, requests, domain, prior sites 4. Persistent Cookies (JavaScript, Flash)
Browser Configurations [41], [37], [42]	1. Browser User Agents, Version, Plugins 2. SSL/TLS Handshake, Ciphers 3. Specific, unique configurations 4. IP-Addresses, Identd, OS
Browser Add-ons & Desktop Agents [43], [3], [21], [40]	1. Tools that observe the users actions 2. Unambiguous device identification 3. All user files and activity available 4. Requires user to install software
Third-Party Networks [44], [40], [41]	1. Social Sharing Buttons 2. Provide content, methods for other websites such as 3rd-party networks 3. Requests script from platforms with personal information
Canvas Fingerprinting [40], [39], [41]	1. Invisible graphic that generate each time on a request (e.g. GET/pixel.gif) 2. Generation depends on hardware, software configurations 3. Easy to identify navigation paths
Ultrasound Tracking [45], [46], [47]	1. uBeacons emitted on websites or TVs 2. Determine what user consume 3. Multi device tracking up to 7 meters 4. Not noticeable (18-20 kHz) 5. Currently on Millions of Smartphones 6. Software from 10 companies (e.g. Silverpush, Shopkick, Signal360, Audible Magic)

TABLE III
HYBRID USER PROFILE ACQUISITION TECHNIQUES OVERVIEW.

Concept	Characteristics
Direct Database Access [50], [51]	1. No limitations (e.g. privacy settings) 2. Partnering with platform owner 3. Research data donations (i.e. very large, structured, anonymized, aggregated)
Sanctioned API [52], [51], [53], [54]	1. Social Login (SSO) a) Explicit techniques unnecessary b) Privacy considerations (33c3-7827) 2. Data / Stream collection (structured) 3. Incomplete data availability (limitations)
Interface Crawling [53], [51], [55]	1. Web scraping techniques 2. Terms and Conditions, legal uncertainties (Robots.txt) 3. Applications and libraries available (e.g. http://socialmediadata.wikidot.com)
Social Media Data Vendor [56], [57], [58]	1. Global data vendors or marketplaces allow to buy or sell data, e.g. Acxiom, Experian, Merkle and Epsilon

investigations and much more [5]. As with the other profile types, the availability and extent of the profile data depend on the users attitude towards entering information and making it visible in his or her profile [49]. Furthermore, the data can be very noisy, contain mistakes or made-up statements, and spammers generate a high volume of inappropriate or unusable data.

In Table III four hybrid techniques to acquire user profiles from social media platforms are presented (i.e. Direct Database Access, Sanctioned API Access, User Interface Crawling and Data Brokers.)

IV. Few Words on Privacy Challenges in User Profile Acquisition

The frame conditions of the user profile acquisition are defined by the laws and the concerns of the users[2]. In all user profile acquisition activities, the privacy concerns of users have to be taken very seriously and should be addressed proactively, even if the actual risks of privacy invasion are low. Security and privacy infrastructures such as privacy policies [60] can be established to inform users about the profile acquisition, their later usage and the intended purposes of that, and the benefits for the users of providing (personal) information in order to establish trust and comply with privacy laws in many countries [22].

In many cases, the consent of the user is recommended or obligatory. It is described as the most critical factor of legitimizing on-line user identification and profile acquisition techniques [17]. The usage and processing of such data might entail specific disadvantages for the users after they have (unwittingly) provided the data (cf. section III). These disadvantages can be of an ethical, legal, material or ideal matter. Hazards and risks that may be associated with this are for instance: stalking, loss of freedom, weakening of democracy, manipulation of users and information filtering, economic disadvantages, discrimination, impact on the health of users, lending, economic espionage, and preparation of IT attacks or criminal offenses [40].

In the private sector, a comprehensive consent of the user is provided by him or her agreeing to the terms and conditions of using a platform or website, e.g. Facebook or, Google, to collect process and store any kind of data [61].

V. Conclusion

User profiles and thus also user profile acquisition techniques are of great importance in dealing with information overload and information management in various ways, notably in the context of information agents which must direct their tasks towards user's interests and needs. In this sense, this work presents a framework of user profile acquisition which describes the most important aspects and frame conditions of user-related information gathering.

Three user profile types were identified and distinguished from each other, namely, explicit user profiles, which require the users active participation, implicit user profiles, which are inferred by observing the users behavior and by closely related metadata, and hybrid user profiles, which are the combination of the implicit and explicit ones and are classified as the outcome of social media platforms.

