COMUN. MÍDIA CONSUMO, SÃO PAULO, V. 21, N. 62, P. 558-580,SET./DEZ 2024
DOI 10.18568/CMC.V20I62.2873

Contextualizing solove's taxonomy of privacy in the data life cycle

Contextualizando a taxonomia da privacidade de solove no ciclo de vida dos dados

Ricardo César Gonçalves Sant'Ana
Dayane de Oliveira Martins

Abstract: *In the Data Life Cycle privacy is a factor that can permeate all phases. However, understanding the concept of privacy can be a complex task. Daniel Solove (2006) describes the so-called Taxonomy of Privacy, which addresses the complexity of privacy breaches. The aim of this study is to relate the scenario proposed in a fictional narrative to the context of privacy breaches. Through a process of segmentation, the episode «Joan is awful» from the series Black Mirror, was divided into 33 sequences to identify the circumstances in which the privacy of the main character of the episode (Joan, by Annie Murphy) was breached, in terms of the Taxonomy of Privacy, bringing greater concreteness to the issue of privacy. Out of the 16 subgroups proposed in the Taxonomy of Privacy, it was observed that the breach of Joan's privacy occurred in 5 distinct subgroups throughout the episode.*

Keywords: *Privacy; Data Life Cycle; Taxonomy of Privacy.*

Resumo: *No Ciclo de Vida dos Dados a privacidade é um fator que pode permear todas as fases. Contudo, compreender o conceito de privacidade pode ser uma tarefa complexa. Daniel Solove (2006) descreve a chamada Taxonomia da Privacidade, que aborda a complexidade da quebra de privacidade. O objetivo do presente trabalho é relacionar o cenário proposto em uma narrativa ficcional com o contexto de quebra de privacidade. A partir de um processo de decupagem, segmentou-se o episódio Joan is awful, da série Black Mirror, em 33 sequências, a fim de identificar em que circunstâncias a privacidade da personagem principal* *do episódio (Joan, interpretada por Annie Murphy) foi quebrada, nos termos da Taxonomia da Privacidade, trazendo maior concretude a questão da privacidade. Dos 16 subgrupos propostos na Taxonomia da Privacidade, percebeu-se que a quebra da privacidade da personagem Joan ocorreu em 5 subgrupos distintos ao longo do episódio.*

Palavras-chave: *Privacidade; Ciclo de vida dos Dados; Taxonomia da Privacidade.*

Introduction

Science fiction audiovisual works have been a relevant part of popular culture, offering imaginative visions of the future, advanced technologies, and exploration of complex scientific concepts (MENEZES E ARAÚJO, 2018). Information and Communication Technologies (ICT) is one of the themes addressed in audiovisual productions of the science fiction genre, as in the case of the series Black Mirror (RODRIGUES; SANT'ANA, 2019).

In both audiovisual works and reality, the creation of new methods and techniques resulting from the evolution of information technology and the growing prominence of ICT, the intensification of data generation, access, collection, and storage in ever quicker pace and larger volumes, configured the so-called Big Data phenomenon (SANT'ANA, 2016), allowing the identification of individuals through the analysis of large volumes of data, exposing privacy.

The issue of privacy has also been one of the themes portrayed in science fiction audiovisual works, for example, when characters use certain ICTs that lead to privacy breaches.

Daniel Solove, a researcher on the topic of privacy, asserts that "privacy is a very complicated concept to summarize in a single essence. Attempts to find such an essence often end up being very broad and vague, with little use for addressing concrete issues" (SOLOVE, 2006, p. 485). Attempting to explain (the breach of) privacy, the author develops the so-called Taxonomy of Privacy (SOLOVE, 2006), stating that a privacy breach does not occur based on a single criterion.

This work aims to relate the scenario proposed in a fictional narrative (audiovisual work) with the context of breach of privacy. We use the episode Joan Is Awful from the series Black Mirror to identify what circumstances compromise the private sphere of the episode's main character (Joan, Annie Murphy) according to the Taxonomy of Privacy, bringing greater concreteness to Solove's (2006) proposed framework.

The present study limits its sample to the episode Joan Is Awful from the series Black Mirror. We adopted the exploratory and descriptive analysis methodology used by Rodrigues and Sant'ana (2019) in a similar analysis in which the authors' observations segmented an episode into sequences. This decoupage process considers that one or more scenes interconnected by the continuity of the action in the audiovisual narrative compose a dramatic unit. Time and space can vary between scenes, but the sequence maintains a logical continuity (AUMONT; MARIE, 2007). In each of the sequences, we identified characteristics with a focus on the circumstances in which the main character's privacy may have been violated (yes), one cannot say that it was violated (no), or we disregarded it because there was no direct relationship with the character Joan (Annie Murphy) (not applicable) based on the groups and subgroups brought up in the Taxonomy of Privacy (SOLOVE, 2006), addressed in the third section.

The Data Life Cycle (DLC) is a cyclical structure composed of four phases: Collection, Storage, Recovery, and Disposal (SANT'ANA, 2019). In this structure, six transversal factors permeate all phases (privacy, integration, quality, copyright, dissemination, and preservation). Based on the groups Solove (2006) identifies, this research emphasizes the factor of privacy in the collection and recovery phases. We did not analyze the storage and disposal phases because these are phases that occur within the DLC holder's space.

