Incidental data: observation of privacy compromising data on social media platforms

Stefan Kutschera

Received: 23 August 2022 / Accepted: 14 October 2022 / Published online: 10 November 2022
© The Author(s) 2022

Abstract  Social media plays an important role for a vast majority in one’s internet life. Likewise, sharing, publishing, and posting content through social media has nearly become effortless. This unleashes new threats as unintentionally shared information may be used against oneself or loved ones. With open-source intelligence data and methods, we show how unindented published data can be revealed and further analyze possibilities that can potentially compromise one’s privacy. This is backed up by a popular view of interviewed experts from various fields of expertise. We were able to show that only 2 hours of manually fetching data are sufficient to unveil private, personal information that was not intended to be published by the person. Two distinctive methods are described with several approaches. From our results, we were able to describe a 14-step awareness guideline and proposed a change of the law within Austrian legislation. Our work has shown that awareness among persons on social media needs to be raised. Critical reflections on our work has revealed several ethical implications that have made countermeasures necessary; however, it can be assumed that criminals do not to these.

Keywords  Privacy · Security · Incidental data · Cyber-security · GDPR · OSINT · Social media
1 Introduction

1.1 Problem statement

As more and more people become connected to the internet, the obstacles to publishing one’s own created content become less complicated. Likewise, more individuals can consume the published content. The simplicity of a connected world has heavy implications for privacy, security, and safety. Not everyone who contributes content has the time or awareness to cross-check for any unwanted data included in the published social media posting. Besides the desired information that a posting should transfer, the unaware and unwanted data published may be used to gather even more information, for instance, previous owners, cost of recent renovations, property tax, telephone number, and date of birth, to name a few.

Unwanted data within a public social media posting may be used for targeted attacks, as criminals can hide behind their computers and gather information that would have needed careful physical observation in front of a victim’s location. Such possible attacks might include, but are not limited to, burglary, phishing, or identity theft.

1.2 Incidental data

The term incidental is used in the context of a medical examination; it means that something fortuitous has been found; thus the term “incidental findings”. However, incidental was also used to describe unanticipated observations during an on-premise information security audit. For example, an x-ray of a broken chest reveals a suspicious tumorous tissue within the lungs, or looking behind a server rack reveals a wet spot indicating a water leak. Taking these points into consideration “incidental data” in our work and beyond shall be known as unintentional hidden information in a public posting on the internet or on social media platforms. The following situation could be an example: a person wishes to show how red, healthy, and tasty the tomatoes are growing in their own garden with a video or a photo but unknowingly reveals the residential address of their home and also a great deal more private information.

1.3 Research questions

The following research questions will be answered in this context as part of the underlying research conducted as part of the author’s master’s thesis.

1. What is the legal aspect of unveiling personal information found in social media postings?
2. To what extent is the information that has been gathered of value to a person or an organisation with criminal intentions?
3. What methods can extract incidental data in a maximum of 2 hours using OSINT data?
1.4 Hypothesis

Postings that are made freely available on social media by a person frequently include incidental data that can be accessed and also used to harm their author.

1.4.1 Null-hypothesis

Out of an arbitrary small selection of public social media accounts, none include compromising private information incidental data.

1.5 Methodology

Open-source intelligence (OSINT) data and methods are used to manually gather and fetch data from various social media platforms. After allocating personal data that may be posted unintentionally, this data is used to extract further information from freely available information or information services. In order to have comparable results and not extensively fetch data from a single profile, the gathering process should not exceed 2 hours. This threshold is also a countermeasure to prevent allegations of stalking. The popular view from experts will be obtained from interviews and analyzed using qualitative content analysis. [4] The incidental data can subsequently be evaluated to assess its economic value from a criminal perspective. In addition, a fictitious scenario based on this work will reveal the legal aspect of fetching incidental data in the course of a case law survey.

1.5.1 Ethics and morality

Personal data is a precious asset in the today’s world, not only for individuals but also for companies, and in both cases, criminals can be at work using personal data in an attempt to harm people. [5] It became evident in the course the research that our work brings a considerable degree of responsibility with it and that an evaluation of the ethical aspect needs to be carefully considered. In this context, we decided to undergo an ethics self-assessment test. [6] The ethics self-assessment test consists of 20 main questions, where an answered “Yes” implicitly or directly indicates a particular risk. As Table 1 shows, our work needed to be answered in 9, thus 45% of all 20 main questions with “Yes”. Coupled with the questionnaire, measures that counteract certain specific risks are also mentioned. We introduced specific measures derived from the feedback of the ethics self-assessment test. For example, we limited the number of hours spent on fetching incidental data. Furthermore, in the

Table 1  This table gives an overview of how we answered the questions from the ethical self-assessment test and the overall ratio of Yes/No is. As in this case, the ratio is 9/11

| Question # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | ∑ 20 | ∑ 9 | ∑ 11 |
|------------|---|---|---|---|---|---|---|---|---|---|------|-----|-----|
| Questions  | 3 | 2 | 1 | 2 | 1 | 5 | 3 | 1 | 1 | 1 |     |     |     |
| Yes        | 0 | 1 | 0 | 2 | 0 | 2 | 2 | 1 | 1 | 0 |     |     |     |
| No         | 3 | 1 | 1 | 0 | 1 | 3 | 1 | 0 | 0 | 1 |     |     |     |
case of persons without a professional security and privacy background, we asked for permission to fetch incidental data from the social media profile in question.