Accordingly, several acquisitions and enrichment techniques for each of the profile types were examined regarding their advantages, disadvantages, and challenges. Furthermore, this put forward overview showed that a possible shift from explicit to implicit and especially to hybrid user profiles and acquisition techniques could be identified (cf. Section III). Currently, the

improvements in well-known and modern techniques of the latter two categories are mainly in the scope of research.

Finally, privacy challenges and legal frame conditions and requirements that apply to user profile acquisition were highlighted. From the current development of the legal frame conditions, it can be inferred that privacy and data protection are gaining more important nowadays.

For further research, two areas might be of interest: privacy preserving user profile acquisition techniques might close the gap between the demand of holistic user profiles and the users privacy concerns; and, as the role of machine learning is vastly increasing, the investigation of learning algorithms for automating user profile acquisition is of great interest and may provide useful insights for researchers and practitioners.

References

[1] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Boston, MA, USA: ACM Press, 1999, vol. 29, no. 23.

[2] H. Chen, V. C. Storey, and R. Chiang, "Business Intelligence And Analytics: From Big Data To Big Impact," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012.

[3] S. Gauch, M. Speretta, A. Chandramouli, and A. Micarelli, "User Profiles for Personalized Information Access," *The Adaptive Web*, vol. 4321, pp. 54–89, 2007.

[4] D. Godoy and A. Amandi, "User profiling in personal information agents: a survey," *The Knowledge Engineering Review*, vol. 20, no. 04, p. 329, 2005.

[5] M. Hildebrandt and S. Gutwirth, *Profiling the European Citizen - Perspectives, Cross-disciplinary*. Springer Netherlands, 2008.

[6] P. Brusilovsky, A. Kobsa, and W. Nejdl, *The Adaptive Web: Methods and Strategies of Web Personalization*, 2007, vol. 4321.

[7] M. Hosseini and M. Shabani, "New approach to customer segmentation based on changes in customer value," *Journal of Marketing Analytics*, vol. 3, no. 3, pp. 110–121, 2015. [Online]. Available: http://www.palgrave-journals.com/jma/journal/v3/n3/abs/jma201510a.html

[8] M. Gao, K. Liu, and Z. Wu, "Personalisation in web computing and informatics: Theories, techniques, applications, and future research," *Information Systems Frontiers*, vol. 12, no. 5, pp. 607–629, 2010.

[9] M. Eirinaki and M. Vazirgiannis, "Web mining for web personalization," *ACM Transactions on Internet Technology*, vol. 3, no. 1, pp. 1–27, 2003.

[10] Simon Child, Kirill Ulitin, Gabriele Malaspina, artworkbean, Matt Caisley, Alex Tai, Icon Island, Gregor renar, Josh Sorosky, Antar, iconsphere, Hector, Rflor, Oliver Kittle, Arthur ShlainGrassegger, Hannes and Krogerus, Mikael, Arthur Shlain, Josh Sorosky, iconsphere, Icon Island, Alex Tai, Simon Child, Matt Caisley, Hector, Rflor, Nancy, Thomas Helbig, "Authors of the used icons - in whole or in a modified form - of Figure 1 and Figure 3. Each author can be found on the Noun Project website." 2016. [Online]. Available: https://thenounproject.com/

[11] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," *AI magazine*, pp. 37–54, 1996.

[12] S. Kanoje, S. Girase, and D. Mukhopadhyay, "User Profiling Trends, Techniques and Applications," *International Journal of Advance Foundation and Research in Computer*, vol. 1, no. 11, pp. 2348–4853, 2014.

[13] J. Tang, L. Yao, D. Zhang, and J. Zhang, "A Combination Approach to Web User Profiling," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 5, no. 1, p. 2, 2010.

[14] W. Fan, M. D. Gordon, and P. Pathak, "Effective profiling of consumer information retrieval needs: A unified framework and empirical comparison," *Decision Support Systems*, vol. 40, no. 2, pp. 213–233, 2005.

[15] D. Poo, B. Chng, and J.-M. Goh, "A Hybrid Approach for User Profiling," *SciencesNew York*, vol. 4, no. C, pp. 103–111, 2003.

[16] A. Hawalah and M. Fasli, "Dynamic user profiles for web personalisation," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2547–2569, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2014.10.032

[17] C. Markou, *Reforming European Data Protection Law*, 2015, vol. 20. [Online]. Available: http://link.springer.com/10.1007/978-94-017-9385-8

---

[2]An overview of the users concerns can be found in Special Eurobarometer 431 [59]

[18] V. Grouès, Y. Naudet, and O. Kao, "Combining linguistic values and semantics to represent user preferences," *Proceedings - 2011 6th International Workshop on Semantic Media Adaptation and Personalization, SMAP 2011*, pp. 27–32, 2011.