The last section presents discussions and reflections on how users can experience privacy breaches in various circumstances, especially without knowledge or consent. In this sense, we hope to bring greater concreteness to the Taxonomy of Privacy (SOLOVE, 2006), even if based on the observation of a fictional work.

The data life cycle (DLC) and the privacy factor based on Solove's taxonomy

With the creation of new methods and techniques resulting from the evolution of computing, especially Information and Communication Technology (ICT), a transformation in data processing can be seen, which involves a quantitative and qualitative change. Thus, data processing occurs in less time (quantitative), and the results obtained are more accurate (qualitative) (DONEDA, 2021).

The data itself is purely objective, does not have a high intrinsic semantic load, and is independent of the user, but constitutes the raw material for a series of possible interpretations, as well as measures or facts represented by numbers, words, sounds, and even images that can support the production of new information (SOUZA; ALMEIDA, 2023). Thus, the present research understands the term data as

> a content unit necessarily related to a given context and composed of the triad entity, attribute, and value in such a way that, even if the details about the content's context are not explicit, they must be implicitly available to the user, thus allowing for their full interpretation (SANTOS; SANT'ANA, 2015, p. 205).

In this study, we adopted the concept of data as the fundamental element in information generation, consisting of the entity-attribute-value (EAV) triad. In these terms, the triad comprises a minimum set of symbols that can be taken as a content unit, requiring the identification of the context to which it belongs (SANTOS; SANT'ANA, 2015).

In this way, Information Science (IS) can contribute by seeking a balance between access and the intense use of personal data in certain contexts. Among the possibilities of contextualization is the delimitation of the phases and factors that permeate the structure of the DLC. Specifically, in this research, the privacy factor stands out when connecting with the subgroups of Solove's privacy taxonomy (2008).

Privacy as a transversal factor of the DLC

Privacy is one of the factors of the DLC, observable in all phases, and is a common concern in most global legislation (DONEDA, 2021). Constitutions, laws, and regulations seek to protect the privacy of their citizens. For example, the 1948 United Nations Universal Declaration of Human Rights states, "No one shall suffer arbitrary interference with his privacy, family, home or correspondence, nor attacks upon his honor and reputation". The Brazilian Constitution of 1988, in its article 5th, also guarantees that "X - the intimacy, private life, honor and image of individuals are inviolable, and the right to compensation for property or moral damages resulting from the violation thereof is ensured".

Professor Daniel Solove (2008) stated that when he began his studies on the subject, he looked for a definition of "privacy" but found no satisfactory concept when delving deeper into the issue.

In this sense, one wonders: why does the definition of privacy seem so common and, at the same time, so complex? According to the author, "Often, privacy problems are merely stated in knee-jerk form: "That violates my privacy!" (SOLOVE, 2008, p. 7). Thus, instinctively, one knows that certain situations can violate privacy. For example, when companies collect personal data without the holder's authorization or knowledge, it can be said that there has been a breach of privacy (SOLOVE, 2008, p. 7). But how can one technically present the concept of privacy? Regarding the diffuse nature of the concept of privacy, Solove says it seems to encompass everything and, therefore, seems to mean nothing (SOLOVE, 2008). Thus, the perception that "The term "privacy" is an umbrella term, referring to a wide and disparate group of related things. The use of such a broad term is helpful in some contexts yet quite unhelpful in others." (SOLOVE, 2006, p. 485). Several situations can represent a breach of privacy:

- A newspaper reports the name of a rape victim.
- Reporters deceitfully gain entry to a person's home and secretly photograph and record the person.
- New X-ray devices can see through people's clothing, amounting to what some call a "virtual strip-search."
- The government uses a thermal sensor device to detect heat patterns in a person's home.
- A company markets a list of five million elderly incontinent women.
- Despite promising not to sell its members' personal information to others, a company does so anyway." (SOLOVE, 2006, p. 481)

Warren and Brandeis authored the article "The Right to Privacy" (1890), warning new technologies like instant photography could invade the sacred precincts of private and domestic life when disclosed in the press, for example. Thus, the breach of privacy also began to be seen as intangible harm, expanding the notion of harm that had been only physical up to that point. The authors noted that the law and regulations should recognize non-physical harm to the same extent as they did physical harm.

In the case of privacy, according to the authors, it involves "injury to feelings". Privacy, therefore, is related to the protection granted to thoughts, feelings, and emotions expressed by any means and is one of the instances of application of the right to be alone, the right not to be disturbed, as defended in the United States for the first time by Justice Thomas Cooley of the Michigan Supreme Court (1888).

Alan Westin (1967), in the same sense, identified four basic states of individual privacy: (1) solitude, the individual is separated from the group and is free from observation or interaction with other people; (2) intimacy, the person has the option of choosing with whom he wants to relate in a reserved, intimate way; (3) anonymity, the individual expresses themselves publicly (through acts or other manifestation) but their identity remains hidden; and (4) reserve, the creation of a psychological barrier against unwanted intrusion.