2 Related work

2.1 Open-source intelligence – OSINT

The process by which intelligence is produced and collected from publicly available information and distributed to persons with specific intelligence requirements is specified as OSINT. OSINT can be executed by a single person, a loose collective of individuals, or groups. For example, an attorney or a civil engineer using OSINT would mean official documents would be in use, whereas for a civil foreign intelligence service, OSINT would be a collection of news broadcast from a foreign country. Moreover, creating reports with the information gathered is also covered within the term OSINT due to the risk of information becoming unavailable. [7–9] Further, OSINT is used among the supervisory boards¹ to gather information on future partners. OSINT also serves as a foundation for investigations during merging and acquisition transaction disputes on the integrity of the person of interest². Besides this, the decision of the Verfassungsgerichtshof³ (Austria) VfGH G 223/2016 shows that public, freely accessible data, such as that found during OSINT searches are per se not regulated by Austrian law § 1 (1) Datenschutzgesetz⁴ (DSG). However, this changes as soon as authorities collect and store such data, as according to VfGH G 223/2016 such data processing falls under the protection Art.8 ECHR the right to respect for private and family life.

Nevertheless, according to [9], when OSINT information is collected, a reproducible procedure cannot be drawn in general. Despite this fact, a generalisation of specific steps can be made as shown in Table 2.

2.2 Value of personal data

OSINT information by definition adds a certain value to the investigation conducted. In addition to this, personal, financial, or health data may be used to gain an advantage, to profit, or to harm others by an intruder. [5, 9] Moreover, a simple identity mistake can have an enormous impact in one’s life, as happened in the case of David Quintavalle, who was later proved innocent by the Federal Bureau of Investigation (FBI). [10, 15:19min]

As made clear by Heimo Flechl, BA MA, Head of the Unit of OSINT & Crime Trends within the Criminal Intelligence Service Austria, gathered incidental data is worth very little when offered for sale on the dark web. Nevertheless, this situation can change when a method such as “Crime as a Service” is used or requested. “Crime

---

¹ See Wolfgang Helpa, Risikosteuerung und Informationsmanagment, Aufsichtsrataktuell 2020 H 6, 28.
² Jan Odewald/Alexandra Hullot, Dispute Resolution im Post-Merger-Umfeld, CB 2020, 242.
³ Translation [GER] to [ENG]: “Verfassungsgerichtshof” – “Austrian Constitutional Court”.
⁴ Translation [GER] to [ENG]: “Datenschutzgesetz” – “Data Protection Law”.
| # | Step                          | Description                                                                 |
|---|-------------------------------|-----------------------------------------------------------------------------|
| 1 | Triage                        | Creation of mission overview. Definition of expectations, questions.          |
| 2 | Preparation of tools          | Programs, tools, and services that are needed to support and answer the Triage step are made ready for usage. |
| 3 | Closed-source data queries    | (Optional) Databases locked through either payment or governmental access are queried against. |
| 4 | OSINT: Query all known identifiers | Necessary and available databases, sources and techniques are executed. |
| 5 | Collection                    | The leads serve as an input for the collection or as new evidence that can be used again in step 4. |
| 6 | Analysis                      | All gathered data is used to depict “bigger picture” and provide a good overview |
| 7 | Reporting                     | All gathered data is used to answer definitions from Triage within a final report. |
| 8 | Cleanup/archiving             | Virtual machines are either reverted and archived to provide an unbiased starting point. |

as a Service” may, in this particular case, describe the information gathering process on a contractual basis. In such a case, the value of this illicitly gathered information can range in value from under 100 EUR to several thousand EUR. Another aspect could be not to sell but to exploit the information to harm a person in numerous ways or occasions, such as causing repair costs after a forceful entry, bringing about loss of reputation or the loss of a job. Whatever the case, however, cash is still in favour among criminals. [3, 11–13]

### 2.2.1 Public versus social watchdogs

A phenomenon of the digital revolution was described in an analysis of case law on Art. 10 ECHR (freedom of expression). It was found during an analysis of case law on Art. 10 ECHR (freedom of expression) that certain professional groups enjoying even greater protection or freedom are divided into two groups, so-called “public watchdogs” and newly found so-called “social watchdogs”. Whereas public watchdogs are journalism in the classic sense, made available to the general public in print or digital broadcast media, social watchdogs are citizens without a professional journalistic background. They must have a similar broadcast range compared to their shared content and opinion on social media with public watchdogs. However, current legislation grants public watchdogs more freedom. Furthermore, the standard of due diligence is no higher for social watchdogs than it is for their professional counterparts, the public watchdogs. [14]

### 2.3 Human Rights Convention

Since human rights are the most fundamental rights any person can have, their inviolability would appear to be an obvious requirement. However, those rights are
not absolute but tend to interfere with each other in such a way that clear borders cannot be drawn. This interference is especially true for Art.8 ECHR the right to respect for private and family life and Art.10 ECHR the freedom of expression, shown as follows.

With a look at case law, it becomes clear that even the private life of publicly well-known people is protected. The case Alkaya v. Turkey (42811/06 ECHR), and its ruling supports this, as a well-known film and theatre actress from Turkey had a home robbery while at home, and a national newspaper later published pictures and information on her whereabouts. Further, the ruling of Satakunnan Markkipäörssi Oy and Satamedia v. Finland (931/13 ECHR), the ECHR, stated in the court assessment “134. The fact that information is already in the public domain will not necessarily remove the protection of Article 8 of the Convention.” In contrast, the ruling of Cengiz and others v. Turkey shows that social media platforms are indispensable tools for exercising the right of freedom of expression. The case of Von Hannover v. Germany clearly hints that Art.10 ECHR and Art.8 ECHR have to be treated with equal respect. In the matter that several photographs were taken and published in a newspaper, the case did not withstand the balance between Art.10 ECHR, the freedom of expression and Art.8 ECHR, the right to respect for private and family life.

3 Implementation

3.1 Manual gathering example #1

In the video “Wild Oklahoma Weather” from the YouTube channel “LiveEachDay”, a backyard scene reveals information on the surroundings of a private home as, for instance, a specific U-shape of the building with a US flag in the middle of the backyard. Albeit, it would be a lengthy process to search for the whole state of Oklahoma, as the title indicates. Notably, the scene where a map from the rain radar within a weather app is shown to the camera as depicted in Fig. 1c, reveals the GPS position.