[19] H. Buxel, "Holger Buxel Customer Profiles Rechtliche Aspekte der Erstellung von Customer Profiles im Inter net in Deutschland Profiles Holger Buxel," no. 160, pp. 3–12, 2002.

[20] N. Höchstötter, "Methoden der Erhebung von Nutzerdaten und ihre Anwendung in der Suchmaschinenforschung," *Handbuch Internet-Suchmaschinen: Nutzerorientierung in Wissenschaft und Praxis*, pp. 175–203, 2009.

[21] F. Ricci, K. Bontcheva, O. Conlan, and S. Lawless, "User Modeling, Adaptation and Personalization: 23rd International Conference, UMAP 2015 Dublin, Ireland, June 29 - July 3, 2015 Proceedings," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9146, pp. 302–314, 2015.

[22] A. Kobsa, J. Koenemann, and W. Pohl, "Personalised hypermedia presentation techniques for improving online customer relationships," *The Knowledge Engineering Review*, vol. 16, no. 02, p. 111, 2001.

[23] I. Peters, *Folksonomies - Indexing and Retrieval in Web 2.0 - Knowledge & Information*, G. De and Saur, Eds., 2009.

[24] W. Woerndl and B. Lamche, "User Interaction with Context-aware Recommender Systems on Smartphones," *Icom*, vol. 14, no. 1, pp. 19–28, 2015.

[25] H. Steck, "Evaluation of recommendations: Rating-Prediction and Ranking," *Proceedings of the 7th ACM conference on Recommender systems - RecSys '13*, pp. 213–220, 2013.

[26] X. Amatriain, J. Pujol, J. Pujol, N. Tintarev, N. Tintarev, N. Oliver, and N. Oliver, "Rate it Again: Increasing Recommendation Accuracy by User re-Rating," *Proceedings of the 2009 ACM conference on Recommender systems*, no. 4, pp. 173–180, 2009.

[27] J. W. Byers, M. Mitzenmacher, and G. Zervas, "The Groupon Effect on Yelp Ratings : A Root Cause Analysis," *Work*, vol. X, no. X, pp. 1–21, 2012. [Online]. Available: http://arxiv.org/abs/1202.2369

[28] B. Fu, J. Lin, L. Li, C. Faloutsos, J. Hong, and N. Sadeh, "Why people hate your app: Making sense of user feedback in a mobile app store," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2013, pp. 1276–1284.

[29] D. Pagano and W. Maalej, "User Feedback in the AppStore: An Empirical Study (submitted)," *RE '13: Proceedings of the 21st International Requirements Engineering Conference*, pp. 125–134, 2013.

[30] E. Frias-Martinez, M. Cebrian, J. M. Pascual, and N. Oliver, "Explicit vs. Implicit tagging for user modeling," *CEUR Workshop Proceedings*, vol. 478, 2009.

[31] F. Abel, "Contextualization, User Modeling and Personalization in the Social Web - From Social Tagging via Context to Cross-System User Modeling and Personalization -," Doktor der Naturwissenschaften, Gottfried Wilhelm Leibniz Universität Hannover, 2011.

[32] K. Burke, K. Hammond, and B. Young, "The FindMe Approach to Assistive Browsing," *IEEE Expert*, vol. 12, no. 4, pp. 32–40, 1997.

[33] D. Jannach, G. Kreutler, I. R. Systems, and M.-d. W. A. Development, "Rapid development of knowledge-based conversational recommender applications with advisor suite," *Journal of Web Engineering*, no. 1, pp. 1–30.

[34] P. Grasch, A. Felfernig, and F. Reinfrank, "ReComment: towards critiquing-based recommendation with speech interaction," *Proceedings of the 7th ACM conference on Recommender systems - RecSys '13*, pp. 157–164, 2013.

[35] L. Han, G. Chen, and M. Li, "A method for the acquisition of ontology-based user profiles," *Advances in Engineering Software*, vol. 65, pp. 132–137, 2013.

[36] Google, "Die User-ID-Funktion," p. 1, 2016. [Online]. Available: https://support.google.com/analytics/answer/3123662?hl=de

[37] G. Skouma and L. Léonard, "On-line Behavioral Tracking: What May Change After the Legal Reform on Personal Data Protection," in *Reforming European Data Protection Law*, 20th ed., S. Gutwirth et al. (eds.), Ed. pringer Science+Business Media Dordrecht 2015, 2015, ch. Chapter 2, p. 409.

[38] M. Trusov, L. Ma, and Z. Jamal, "Crumbs of the Cookie: User Profiling in Customer-Base Analysis and Behavioral Targeting," *Marketing Science*, vol. 35, no. 3, pp. 405–426, 2016.