For Westin (1967), privacy is related to the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them reaches others. Westin (1967) also expressed his concern about preserving privacy in the face of new surveillance technologies.

However, according to Solove (2006), the categories presented focus mainly on spatial distance and separation but fail to capture the different dimensions of informational privacy.

Despite the important considerations brought by the authors mentioned above, Solove (2006) states that Prosser, for example, wrote his considerations on privacy more than 40 years ago, and new technologies, especially ICT, have given rise to a set of new damages to privacy, making the construction of this concept even more complex.

Solove's taxonomy of privacy

In an attempt to understand privacy, Solove (2006) presents the Taxonomy of Privacy, shifting the focus away from a single definition for the term to look at activities that affect privacy. That is, he establishes a plural concept based on actions that can harm or violate an individual's privacy. Thus, Solove states that:

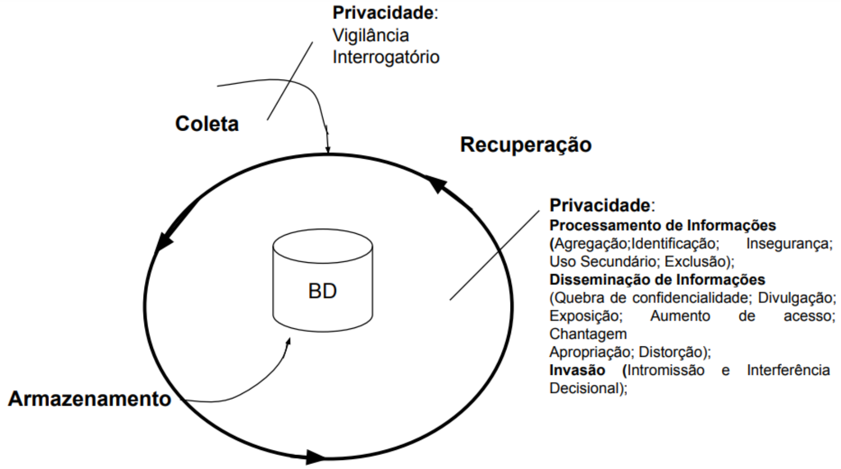
> In terms of generality, I argue that privacy should be conceptualized from the bottom up rather than the top down, from particular contexts rather than in the abstract. All conceptions must exist at some level of generality, however, so my theory generalizes beyond the myriad of specific contexts. (SOLOVE, 2008, p. 9)

Solove's (2006) taxonomy aims to simplify the understanding of situations that can violate users' privacy. Thus, the author argues that the focus should be on privacy issues rather than seeking to locate a single conceptual terrain.

The author's taxonomy of privacy classifies four basic groups of activities that violate privacy, namely: (1) information collection, (2) information processing, (3) information dissemination, and (4) invasion.

In Figure 1 below, we adapted the DLC (SANT'ANA, 2016), highlighting the transversal factor of privacy in the collection and recovery phases, as brought by Solove (2006), as follows:

Figure 1 – Privacy as a transversal factor in the data collection and recovery phases


Privacidade:
Vigilância
Interrogatório
Coleta
Recuperação
BD
Privacidade:
Processamento de Informações
(Agregação;Identificação; Insegurança; Uso Secundário; Exclusão);
Disseminação de Informações
(Quebra de confidencialidade; Divulgação; Exposição; Aumento de acesso; Chantagem
Apropriação; Distorção);
Invasão (Intromissão e Interferência Decisional);
Armazenamento

Source: Adapted from SANT'ANA, 2016.

According to Solove (2006), the breach of privacy can begin as early as the **collection of information**. The **collection phase** is one of the phases of the DLC (SANT'ANA, 2016) that has privacy as a transversal factor since it is necessary to identify, in the sources used, the breach of privacy of individuals related to the data collected (SANT'ANA, 2016). Data collection that violates an individual's privacy can occur in two ways: Surveillance or Interrogation. Surveillance is perceived when data collection occurs through watching, monitoring, listening to, or recording an individual's activities. Rodrigues and Sant'Ana (2015) give examples of when surveillance can occur.

> For example, a service available on the internet can (...) perform surveillance actions such as targeting content based on data collected about the user's routes (including geographic coordinates, humidity, atmospheric pressure, and altitude); information about the data network and the device used to access it; history of voice commands; tastes and experiences concerning visited locations; time spent in a public or private place; information about the network connection; metadata of images, audios, and videos, among others (RODRIGUES; SANT'ANA, 2015, p. 3).

Interrogation, in turn, consists of collecting data through questions or interviews and can be treated as the pressure suffered by the individual to provide some information about themselves. An everyday example occurs when websites require users to fill out forms as a mandatory condition for granting them access (RODRIGUES; SANT'ANA, 2015).

Solove (2006) recognizes **information processing** as a second group of activities. As previously stated, this research adopts the phases of the DLC (SANT'ANA, 2016) and, therefore, we must relate the information processing described by Solove in the Taxonomy of Privacy (2006) as belonging to the **recovery phase** since it makes the data already stored available for access and use by the holders (SANT'ANA, 2016).

In the recovery phase, which involves processing, privacy is also a transversal factor since "[...] those involved with the content to be made available must be considered, identifying structures and possible users" (SANT'ANA, 2016, p. 18).