With this information, the search radius can be drastically reduced as shown in Fig. 2a–c using a zoomed, rotated, and cropped image as shown in Fig. 2a. Figure 2a is eventually used to morph and match the corresponding map from Google Maps as shown in Fig. 2c).
Fig. 2 Depicts the morphing process of a screenshot with a map provider as a second data source. a Zoomed and rotated snippet of the map and GPS marker visible on the smartphone from screenshot visible in Fig. 1c [15]. b Showing the same city fragment of Tulsa, Oklahoma, USA on Google Maps as visible in the screenshot from Fig. 2a [16]. c Utilizing the distinct street arrangements a morphed images was created revealing a usable search area on the map [15, 16].

Since the search radius for the home was drastically reduced due to the GPS location provided, and eventually the address was discovered within 10 minutes using satellite images, as seen in Fig. 3. By using the address as an additional search input, it was possible to find information on the foreclosed previous owners. [17] In these foreclosure details, ultimately, the real estate parcel number “9941****290” allowed us to query the public register of Tulsa City, unveil information about the previous and current owners, with the implication of these being members of a close family in the listed current and previous owners with the inclusion of tax information, home improvements, and a list of further documents and pictures, as

Fig. 3 Screenshot from Bing Maps of the search result for a physical location after analysing satellite images for features as shown in Fig. 1a,b. [18]
Fig. 4  Diagram of prices from June 2016 to June 2017, where the owner of the YouTube channel LiveEachDay allegedly bought the property. [19, 20]

Fig. 5  Excerpt from the public register of Tulsa City showing sales on June 2017 to the owner of the YouTube channel LiveEachDay. [21]
Fig. 6 Several posts of Mr. Hunt depicting images from the alleged front respectively backyard of his home. a Car in front of alleged home [22], b Alleged backyard showing the sea view [23], c Alleged backyard showing a pool [24]

seen in Figs. 4–5. The evidence proved consistent, as the information from the public register of Tulsa City matched the foreclosure information.

3.2 Manual gathering example #2

In the following example, posts on the Twitter profile of Mr. Troy Hunt, a public figure known as an Australian web security consultant, outreach on security topics, and creator of “Have I Been Pwned”, are observed for indications on incidental data.

In a tweet, Mr. Hunt was asked if he still owns his Mercedes car; he replied with an image showing the car in front of a house, as shown in Fig. 6a. With the implied assumption that the house that can be seen here might well be his home, an OSINT search, including posted images from his backyard, was used; see Fig. 6a–c. Looking for the striking skyscraper pattern in the background as visible in Fig. 6c using the Google Maps 3D-view, an address was found within 25 minutes. Moreover, by utilizing the address, it was possible to gather information on alleged relatives, property price, size, date of purchase, and telephone numbers of Mr. Hunt and his alleged relative as visible but redacted in Table 3. Further, the telephone number

| Table 3 | Redacted information collected from various public, freely available sources |
|---------|--------------------------------------------------------------------------------|
| Overview| Troy Hunt                                                                        |
| Public channel | Troy Hunt                                      |
| Full name      | Troy Hunt                                      |
| Birthday       | —                                              |
| Address        | 5**** A**** D****, 4217, Surfers Paradise, Queensland, AUSTRALIA               |
| Relative       | K**** Hunt                                     |
| Additional information | Property Lot **** on RP****, 721 sqm          |
| Additional information | Property bought on 23rd March 2018 for **** AUD |
| Additional information | Phone (Troy Hunt): +61****76                  |
| Additional information | Phone (K**** Hunt): +61****88                |
was also linked to messenger apps such as Signal and Telegram, where the profile picture depicting Mr. Hunt was visible.

The most surprising fact was the announcement of an upcoming adventure respectively vacation with a detailed itinerary of around 9000 km and later updates from various locations during his ongoing trip, as seen in Fig. 7a–d. In Fig. 7b, a detailed itinerary shows feature destination ahead of time. Evidence implicitly confirms his physical presence later in the Yulara area within the Northern Territory.
At first glance, Fig. 7a is inconsistent with respect to screenshot time and publishing time, as the screenshot time is clearly ahead of the posting time. However, it turned out that posting times on Twitter match the time zone of the viewer. When the time zone is manually changed to UTC+9:30, matching the Northern Territory time zone, the posting time was 18:07, thus 1 minute after the posted screenshot on Twitter was
taken; see Fig. 7a. The evidence found strongly hints that the adventure trip was, indeed, posted ahead of time and with current updates of different locations.

### 3.3 Off-the-grid

We made a relatively quick setup to show how quick and easy it is to hide one’s identity and go, so-to-speak, off-the-grid. Assuming that the internet service provider (ISP) has the technical ability to track our online behaviour and legal obligation to hand over information to authorities, and further, the fact that the onion-routing (TOR) browser has publicly known entry and exit nodes, the ISP shall not be able to track that we are entering into a TOR node. In order to achieve such a disguise, we decided to create local layers, as seen in Fig. 10. We eliminated potentially identifying pieces of information lurking on the device used by removing the internal storage disk and booting the operating system “Tails” [30] entirely from a USB-live stick. Before accessing any internet services, a virtual private network (VPN) was installed on the router, allowing us to hide our traffic from the ISP. In the final step, we accessed the TOR network through our VPN connection, which allowed us to hide traffic from the VPN provider. Such setup is visually displayed in Fig. 9. In addition, Open-VPN software provides a functionality with the name “remote-random” where a new random server in an arbitrary country is selected during the boot-up of the router. In extreme cases, our approach can be further strengthened as free public internet access might be used where the VPN needs to be established on the mobile computer. However, in this case, surveillance cameras might be used to detect our identity and, hence, need to be strictly avoided. [3]
3.4 Interviews and statements

The following expert interviews were carried out in alignment with the presence on social media or the profession.