[39] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild," *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS '14*, pp. 674–689, 2014.

[40] M. Schneider, M. Enzmann, and M. Stopczynski, "Web-Tracking-Report 2014 02/2014," Fraunhofer-Institut fur Sichere Informationstechnologie SIT, Darmstadt, Tech. Rep., 2014.

[41] T. Mauerer, "Web Privacy," *Network Architectures and Services*, no. September, pp. 25–32, 2015.

[42] P. Eckersley, "How unique is your web browser?" *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6205 LNCS, pp. 1–18, 2010.

[43] H. Buxel, K.-P. Wiedmann, T. Frenzel, and G. Walsh, *Konsumenten-verhalten im Internet - Konzepte - Erfahrungen - Methoden*, 1st ed., B. Rascher and J. Hinrichsen, Eds. Wiesbaden: Gabler Verlag, Springer Science+Business Media, 2004.

[44] S. Puglisi, D. Rebollo-Monedero, and J. Forné, "On Web User Tracking: How Third-Party Http Requests Track Users' Browsing Patterns for Personalised Advertising," 2016.

[45] V. Mavroudis, S. Hao, and Y. Fratantonio, "The Ultrasound Tracking Ecosystem," 2016.

[46] V. Mavroudis and F. Maggi, "Talking Behind Your Back Attacks & Countermeasures of Ultrasonic Cross-Device Tracking," 2016.

[47] S. Englehardt and A. Narayanan, "Online Tracking," *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, no. 1, pp. 1388–1401, 2016.

[48] W. Premchaiswadi and P. Poompuang, "Hybrid Profiling for Hybrid Multicriteria Recommendation Based on Implicit Multicriteria Information," *Applied Artificial Intelligence*, vol. 27, no. 3, pp. 213–234, 2013.

[49] I. Heimbach, J. Gottschlich, and O. Hinz, "The value of user's Facebook profile data for product recommendation generation," *Electronic Markets*, vol. 25, no. 2, pp. 125–138, 2015.

[50] J. Crump, "What are the police doing on Twitter? Social media, the police and the public," *Policy & Internet*, vol. 3, no. 4, p. 7, 2011.

[51] B. Rieder, "Studying Facebook via Data Extraction: The Netvizz Application," *Proceedings of WebSci '13, the 5th Annual ACM Web Science Conference*, pp. 346–355, 2013.

[52] B. Batrinca and P. C. Treleaven, "Social media analytics: a survey of techniques, tools and platforms," *AI and Society*, vol. 30, no. 1, pp. 89–116, 2014.

[53] M. a. Russell, *Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites*, 2011. [Online]. Available: http://books.google.com/books?id=S5TZeDRYWdMC&pgis=1

[54] D. Fett, R. Küsters, and G. Schmitz, "A Comprehensive Formal Security Analysis of OAuth 2.0," *Proceedings of the 2016 ACM Conference on Computer and Communications Security (CCS)*, vol. 1, pp. 1–12, 2016. [Online]. Available: http://arxiv.org/abs/1601.01229

[55] D. Freelon, "Social media data collection tools," p. 1, 2016. [Online]. Available: http://socialmediadata.wikidot.com/

[56] Federal-Trade-Commission, "Data brokers: A call for transparency and accountability," *Data Brokers and the Need for Transparency and Accountability*, no. May, pp. 1–101, 2014.

[57] P. Dixon and Rockefeller, "Testimony of Pam Dixon Executive Director, World Privacy Forum - Before the Senate Committee on Commerce, Science, and Transportation - What Information Do Data Brokers Have on Consumers, and How Do They Use It?" *World Privacy Forum*, vol. 1, no. 1, pp. 1–18, 2013.

[58] o. C. S. U.S. Senate Committee and Transportation, "A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes," Committee On Commerce, Science, And Transportation: Office Of Oversight And Investigations Majority Staff, Tech. Rep., 2013.

[59] European Union, "Special Eurobarometer 431 - Data Protection Report," European Commission, Directorate-General for Justice and Consumers and co-ordinated by the Directorate-General for Communication, Tech. Rep. 14, 2015.

[60] N. J. King and J. Forder, "Data analytics and consumer profiling: Finding appropriate privacy principles for discovered data," *Computer Law and Security Review*, vol. 3, 2016.

[61] O. Heuberger-Götsch, "Der Wert von Daten aus juristischer Sicht am Beispiel des Profiling," in *Datenschutz in Zeiten von Big Data*, hmd ed., D. Fasel and A. Meier, Eds. Zürich , Schweiz: Springer Fachmedien Wiesbaden 2016, 2016, ch. 5, pp. 83–105.