Solove (2006) recognizes information processing as a second group of activities. As previously stated, this research adopts the phases of the DLC (SANT'ANA, 2016) and, therefore, we must relate the information processing described by Solove in the Taxonomy of Privacy (2006) as belonging to the recovery phase since it makes the data already stored available for access and use by the holders (SANT'ANA, 2016).

In the recovery phase, which involves processing, privacy is also a transversal factor since "[...] those involved with the content to be made available must be considered, identifying structures and possible users" (SANT'ANA, 2016, p. 18).

Solove (2006) highlights five subgroups observable in information processing (recovery phase of DLC). Aggregation means crossing data from multiple sources, aiming to reveal hidden facts about the individual, facts that would not be revealed if analyzed in isolation (RODRIGUES; SANTANA, 2015, p. 40). According to Solove (2008), an information piece here and there may not say much about a person. However, once combined, various pieces can form a portrait of that person and reveal new information they did not expect third parties to know. Thus, Solove explains that aggregating information is not a new activity, as combining several pieces of data is always possible, putting two and two together to show us something new about the person (SOLOVE, 2008, p. 118).

In a didactic example, Solove (2008) explains that e-commerce uses aggregated data when suggesting products to an individual that they may be interested in based on their previous purchases. Aggregation violates the privacy of individuals when it unexpectedly combines data, revealing facts previously unknown to third parties. Identification means connecting data to individuals, relating or (re)identifying information to an individual based on data linkage. Solove (2008) explains that identification can be like aggregation since both involve data combinations. However, they differ in the sense that identification implies recognizing a person. For example, exhaustive aggregations of data about a person may occur in several databases, but this aggregation will not necessarily connect (identify) a person in their day-to-day life. Identification without aggregation exists, for example, at checkpoints, where people identify themselves, but there is not necessarily a repository of data about them (SOLOVE, 2008). Insecurity involves carelessness in protecting stored information against leaks and improper access, that is, from the moment a "[...] network is the target of unauthorized external data collection through techniques such as exploits, the result is a leak of personal data from which there is no possibility of returning to the previous stage" (RODRIGUES; SANT'ANA, 2015, p. 4). Secondary Use involves processing the collected information for a different purpose without the subject's consent. Exclusion concerns activities in the life cycle of their data opaque to users, such as data storage and sharing, and the lack of participation of these users in decisions regarding the collection, storage, recovery, and disposal of their data (RODRIGUES; SANT'ANA, 2015).

It is worth stressing that these activities do not involve data collection since that happened in the previous phase. Instead, these activities involve how the data is processed.

The third group refers to **Information Dissemination**, which is also related to the DLC **recovery phase**. It involves publishing, exposing, and disseminating information about individuals or the threat of doing so. This group comprises the following seven subgroups. Breach of Confidentiality breaks the trust between the parties regarding their commitment to confidential information. Breach of confidentiality can occur when a given service commits not to share information about its users with partners, but the user starts to receive advertisements from these partners.

Disclosure occurs when accurate information about someone comes to light to others, affecting the way others judge their character, or "when the repertoire of information available to their peers and their peers' connections is not transparent to users" (RODRIGUES; SANT'ANA, 2015, p. 4), resulting in judgments about their character based on the information revealed. Exposure involves the revelation of an individual's emotional or physical attributes, such as nudity, pain, or bodily functions, as occurs with the disclosure of intimate photos of an individual. Solove (2008) states that exposure can be like disclosure, as both involve sharing accurate information about a person. However, exposure involves information about our bodies and health, while disclosure includes a broader range of data. Increased Access means expanding the accessibility of information beyond what the parties involved expect. In this case, a breach of privacy occurs when, for example, a website shares its users' personal data with its partners beyond what users expect or what is necessary for the execution of the service. Blackmail refers to activities of control, domination, intimidation, or threats to individuals or groups by third parties, such as when criminals threaten an individual, extorting them to pay a determined amount to avoid disclosure of their information. Appropriation involves using an individual's identity to serve the objectives and interests of another or to endorse a service or product without the individual's due consent. Distortion consists of disseminating false, misleading, or contradictory information about individuals.

Finally, **Invasion**, also in the **recovery phase of the DLC**, encompasses activities that invade users' privacy. The group unfolds into two subgroups. Intrusion means invasive acts that disturb an individual's peace or solitude, such as "using services with the purpose of recording data about actions in a given environment, without the consent of the parties (...)" (RODRIGUES; SANT'ANA, 2015, p. 6). Decisional Interference involves government interference in private matters, that is, non-consensual intrusion by government agencies into an individual's life.

Therefore, the taxonomy of privacy can demonstrate connections between different harms and problems. In this way, various situations can be referred to as breaches of privacy because there are substantial similarities between them as much as divergences.

Analysis of *Black Mirror*'s episode *Joan is awful*

The Black Mirror series, created by Charlie Brooker, premiered in December 2011 and became a worldwide success. Each episode is independent and presents a unique story addressing various themes, such as social isolation, privacy, politics, and artificial intelligence. Generally, the series' episodes portray a dystopian reality where technology plays a central role in people's lives (RODRIGUES; SANT'ANA, 2019).