3.4.1 Interview, Scott Helme

Mr. Scott Helme, a security researcher, entrepreneur, and international speaker who provides training on hacking and encryption, was available to answer questions regarding findings from his Twitter and LinkedIn accounts. A post where the reflection on his car revealed details of his renovated home ultimately allowed findings such as his full home address, complete birthdate, detailed floor planning, and the price and date of the alleged purchase within a 40 min search. Mr. Helme confirmed the findings and was resigned to legislation in the UK that required such information to be made public. He is nonetheless well aware that such information can be used to plan a burglary, to impersonate him, or to produce personalised fraud letters sent via mail directly to his home. He also said the effort to hide or remove his data appeared to him as unbearable. [3]

3.4.2 Interview, Dr. Ries Bouwman

International entrepreneur and co-founder of “Omi’s Apfelstrudel” was asked for his view on privacy for SMEs. He was aware of the information that had been found and was happy that it was easily available to the public, as none of the data represents a threat but rather an opportunity for business partners to get in touch with him. Moreover, he uses public information to track whether he or his business partners previously marked a new business as untrustworthy. [3]

3.4.3 Interview, Dr. Vesna Krnjic

Dr. Vesna Krnjic, a researcher for privacy and security, at the Institute of Software Technology at Graz University of Technology, was not surprised by the information found after an OSINT search of 1 hour and 50 minutes among her social media profiles. However, what she found most interesting was the approach used. Dr. Krnjic fully supports the proposed change of law5, as she sees no need for SMEs without direct contact to customers to publicly disclose the company address when it matches the private home address of the owner. In addition, Dr. Krnjic believes the mobile phone signature offers the infrastructure needed for easily supporting an additional layer of security for such vulnerable information, as the current situation allows data gathering without registration or limitation. Ultimately, Dr. Krnjic points out that in today’s internet life, anyone and everyone might potentially reveal personal information about themselves without being aware of it. Regarding the lack of consciousness about these issues in the broad mass of internet users, she feels this is a little frightening when she reflects on the information that can be revealed.

---

5 See Sect. 5.4 [3].
by photos posted online, especially when considering the involvement of artificial intelligence or the use of unlimited time for analysis. [3]

3.4.4 Interview, Austrian SME entrepreneur

The Austrian SME Entrepreneur⁶, who owns several homepages and hosts her podcast, states that social media play a crucial part in her daily business. Travel enthusiasts but also professional scuba divers appreciate her expertise the most. The data found was already known to be publicly available; however, the approach used was new to her. For her, the obligation within Austria’s legislation to publish an imprint is most polarizing, as it is helpful but also concerning at the same time. Her measures against criminals or attacks include but are not limited to: never posting pictures of a destination that she is currently reviewing; posts are published according to a different time zone to prevent information leakage on observed posting times; legally binding her business partners not to release any statements or information revealing her current or future locations. [3]

3.4.5 Interview, Henry from Techlore

Henry, the owner of Techlore, an organisation in the United States that provides content and consultation on privacy and security, states that physical danger is the most likely threat behind publicly available data, which can be compensated with local security protections. He has mixed feelings regarding the obligation of publicly available company information or another protection layer. In Henry’s vision, education and the availability of alternative possibilities is a solution to the problem. Hence, make the general public and business owners aware of security and privacy threats and, in addition, giving easy possibilities to implement proper measures such as creating businesses outside one’s home address. [3]

3.4.6 Interview, Heimo Flechl

Mr. Flechl states that OSINT cannot be declared as criminal activity per se but that it certainly will be used for further criminal activities. The monetary value varies from subject to subject and may have an additional impact on integrity or reputation. However, damage can also occur on a physical or psychological level. With this in mind, Mr. Flechl reminds us that arbitrarily collected data may not be of value on the black market but may well be crime as a service. For example, crime as a service could be in the form that one criminal creates and sells a Trojan that collects banking data from victims and maybe infects contacts stored on the device of the victim. The collected data is then bought and used by another criminal, who exploits the information and with the possible involvement of other criminals who are able to perform money laundering. [3]

For Mr. Flechl, the imprint of companies is one piece of information, but it is certainly not enough to gain the trust of employees in order to perform a successful

⁶ Due to privacy concerns, it was decided to redact personal information.
fraud. A question that must be asked is whether the information that is given on the company page is necessary for fulfilling the legal or business requirements and to gain the trust of customers. This is also the case for individual persons who are sharing information on various social media platforms, when a closer look at the primary subject of the posting might reveal minor yet still visible and revealing details, such as street names, house numbers, famous buildings, prominent landmarks, and security measures, but also even the absence of such details. [3]

3.4.7 Conclusion of the interviews in Sects. 3.4.1–3.4.6

Summarizing of the interview contents has emphasised that privacy is a polarizing topic. On the one hand, entrepreneurs use publicly available business information to verify potential partners’ integrity and decide whether to conduct business with them, as the interviews with Dr. Bouwmann and the Austrian SME Entrepreneur have shown. On the other hand, the expertise of Mr. Flechl shows how criminals can misuse this information.

4 Evaluation

4.1 K-index table

We found no method or formula that satisfied our requirement for making incidental data comparable and to correlate it with one’s broadcast range. This is the ultimate reason behind our approach of the K-index. Stated simply, the K-index is a weighted sum of different categories on incidental data where in one case the value is capped. At the same time, the Relative K-index is set in relation to the broadcast range of the person. However, we had issues implementing caped values, as the PostgreSQL database used limited us in using only mathematical operations and no if statements within a query, which led to the adapted limitation equation shown in Eq. 1.

\[
k_{\text{index}} = \frac{2 \cdot 1}{1 + e^{-20 \cdot (\text{k.addresses})}} - 1 \pmod{2} + \sum_{i=1}^{n} k_i \cdot x_i,
\]

\[
\text{rel.k.index} = \frac{\text{broadcast.range} \cdot 0.000001}{100} \cdot k_{\text{index}}.
\]

| k-array          | weight |
|-----------------|--------|
| k.name          | 3      |
| k.birthdate     | 4      |
| k.relatives     | 0.5    |
| k.add.info      | 0      |
| k.sensitive.info| 5      |
4.2 Mock-up case Austria–Austria

The gathering and extraction of incidental data from social media platforms clearly raises legal questions. With this in mind, we attempted to summarise the most relevant facts and created a fictional case involving two Austrian citizens, where the plaintiff is the owner of social media accounts from which incidental data has been extracted and to show what legal aspects need to be taken into consideration in such a case.