The series, currently available on the streaming platform Netflix, launched its sixth season in June 2023 with five episodes that tell independent stories.

The first episode of the series' sixth season, which we analyze in this work, is called Joan is Awful. Ally Pankiw directs the episode in question. According to the official synopsis, the episode shows how "An ordinary woman is shocked to discover that a global streaming platform has released a prestigious television adaptation of her life".

The episode, composed of 33 (thirty-three) sequences, according to our observations, portrays the story of Joan, played by Annie Murphy, a woman, in theory, ordinary, who also has ordinary routine activities: the opening sequences show Joan waking up, turning off her alarm clock; having her first meal of the day, prepared by her boyfriend; after that, she leaves home to go to work. Her life is, apparently, ordinary and would not be an interesting plot for any series, in theory. However, when choosing a series to watch with her boyfriend, Joan sees the premiere of Joan is Awful, whose protagonist physically resembles her and bears her name.

While watching the series, Joan realizes a streaming platform has transformed her life into a series starring a popular actress (Salma Hayek). The series exposes details of Joan's daily life, more dramatically, just a few hours after they happened in "real" life, bringing repercussions among Joan's friends and coworkers, who recognize her. As a result of the series, Joan is fired from her job (for having supposedly "revealed" her company's trade secrets in the series), and her boyfriend ends their relationship (due to a betrayal also portrayed in the series).

Thus, episode after episode reveals details of her private life, and Joan seeks out her lawyer, hoping to prevent further exposure of her life. However, the lawyer claims they can do nothing since Joan consented to the fictitious streaming company collecting, storing, and retrieving her data. Since she has no legal means of resolving the issue, Joan seeks on her own to destroy the quantum computer responsible for collecting, storing, and retrieving her data, thus ending the breach of her privacy by destroying it.

Of the 33 (thirty-three) sequences analyzed in this study, sequences 22, 24, and 25 were disregarded (not applicable) since the protagonist, Joan, does not appear in the scene nor is indirectly involved. It is worth noting we only focused on the breach of Joan's privacy in this study. We disregarded the breach of privacy of other characters.

Our analysis did not observe breaches of Joan's privacy in the following subgroups of Solove's Taxonomy of Privacy (2006):

I. In the interrogation subgroup (collection phase; collection group), the main character did not undergo any interrogation or give any interviews. All of her data, when collected, was collected through surveillance.

II. In the aggregation subgroup (recovery phase; processing group), the series merely reproduces Joan's life, and we cannot say there was a combination of data from different sources nor that a combination generated unexpected data.

III. In the insecurity subgroup (recovery phase; processing group), no third party invaded the system of the fictitious streaming service responsible for storing the main character's collected data.

IV. In secondary use (recovery phase; processing group), Joan did not initially know she was under surveillance. In theory, she did not know the purposes of this surveillance and, therefore, could not have her data distorted. After learning about the surveillance, the lawyer shows that Joan gave her consent for the purposes used by the streaming company. The legality or effectiveness of this consent was not the subject of analysis in this research, which we limited to the breach of privacy.

V. In exclusion (recovery phase; processing group), a lawyer informs Joan that she had consented to the terms of use; that is, she had agreed to the platform using her data and, in theory, would know how the streaming platform could use it.

VI. In breach of confidentiality (recovery phase; dissemination group), the character Joan theoretically consented to sharing her data with third parties, since, according to the lawyer who analyzed the document, the streaming service informed her about the possibility of using the data to produce a series.

VII. In exposure (recovery phase; dissemination group), the series did not disclose intimate photos of Joan since all the content shown in the fictional series was a cinematic and dramatized reproduction.

VIII. In increased access (recovery phase; dissemination group), the platform did not theoretically breach confidentiality since the terms of use duly described the extent of access to Joan's data. She formally consented to all access and use of her data.

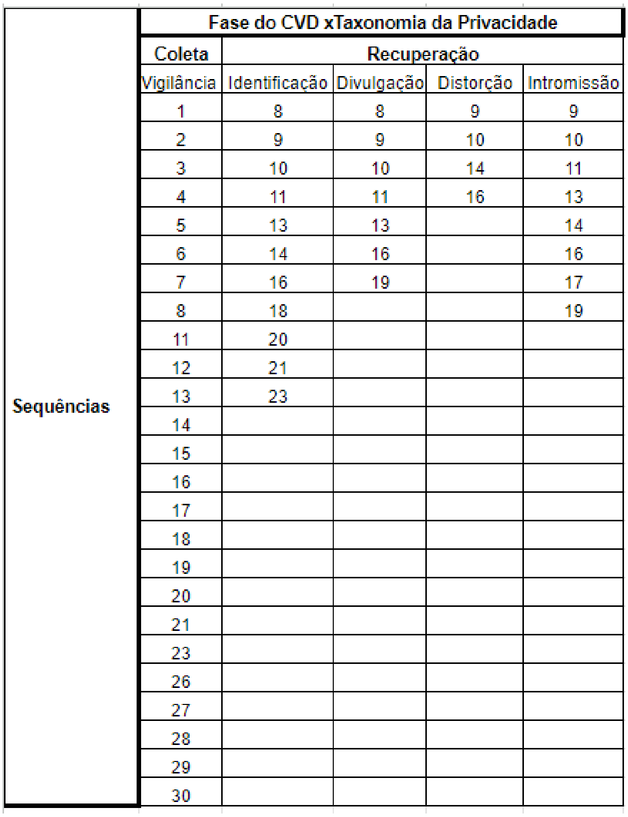
IX. In the blackmail subgroup (recovery phase; dissemination group), the main character did not suffer an attempt at extortion to prevent her data from being disclosed.