4.2.1 Facts

In Table 4 the plaintiff’s social media accounts are described.

The plaintiff owns several social media accounts and has developed his business from a simple hobby to a legal entity that was established in 2014. Due to financial limitations and non-existing physical customer contacts, the plaintiff decided to use a separate room in his own home for his business. The plaintiff uploaded a 15:20 minute video with the title “My beautiful tomatoes are growing so well” onto the video platform YouTube in January 2019. The video had 12,082,017 views. The garden in the back of the plaintiff’s home is intentionally visualised in the video. However, unintentionally, the video shows two sides of the plaintiff’s house, including its colour and structure, as well as a five-floor tall building and mountains with specific patterns in the vicinity. The plaintiff also mentions issues with stalkers in the past, which following a newspaper interview, led to only the given name being published, instead of the full name.

The defendant accessed the aforementioned video and created screenshots at 1:25, 5:55, and 13:12 min. The second attempt to geolocate the address of the plaintiff was successful after the name of the channel was typed into the Google search engine and ultimately hinted at by the auto-completion. The hinted full name a query against the Austrian Business Licence Information System (GISA\(^7\)) resulted in a hit that was further verified using VPL. The address was then for a cross-check entered into the public land register of Styria (GIS Cadastre Styria\(^8\)), where the property owner’s name matched the alleged full name of the plaintiff. As a result, the defendant

\(^7\) Translation [GER] to [ENG]: “Gewerbeinformationssystem Austria – GISA” – “Austrian Business Licence Information System”.

\(^8\) Translation [GER] to [ENG]: “Geo Informations System Kataster Steiermark” – “Geo Information System Cadastre Styria”.

 Springer
Table 4  List and properties of the accounts owned by the plaintiff

| Platform | Public account name | Media reach | Content       |
|----------|---------------------|-------------|---------------|
| YouTube  | MU account YT       | 950,000     | 450 videos    |
| Twitter  | MU account TW       | 60,000      | 1,800 tweets  |
| Facebook | MU account FB       | 60,000      | unknown       |
| Reddit   | MU account RD       | unknown     | 1,000 comments|

assumed that the information found on the address, full name, date of birth, and owned property of the plaintiff was accurate and cross-confirmed throughout several publicly available databases.

4.2.2 Legal obligations

**YouTube** In accordance with the terms of service\(^9\) (TOS) the uploading user is still responsible for the content. However, the viewing users have restrictions as stated in the TOC as follows “*The user is not allowed to access, reproduce, download, distribute, transmit, display, sell, license, modify, adapt, or otherwise use any portion of the service or content except in the manner permitted by the service.***” [31]

**Google Search Engine** The TOS of the Google search engine is not per se violated by hinting at a full name. However, if the search results are used in a violation of applicable law, TOS would inevitably be violated as well. However, if the defendant were to use the search results to violate applicable law, for instance by stalking or harassing the plaintiff, a violation of the Terms of Service would be inevitable. [32]

**Gewerbeinformationssystem Austria (GISA)** Austrian law § 365e Gewerbeordnung\(^10\) 1994 (GewO 1994), BGBl I 45/2018, grants the unrestricted access to GISA. However, information as specified in § 365a (2) Figs. 1–8 GewO 1994, among others, is the private home address and shall only be provided if the requesting person substantiates a legitimate interest in the information.

**Public land register (“GIS Cadastre”)** Public land registers are covered by Allgemeines Grundbuchsgesetz\(^11\) 1955 (‘GBG 1955’) BGBl I 81/2020. The access of the public land register is public as governed by § 7 GBG 1955. Nonetheless, a sudden deactivation of the public land register service “GIS Cadastre” happened but was not covered by a recent change of legislation. A request for comment was issued to the Styrian Provincial Government and answered by Dipl. Ing. Dr. Rudolf Aschauer and Dipl. Ing. Harald Grießler, under case number “ABT17-25947/2014-401” as follows: “[...] Due to the data protection regulation and increased inquiries, the constitutional service of the province re-evaluated the question of access to land register data via the GIS Styria. The land register access continues to be available to the public on a limited basis via the other known platforms. [...]”. [3]

**Regulation EU 2016/679 GDPR & DSG** As the defendant, as well as the plaintiff, are citizens of Austria, the Regulation EU 2016/679 General Data Protection

\(^9\) Note that the legally binding version of 22nd of July 2019 is written in German.

\(^10\) Translation [GER] to [ENG]: “Gewerbeordnung” – “Trade Code”.

\(^11\) Translation [GER] to [ENG]: “Allgemeines Grundbuchsgesetz” – “Land Register Law”.

\[ Springer \]
Regulation (GDPR) is inter alia Art. 3 applicable. Terms, such as *personal data* and *processing*, used within Regulation EU 2016/679 GDPR are defined within Art. 4. *Personal data* relates to any identifiable or identified information of a natural person. Whereas operations performed either manually or automatically within a set of operations or a single operation on personal data is defined as *processing*. Such processing operations are defined by Art. 4 Sect. 2 of Regulation EU 2016/679 GDPR, as the collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure, or destruction of personal data. [33] Art. 15 Sect. 1 Regulation EU 2016/679 GDPR grants a data subject the right to ask for confirmation and access to personal data in possession of the data controller from the data controller. Inter alia, Art. 2 Sect. 2 Lit. c Regulation EU 2016/679 GDPR, grants an exception in case of “a natural person in the course of a purely personal or household activity”. [33] Moreover, based on Regulation EU 2016/679 GDPR, the ruling of Oberster Gerichtshof[12] (Austria) OGH 6 Ob 150/19f refers in Sect. 5 of the reasoning inter alia to the decision of European Court of Justice (ECJ) C-212/13, where the latter is based on the predecessor of Regulation EU 2016/679 GDPR, namely Directive 95/46/EC. As a matter of fact, the ruling of OGH 6 Ob 150/19f, refers in Sect. 5 of the reasoning inter alia to the decision ECJ C-212/13, where the principle of household activity cannot be claimed if the personal data gathered is also used in order to protect property. This jurisdiction makes clear that household activity has a restricted meaning. In view of the aforementioned, two conclusions can be drawn. Firstly, ECJ rulings based on Directive 95/46/EC can still be used to help interpret the successor Regulation EU 2016/679 GDPR. Secondly, a natural person who does not solely process data based on private and household activity, must abide by Regulation EU 2016/679 GDPR, private and household activity, however, needs to be viewed as limited, as property protection is not included. This ultimately implies that the defendant (Sect. 4.2.1) needs to be viewed as a controller and eventually shall grant rights and protect personal data to the data subject as per Regulation EU 2016/679 GDPR. At the same time, as discussed in Sect. 2.1, the decision of the VfGH G 223/2016 hints that public freely accessible data are per se is not regulated by Austrian law § 1 (1) (DSG).

5 Results

5.1 Incidental data

It is evident that publicly posted information may be legally accessed and processed with a legitimate interest by persons within the OSINT field, such as executives, police, and detectives, as can be seen in Sect. 2.1. Moreover, Regulation EU 2016/679 GDPR states that information that was posted by a person itself has lost certain data protection rights. Not to mention that Austrian law § 1 (1) DSG does not regulate publicly available data, as discussed in Sect. 4.2.2. Therefore, a person who posts

[12] Translation [GER] to [ENG]: “Oberster Gerichtshof” – “Supreme Court” –.
his or her information have a lower protection value in terms of data protection. This is further supported by VfGH G 223/2016, as this decision only sees publicly available data protected under Art.8 ECHR the right to respect for private and family life in the case that the data processor is an authority. As shown in Sects. 3.1, 3.2, and 3.4.3, we were able to find incidental data among a small number of social media accounts. Given this fact, we need to reject our null hypothesis from Sect. 1.4 and confirm the hypothesis that social media postings may include incidental data which, in addition, may have the potential to cause harm. Incidental data was found among entrepreneurs, security experts, politicians, and other individuals with various specialisations. Among the incidental data, information such as private addresses, detailed floor plans, price of the property and/or buildings or flats, security measures, and information on alleged relatives was found. In the case where we found incidental data about the former Chancellor of Austria, who was in office at the time of the analysis, the Austrian Federal Chancellery was constructively involved at an early stage of our findings. Further, we found victims of stalking who remained on social media platforms and showed more awareness towards privacy. Sadly, it was still possible to gather incidental data from content posted after the incidents, such as the private address, detailed information on the main home and other properties, as well as information on relatives. Equally important is the role of public and social watchdogs. However, case law has shown that individuals well known to the public must allow private information to be (re)published under certain circumstances. In this light, incidental data gathered by an online security expert may not withstand the balancing between Art.8 ECHR and Art.10 ECHR.
5.2 Methods for incidental data

We found two methods most useful. Firstly, the identification of a physical location (IPL), where the address of a person is not known but due to an OSINT investigation found with a high level of certainty. The IPL method is visualised in Fig. 13. Secondly, verification of a physical location (VPL) where the address might be known but not confirmed, or the information is outdated. The VPL method is used to verify the address of a person with high certainty, as visualised in Fig. 12.

5.3 Summary of the awareness guideline

Given our observation and expert interviews, we concluded with the following best practice measures to improve one’s online security and privacy behaviour. We acknowledge that some measures are not suitable for every person and want to emphasise that the measures might also strengthen the privacy of others if later published content was taken from the property of someone else.

1. Avoid posting content that includes house numbers or street names.
2. Be on the lookout for reflections in mirrors as well as on surfaces such as cars, windows, vitrines, glasses, sunglasses, or watches.
3. Post content of vacations – if at all – only after the vacation has ended.
4. Avoid repetition of vacations or periods of absence, such as “during New Years Eve I am – always – on a one-week trip”.
5. Posts should be in accordance with a single time zone irrespective of a current and temporary time zone.
6. Avoid posting any information from parcels or letters, such as tracking number, full address, names, or QR codes.
7. Do not post IDs such as driver’s license, personal ID, credit or debit card, etcetera.
8. When posting letters, make sure that the address or sensitive information on the back of the letterhead does not show through.
9. Avoid posting scenes that include location-based map materials, such as navigation maps, weather or fitness apps, et cetera.
10. Close all curtains or post content where no windows are visible.
11. Try tilting the camera angle as low as possible when showing your own property.
12. Be aware that shadows or the sun’s position can also hint at additional information about the location.
13. Do not share fitness routes that start or end at your home location.
14. Do not share information about your own or surrounding WLAN/WiFi SSIDs

5.4 Proposed change of law

With a look at recent changes in the accessibility of the public land register as discussed in Sect. 4.2.2 it becomes evident that the term “public access” may also be associated with small payments and/or personalised logins. For instance, in order to get access to an entry within the public land register one must, besides paying a fee, ultimately reveal one’s name in any form to either an agency or provider thereof. With this in mind, the implementation of a change of law as stated below would increase the right to respect for private and family life (Art. 8 ECHR) of entrepreneurs and their families without infringing against the rights of other parties. This protection can be archived by accessing the currently unrestrictedly accessible information by restrictions imposed through the qualified electronic signature as with Regulation EU 910/2014 (e-IDAS). Moreover, this would ensure that a private address will have a higher standard of protection as indicated by § 365a (2) 4 GewO 1994. With the restriction of a login through the qualified electronic signature applied, the information can still be seen as publicly available.