X. In appropriation (recovery phase; dissemination group), Joan supposedly had formal knowledge of how the platform would recover her data. So, there was no misappropriation, and the platform did not violate her privacy.

XI. In the decisional interference subgroup (recovery phase; invasion group), there was no government interference in the main character's life.

Thus, based on Solove's Taxonomy of Privacy (2006), the character Joan has her privacy broken into the following subgroups and respective groups presented below in Table 1:

Table 1 – Sequences according to the DLC and the Taxonomy of Privacy



	Fase do CVD xTaxonomia da Privacidade				
	Coleta	Recuperação			
	Vigilância	Identificação	Divulgação	Distorção	Intromissão
Sequências	1	8	8	9	9
	2	9	9	10	10
	3	10	10	14	11
	4	11	11	16	13
	5	13	13		14
	6	14	16		16
	7	16	19		17
	8	18			19
	11	20			
	12	21			
	13	23			
	14				
	15				
	16				
	17				
	18				
	19				
	20				
	21				
	23				
	26				
	27				
	28				
	29				
	30				

Source: The authors.

Thus, based on Solove's Taxonomy of Privacy (2006), the character Joan has her privacy broken into the following subgroups and respective groups presented below in Table 1:

We consider the streaming service breached Joan's privacy in the following opportunities:

a. Surveillance (collection phase; collection group): the breach began during the collection of Joan's data, which occurred through surveillance (subgroup). Sequences 9 and 10 do not apply to the collection subgroup since the main character is absent. Within the collection phase, through surveillance, for example, the protagonist's privacy suffers a breach because she does not effectively know that her routine is being mapped and monitored through a surveillance process of constant data collection. We observed that data collection (collection phase) through surveillance occurred in 25 (twenty-five) sequences, except for those we previously mentioned. Thus, until the destruction of the quantum computer, the equipment responsible for storing the collected data in sequence 30, Joan's data was, theoretically, being collected non-stop.

b. Identification (recovery phase; processing group): the main character is identifiable in a total of 11 (eleven) sequences with emphasis on sequences 9 to 11 in which people from Joan's social circle recognize her as a "character" in the series.

c. Dissemination (processing group and recovery phase): we observed changes in how others judge the protagonist as a character in 7 (seven) sequences when they begin recognizing her as a "terrible" person.

d. Distortion (recovery phase; dissemination subgroup): we could see the breach of privacy due to distortion in 04 (four) sequences, such as in the sequences in which the series within the series portrays Joan's life exaggeratedly when compared to the original to increase the drama and generate more audience.

e. Intrusion (recovery phase; invasion group): we identified 08 (eight) sequences in which it was possible to glimpse the breaking of Joan's

> solitude, causing her discomfort, since a series was portraying her life to all "subscribers" of the fictional streaming service.

In this sense, we could connect Solove's Taxonomy (2006) to the DLC (SANTANA, 2016) since an individual can have their privacy violated at any stage of the DLC, namely collection, storage, recovery, and disposal. Within each stage mentioned above, privacy violation can happen differently based on the groups and respective subgroups presented by Solove (2006).

Discussion and reflections

The episode under study presents an augmented and dramatized reality of the cycle of information collection, storage, recovery, and disposal of personal data by holders. Still, we experience the reality of having our data used for reasons that are opaque to us.

The user-friendly appearance of ICTs and the ease and speed at which these devices operate fully satisfies users' desire for information. However, users are unaware of how these processes occur and at what cost. The episode under study demonstrates the existence of layers of abstraction through which data flows and how users are unaware (SANT'ANA, 2021).

As shown in the series, a considerable and unknown volume of data is collected and recovered from the use of ICT, an occasion in which the average user is unaware of the volume of data generated in the various and routine operations that they perform in their daily lives, since ICT only presents the bioavailable information on users' screens.

This gap between the information holder's knowledge and the users' ignorance can make the latter vulnerable to the wishes of the former. Thus, even though technology has advanced and become capable of storing large volumes of information efficiently, we are still physically and biologically carriers of the same human weaknesses and dependent on data being treated and presented as bioavailable information (SANT'ANA, 2021).

When consulting their lawyers, the characters of Joan (sequence 15) and actress Salma Hayek, playing herself (sequence 22), are informed that the Terms and Conditions of the fictitious streaming company describe the entire DLC. The Terms and Conditions are a statement issued by information holders and addressed to users detailing how a given service works and how the organization processes personal data. The holders claim these documents are sufficient for the user to understand and effectively comprehend how their data will be processed. The legal validity of this argument is not within the scope of this study.