“Die Gewerbeordnung 1994 – GewO 1994, BGBl. Nr. 194/1994, zuletzt geändert durch das Bundesgesetz BGBl. I Nr. 65/2020, möge wie folgt geändert werden:
Dem §365e wird folgender Abs.6 angefügt:
(6) Ist die Wohnanschrift (§ 365a Abs. 2 Z 4) dieselbe wie der Standort der Gewerbeberechtigung (§ 365b Abs. 1 Z 3), hat die Behörde den Standort der Gewerbeberechtigung gleich wie seine Wohnanschrift (§ 365a Abs. 2 Z4) zu behandeln und nur darüber Auskunft zu erteilen, wenn der Auskunftserwerber ein berechtigtes Interesse an der Auskunft glaubhaft macht.

1. Der Gewerbetreibende hat jederzeit die Möglichkeit der Behörde formlos mitzuteilen, den Standort seiner Gewerbeberechtigung (§ 365b Abs. 1 Z 3) jedermann, auch ohne Glaubhaftmachung von berechtigtem Interesse, zu auskunfts- und
2. Die Mitteilung über §365e Abs. 6 Z 1 an die Behörde kann mündlich, telefonisch, telegrafisch, schriftlich, fernschriftlich, mit Telefax, im Wege automationsunter-

13 Since the Austrian legislation is in force in its entirety only in the German language, we decided to present the proposed change of law also in German. This decision was driven by the desire to make an implementation of the proposed change as effortless as possible.
The English translation of the proposed change of law reads as:

“The Gewerbeordnung 1994 – GewO 1994, Federal Law Gazette No. 194/1994, as last amended by the Federal Law Gazette I 65/2020, shall be amended as follows: The following Paragraph 6 shall be added to § 365e:

(6) If the residential address (Sect. 365a(2)4) is the same as the location of the business license (Sect. 365b(1)(3)), the authority shall treat the location of the business license in the same way as its residential address (Sect. 365a(2)4) and shall only provide information about it if the person requesting the information substantiates a legitimate interest in the information.

1. The trader has, at any time, the possibility of informing the authority informally, the location of its trade license (§ 365b (1) 3) to provide information to anyone, even without credible justified interest.

2. The notification of § 365e (6) 1 to the authority may be made orally, by telephone, telegraph, in writing, by telex, by fax, by means of computer-assisted data transmission or in any other technically feasible manner.”

6 Conclusion and outlook

6.1 Conclusion

All things considered, incidental data can be found among people with different personal and professional backgrounds. Interestingly, expert interviews have shown that publicly available information is a polarizing topic among entrepreneurs, as it can be an ominous threat to one’s private life, or by contrast it may enhance and secure one’s own business. Controversial is the finding of the pre-announced and updated trip of several thousand kilometres, as an online security expert should lead as a role model.

6.2 Outlook

Despite our tendency to consider incidental data as an awareness problem that can be solved by education, the implementation of a not-yet-existing tool that is able to detect and classify critical content in a posting, as the awareness guide (Sect. 5.3) points out, and warns a user before the post is made public, would be of substantial value. At the same time, the future development of such tools also has the potential to be used against people and, consequently, raise ethical issues. In any case, the first steps that might solve problems surrounding incidental data might eventually be solved by interdisciplinary measures such as education, awareness programs, and, furthermore, be strengthened by legislation, as we proposed in our presented change of law [3].
Acknowledgements  Special thanks are due to Michael Brickmann, Dr. Sabine Proßnegg, and further also Dr. Bernd Pichlmayer for their helpful input and for many hours of fruitful discussions.