However, neither of the two characters had read the documentation; even if the information is formally presented and available to users, there is no guarantee they will read and effectively understand what these voluminous documents describe, which prevents them from acquiring due knowledge about the DLC of their data, operated by large corporations, which occupy the role of holders.

Given the complexity, volume of pages, and often robust language generally common in documents of this nature (as demonstrated in the episode), most users (real and not just fictional ones) do not access this type of information and remain unaware of how companies process their data, for what purposes, and who will have access to it.

Following the lawyer character, we consider Joan supposedly knew how the streaming platform would use her data by accepting the Terms of Use and thus disregarded the secondary use, exclusion, and increased access groups.

However, consent is merely formal and cannot properly inform the user how the company will process their data. The characters Joan (Annie Murphy) and Salma Hayek (playing herself) were surprised by the provisions of the Terms of Use when informed by their lawyers.

In this sense, even though it is a fictional work, the lack of knowledge about the Terms of Use also seems recurrent among real-life users since, after the premiere of the episode Joan is Awful, for example, the search for Netflix's Terms and Conditions of Use increased by 596% (five hundred and ninety-six percent). Therefore, we recommend further studies on the effectiveness of the Terms of Use regarding the assimilation of information by these users.

Final considerations

The perception of privacy is often an intuitive feeling, but it is complex to conceptualize it in only one way. Since privacy is a factor that permeates all phases of the Data Life Cycle (SANT'ANA, 2016), it can suffer breaches in any of these phases.

Drawing from Solove's (2006) considerations, this article examined an audiovisual work, the episode Joan is Awful from the Black Mirror series, seeking to bring greater concreteness to the Taxonomy of Privacy based on the situations experienced by the protagonist Joan (Annie Murphy). Thus, we observed the complexity of the breach of privacy since the violation of the character Joan's privacy occurred in 5 (five) distinct subgroups of the 16 (sixteen) proposed by Solove (2006) in the Taxonomy of Privacy throughout the episode.

The dissemination of audiovisual works like Joan is Awful is a powerful tool for increasing users' awareness of the constant data process based on ICT. Acceptance of the Terms and Conditions of Use is a requirement for accessing the streaming company's services. However, it seems to have been necessary to create an episode, supposedly dystopian, of a series with a large audience to generate effective curiosity among users about what these terms describe.

The relationships between consumption, communication, and society can be seen in an audiovisual work, even if it is fictional. We can identify similarities with what may become a reality. In this sense, we see personal data as the currency paid for using online platforms, websites, and services freely. This data is capable of inferring identities and behaviors from the aggregation of information from different sources.

From an economic perspective, personal data usage could increase the efficiency of commercial transactions in digital networks. By identifying consumer profiles, algorithms available on online platforms could improve their browsing and consumption experience since what they like most would be offered directly without wasting time (SILVEIRA; AVELINO; SOUZA, 2016).

However, despite providing a better user experience, we suggest furthering studies on the most effective ways and means for users to learn about the processing of their data, besides studies focused on the legality of the form of presentation of instruments of this nature since the effectiveness of Law 13,709 of August 14, 2018, General Personal Data Protection Law (LGPD).

References

AUMONT, J.; MARIE, M. *Dicionário teórico e crítico de cinema*. Campinas: Papirus, 2007.

BRASIL. *Constituição da República Federativa do Brasil*, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 11 jul. 2023.

BROOKER, C. *Black Mirror*. Reino Unido. Endemol/Netflix, 2011.

DONEDA, D. *Da privacidade à proteção de dados pessoais*. 2. ed. São Paulo: Thomson Reuters Brasil, 2021.

DUTRA, I. A influência da literatura de ficção-científica na técnica cinematográfica: uma análise de o homem invisível de H.G. Wells e sua transposição fílmica homônima. *Baleia na Rede*, v. 1, n. 9, p. 209-221, 2012.

FOSS, J. C. Netflix: buscas por termos de uso disparam devido a Black Mirror. *Tecmundo*, 18 jun. 2023. Disponível em:
https://www.tecmundo.com.br/mercado/265747-black-mirror-buscas-termos-uso-netflix-disp ararem.htm. Acesso em: 06 ago. 2023

MENEZES, S. S.; ARAÚJO, R. F. Fanfiction de ficção científica: divulgação e incentivo à leitura sobre ciência. *Múltiplos Olhares em Ciência da Informação,* n. Especial, 2018. Disponível em: http://hdl.handle.net/20.500.11959/brapci/106502. Acesso em: 04 ago. 2023.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. *Declaração Universal dos Direitos Humanos*, 1948. Disponível em:
https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos. Acesso em: 12 jul. 2023.

RODRIGUES; F. A; SANTANA, R. C. G. Uso de taxonomia sobre privacidade para identificação de atividades encontradas em termos de uso de redes sociais. En XII CONGRESO ISKO ESPAÑA Y II CONGRESO ISKO ESPAPA-PORTUGAL, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia. Disponível em: https://www.researchgate.net/publication/294728457_Uso_de_taxonomia_sobre_privacidade _para_identificacao_de_atividades_encontradas_em_termos_de_uso_de_redes_sociais. Acesso em: 06 ago. 2023.

RODRIGUES, F.A.; SANT'ANA, R.C.G. Ficção Científica e Realidade da Coleta de Dados em Redes Sociais Online (Black Mirror). In: MORAES, J.A.; RODRIGUES, F. A.; PANTALEÃO, N. C. A.(Orgs.). *Tecnologia e Sociedade*: discussões contemporâneas. São Paulo: FiloCzar, 2019

SANTOS, P. L. V. A. C.; SANTANA, R. C. G. Dado e Granularidade na perspectiva da Informação e Tecnologia: uma interpretação pela Ciência da Informação. *Ciência da Informação*, v. 42, n. 2, 2015. DOI: 10.18225/ci.inf.v42i2.1382. Disponível em: https://revista.ibict.br/ciinf/article/view/1382. Acesso em: 12 fev. 2023.

SANT'ANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. *Informação & Informação*, v. 21, n. 2, p. 116â142, 2016. DOI: 10.5433/1981-8920.2016v21n2p116. Acesso em: 25 jun. 2021.

SANT'ANA, R. C. G. Transdução informacional: impactos do controle sobre os dados. In: MARTÍNEZ-ÁVILA, D., SOUZA, E. A., and GONZALEZ, M. E. Q., eds. *Informação, conhecimento, ação autônoma e big data: continuidade ou revolução?*. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica; FiloCzar, 2019, pp. 117- 128. ISBN: 978-85-7249-055-9. Disponível em: http://books.scielo.org/id/gfrbh/pdf/martinez-9788572490559-09.pdf.
https://doi.org/10.36311/2019.978-85-7249-055-9.p117-128. Acesso em: 05 jun. 2022.

SANT'ANA, R. C. G. A transdução nos processos de mediação e a informação biodisponível. In: SMIT, J. W. et al (org.) *Humanidades digitais, big data e pesquisa científica*. São Paulo : Fundação Fernando Henrique Cardoso (FFHC), 2021. Disponível em: https://fundacaofhc.org.br/files/Humanidades%20Digitais.pdf. Acesso em: 04 jun. 2022.

SILVEIRA, S. A.; AVELINO, R.; SOUZA, J. A privacidade e o mercado de dados pessoais | Privacy and the market of personal data. *Liinc em Revista*, v. 12, n. 2, 2016. DOI: 10.18617/liinc.v12i2.902. Disponível em: https://revista.ibict.br/liinc/article/view/3719. Acesso em: 1 mar. 2024.

SOLOVE, D. J. A *Taxonomy of Privacy*. University Of Pennsylvania Law Review, 2006, 154(3), 477. http://doi.org/10.2307/40041279.

SOLOVE, D. J. *Understanding privacy*. Cambridge: Harvard University Press, 2008.

SOUZA, M. ; ALMEIDA, F. G. O comportamento do termo conhecimento na Ciência da Informação. *Revista Conhecimento em Ação*, Rio de Janeiro, v. 8, n. 1, p. 3–27, 2023. DOI: 10.47681/rca.v8i1.58126. Disponível em: https://revistas.ufrj.br/index.php/rca/article/view/58126. Acesso em: 21 maio. 2024.

WARREN, S. D.; BRANDEIS, L. D. *The Right to Privacy*. *Harvard Law Review*, v. IV, n. 5, 1890. Disponível em:
https://www.cs.cornell.edu/~shmat/courses/cs5436/warren-brandeis.pdf. Acesso em: 21 maio 2024.

WESTIN, A. *Privacy and freedom*. New York: Athenaeum, 1967.

About the authors

Ricardo César Gonçalves Sant'Ana is an Associate Professor at the São Paulo State University, UNESP, Faculty of Science and Engineering, FCE, Tupã Campus. He is a Professor in the Postgraduate Program in Information Science at the São Paulo State University, Marília Campus. He holds a PhD in Information Science (2008) and a Livre-Docência title in Information Systems from UNESP (2017). Sant'Ana specialized in Object Orientation (1996) and Information Systems Management (1998). He is a leader of the Data Access Technologies (GPTAD) Research Group and a member of the New Information Technologies (GPNTI) Research Group. He was President of the first composition of the Monitoring and Evaluation Committee of Undergraduate Courses - CAACG at UNESP between 2018 and 2020. E-mail: ricardo.santana@unesp.br. ORCID: https://orcid.org/0000-0003-1387-4519.

Dayane de Oliveira Martins earned a master's in Information Science from São Paulo State University (Unesp). She specialized in Civil Procedural Law at Anhaguera University, Uniderp (2017), and Civil and Business Law at Instituto Damásio de Direito da Faculdade, IBMEC SP (2020). She earned her bachelor's degree in Law from the Federal University of Goiás (2016). She is a licensed lawyer in Brazil and a member of the Bar Association of the State of Goiás. She currently works as a Technical Manager in a personal data management platform that offers outsourced DPO services (www.dponet.com.br), issuing legal reports. She is a member of the Data Access Technologies (GPTAD) Research Group. E-mail: dayane.martins@unesp.br. ORCID: https://orcid.org/0009-0000-8872-510X.

Date of submission: 9/3/2023
Date of acceptance: 10/28/2024