Funding  Open access funding provided by Graz University of Technology.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Wolf SM, Lawrenz FP, Nelson CA, Kahn JP, Cho MK, Clayton EW, Fletcher JG, Georgieff MK, Hammerschmidt D, Hudson K, Illés J, Kapur V, Keane MA, Koenig BA, LeRoy BS, McFarland EG, Paradise J, Parker LS, Terry SF, Van Ness B, Wilford BS (2008) Managing Incidental Findings in Human Subjects Research: Analysis and Recommendations. J Law Med Ethics 36(2):219–248. https://doi.org/10.1111/j.1748-720X.2008.00266.x
2. Casanovas P, Morris N, González-Conejero J, Teodoro E, Adderley R (2018) Minimisation of Incidental Findings, and Residual Risks for Security Compliance: the SPIRIT Project. In TERECONNURIX (Vol-2309), pp 97–110. ISSN 1613-0073. Online: http://ceur-ws.org/Vol-2309/09.pdf. Accessed 15 Aug 2022
3. Kutschera S (2021) Incidental Data: Detect and Process Personal Information from Social Media Platforms. FH JOANNEUM GmbH, Kapfenberg, Austria
4. Mayring P (2010) Qualitative Inhaltsanalyse: Grundlagen und Techniken. Pädagogik, vol 11. Beltz, Weinheim, Basel
5. Ablon L (2018) Data thieves: the motivations of cyber threat actors and their use and monetization of stolen data. RAND Corporation, Santa Monica, California https://doi.org/10.7249/CT49
6. European Union Funding for Research & Innovation, European Research Council (2018) European Research Council Executive Agency: Ethics Self-Assessment step by step Horizon 2020
7. Akhgar B (2016) OSINT as an Integral Part of the National Security Apparatu. Springer, Cham, pp 3–9 https://doi.org/10.1007/978-3-319-47671-1_1
8. Bayerl PS, Akhgar B (2015) Surveillance and falsification implications for open source intelligence investigations. Commun ACM 58(8):62–69. https://doi.org/10.1145/2699410
9. Bazzell M (2021) Open source intelligence techniques – resources for searching and analyzing online information (Independently published)
10. Bazzel M (2021) The privacy, security, & OSINT show, 214-offensedefense: the capitol siege. Soundcloud global limited & co. KG. https://soundcloud.com/user-98066669/214-offensedefense-the-capitol-siege. Accessed 3 May 2021
11. (2015) Polizei Hagen Nordrhein-Westfalen: Facebook. Meta Platforms Ireland Limited. www.facebook.com/Polizei.NRW.HA/photos/pb.208563659315946.-2207520000.1436266044./439893356 182974/?type=3&theater. Accessed 29 Jan 2021
12. Polizei Hessen PPNH mkr (2020) Polizei Hessen – Urlaub und soziale Netzwerke – Auch Einbrecher nutzen Facebook. https://www.polizei.hessen.de/praevention/sicher-in-den-urlaub/var-dem-urlaub/einbruch-und-diebstahlschutz/broker.jsp?uMen=2a370ae5-c345-7e41-dbd6-b0f60ef798e7&uCon=0b960058-53c8-ae41-8569-08820ef798e7&utmTem=bb71055-bb1d-50f1-5b60-308f1-2860-72700266cb59. Accessed 29 Jan 2021
13. Weber J, Krussbergen EW (2019) Criminal markets: the dark web, money laundering and counter-strategies – an overview of the 10th research conference on organized crime. Trends Organ Crime 22(3):346–356. https://doi.org/10.1007/s12117-019-09365-8
14. Holoubek M (2016) “Social watchdogs” nicht professionell organisierte Medienmacher im Lichte der Medienfreiheit. In: Berk, Holoubek, Leitl-Staudinger (eds) BürgerInnen im Web. REM, vol 14
15. (2018) LiveEachDay: Wild Oklahoma Weather. YouTube. https://youtu.be/FRdxfiZDXwE?t=317. Accessed 15 Dec 2019
16. Google Inc. (2019) Google Inc.: Map of Tulsa, Oklahoma. https://www.google.at/maps/place/Tulsa,+Oklahoma,+USA/@36.152105,-96.438444,9z. Accessed 29 Sep 2019
17. (2019) ForeclosureFreeSearch.com: Foreclosure.com | Foreclosures | Foreclosure Listings. BluSea Holdings LLC. https://www.foreclosure.com. Accessed 15 Dec 2019
18. Microsoft (2022) Microsoft Bing Maps. https://www.bing.com/maps. Accessed 22 June 2022
19. (2019) Trulia: Trulia: Real Estate Listings, Homes For Sale, Housing Data. Zillow, Inc. https://www.trulia.com/. Accessed 15 Dec 2019
20. Zillow LLC (2019) Zillow: Real Estate, Apartments, Mortgages & Home Values. Zillow, Inc. https://www.zillow.com. Accessed 15 Dec 2019
21. Wright JA (2015) Tulsa County Assessor. Tulsa County. https://www.assessor.tulsacounty.org/assessor-property-search.php. Accessed 15 Dec 2019
22. Hunt T (2020) Yep, and the E63S:. Twitter Inc. https://twitter.com/troyhunt/status/1246587017661538304. Accessed 16 Jan 2021
23. Hunt T (2020) So beautiful here today, but so quiet... Twitter Inc. https://twitter.com/troyhunt/status/1246587017661538304. Accessed 16 Jan 2021
24. Hunt T (2021) Book-ending the day moon swane. Twitter Inc. https://twitter.com/troyhunt/status/1246713081419288576. Accessed 16 Jan 2021
25. Hunt T (2021) Nope, walk only. Twitter Inc. https://twitter.com/troyhunt/status/1344925637513940994. Accessed 16 Jan 2021
26. Hunt T (2020) ROAD TRIP!!! Dolphin Spider Shark Crocodile Kangaroo Snake Koala. Twitter, Inc. https://twitter.com/troyhunt/status/1336432552177856513/photo/1. Accessed 16 Jan 2021
27. Hunt T (2020) Australian death frog. Twitter, Inc. https://twitter.com/troyhunt/status/1337695185514348544/photo/1. Accessed 16 Jan 2021
28. Hunt T (2021) Weekly Update 224. YouTube Inc. https://twitter.com/troyhunt/status/1344925637513940994. Accessed 16 Jan 2021
29. (2021) RP Data Pty Ltd: Onthehouse.com.au: Your Home for Property Search. onthehouse.com.au. https://www.onthehouse.com.au. Accessed 16 Jan 2021
30. (2019) Tails: Tor bridge mode. Tails. https://vpnpro.com/blog/nordvpn-security-breach-between-fact-and-fiction/. Accessed 1 Jun 2020
31. (2019) YouTube Inc.: Nutzungsbedingungen. https://www.youtube.com/t/terms/#d676f2b451 (Created 22 July 2019). Accessed 28 April 2020
32. (2020) YouTube Inc.: Google Terms of Service. Google Inc. https://policies.google.com/terms?fg=1#toc-what-we-expect. Accessed 28 April 2020
33. Unger K (2018) Grundzüge des Datenschutzrechts. NWV, Neuer Wissenschaftlicher Verlag, Wien, Graz
34. Schmid F (2019) Sicherheitspanne: ARD-Doku verrät Adresse von Kanzler Kurz. STANDARD Verlagsgesellschaft m.b.H. https://www.derstandard.at/story/2000097259454/sicherheitspanne-ard-doku-verraet-adresse-von-kanzler-kurz. Accessed 7 Oct 2019
35. (2021) Google Maps: Google Maps – Street View. Google Ireland Limited. https://www.google.com/streetview/. Accessed 2 May 2021
36. (2021) YouTube: YouTube. Google Ireland Limited. https://www.youtube.com. Accessed 2 May 2021

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.