Italian Labour Law e-Journal
Issue 1, Vol. 16 (2023)
ISSN 1561-8048
https://doi.org/10.6092/issn.1561-8048/16472

Data privacy and digital work platforms in global perspective.

Sonja Mangold*

Abstract

The article focuses on privacy risks in the global platform labour market. As an empirical analysis of privacy statements shows, German, Chinese, and U.S. crowdsourcing platforms collect large amounts of personal data from workers and other users. Some platforms use detailed monitoring measures of work processes that are questionable from a data protection point of view. German platforms are relatively transparent in their data handling compared to non-European portals. They also stand out positively in terms of self- initiative precautions such as anonymisation and pseudonymisation of personal data. Nevertheless, there are also data protection gaps here. There is thus a need for further legal policy action to improve privacy rights for platform workers in the future.

Keywords: Digital work platforms; Global crowdsourcing market; Workers' privacy; Data protection.

1. Introduction.

In recent years, the innovative digital business model of crowdsourcing platforms has taken root worldwide.[1] Crowdsourcing describes a new way for clients, especially companies (so-called crowdsourcers), to outsource tasks through third-party platforms to a large community of Internet users (so-called crowd workers).[2] The tasks mediated via platforms are diverse. They range from well-paid, highly complex jobs in areas such as innovation or

* Postdoctoral researcher at the Faculty of Business Studies and Economics at the University of Bremen (Germany). This essay has been submitted to a double-blind peer review. It has been developed under the project "Platform work data protection in a national and international perspective", funded by the Hans Böckler Foundation.

[1] *See* World Bank Group, *World Development Report. The Changing Nature of Work*, Washington, 2019.

[2] *See* Jeff Howe's pioneer definition in Howe J., *The rise of Crowdsourcing*, in *Wired*, 1 June 2006, https://www.wired.com/2006/06/crowds/.

software development and design competitions to simple, low-paid micro jobs (such as image labeling or data categorisation). There is also a difference between jobs that are done purely digitally (so-called remote work) and location-dependent activities (so-called gig work).[3]

Crowdsourcing offers various new opportunities. From a business point of view, it can help solve organizational problems, support innovation and create a competitive advantage.[4] From the perspective of workers, labour platforms offer new job opportunities and flexibility in terms of working time and place. [5] On the other hand, from a labour law point of view, the rise of digital platforms poses serious challenges to decent work.[6] Crowd workers are typically considered as self- employed and thus usually excluded from labour law safeguards.

In the recent public debate, questionable working conditions on platforms such as unfair levels of pay or occupational stress and risks are criticised.[7] Another major concern related to decent work is ensuring data privacy and security.[8] Privacy challenges through crowdsourcing include risks of de-anonymisation attacks,[9]data theft or misuse. In addition, there are specific risks for worker privacy. In particular, platforms collect and analyse personal information for the matching of customers, tasks, and crowd workers. The use of digital ratings and reputation systems also produces a great deal of data.[10] Furthermore, platforms employ opaque AI-based managerial algorithms for task matching and monitoring and evaluating worker behaviour.

Privacy risks and surveillance practices on crowdsourcing platforms have been insufficiently researched to date. As far as is known, there is still no systematic empirical research on how platforms actually handle the personal data of users and crowd workers in their business practice. For this reason, there is not yet a sufficient basis for developing problem-specific legal policy responses. Finally, there is a lack of studies that examine data privacy issues in crowdsourcing from an international comparative perspective. A current

[3] *See* Leimeister J.M., Zogaj S., *Neue Arbeitsorganisation durch Crowdsourcing – Eine Literaturstudie,* Arbeitspapier Nr. 287, Hans-Böckler-Stiftung, Düsseldorf, 2013.

[4] Boudreau K.J., Lakhani K.R., *Using the Crowd as an Innovation Partner*, in *Harvard Business Review*, April 2013, 61-69.

[5] *See*, e.g. German Federal Ministry of Labour and Social Affairs, *Fair work in the platform economy*, 2020, available at https://www.denkfabrik-bmas.de/en/topics/platform-economy/summary-of-the-key-issues-paper-fair-work-in-the-platform-economy.

[6] *See*, for instance, Gyulavári T., Menegatti E. (eds.), *Decent Work in the Digital Age. European and Comparative Perspectives*, Hart Publishing, Kemp House, 2022.

[7] *See*, for instance, Leist D., Hießl Ch., Schlachter M., *Plattformökonomie - eine Literaturauswertung*, Bundesministerium für Arbeit und Soziales - Universität Trier, Institut für Arbeitsrecht und Arbeitsbeziehungen in der Europäischen Union (IAAEU), Berlin, June 2017.

[8] *See*, e.g. Policy Input by IT for Change, *Decent Work vis-a-vis Workers' Data Rights and Social Security Concerns in an Algorithmified Workplace. For the ILO Meeting of Experts on Decent Work in the Platform Economy*, October 2022, available at https://itforchange.net/decent-work-vis-a-vis-workers%E2%80%99-data-rights-and-social-security-concerns-an-algorithmified-workplace; Kandappu T., Friedman A., Sivaraman V., Boreli R, *Privacy in Crowdsourced Platforms*, in Zeadally S., Badra M. (eds.), *Privacy in a Digital, Networked World. Computer Communications and Networks*, Springer, Cham, 2015, available at https://doi.org/10.1007/978-3-319-08470-1_4.

[9] Lease M., Hullman J., Bigham J., Bernstein M., Kim J., Lasecki W., Bakhshi S., Mitra T., Miller R., *Mechanical Turk is Not Anonymous*, in *Social Science Research Network*, March 6, 2013, available at https://ssrn.com/abstract=2228728.

[10] *See*, e.g. Gogola M., *Digitale Ratings als rechtliche Herausforderung*, in Bader V., Kaiser S. (eds.), *Arbeit in der Data Society*, Zukunftsfähige Unternehmensführung in Forschung und Praxis, Springer Nature, Berlin, 2020, 173-188.

research project at the University of Bremen (Germany),[11] aims to close these knowledge gaps. Our study operates at the intersection of law and empirical economic research.[12] For the first time, we comprehensively investigate data protection risks on German, Chinese, and U.S. crowd work platforms. The market for crowdsourcing portals has grown steadily in all three countries in recent years.[13] Researching data privacy issues on platforms in Germany, China and the USA seems worthwhile not least because the data protection frameworks differ significantly.[14] In Germany, the EU General Data Protection Regulation (GDPR) provides comprehensive legal standards to protect the privacy of platform users and workers. The GDPR contains a number of core principles related to data processing that platform companies must observe, such as data minimisation, transparency or valid consent. China's data protection regime was fragmented and opaque for a long time. In recent years, however, there have been tendencies toward legal standardisation. In 2021, China's National People's Congress Standing Committee adopted the Personal Information Protection Law (PIPL)[15] which lays down for the first time a comprehensive set of rules around data protection in the digital economy. In the USA, so far, there is no federal omnibus data protection law. Legal requirements on data privacy that may affect platforms are scattered across numerous sectoral and state laws. California is considered a pioneer regarding consumer privacy in the digital economy.[16] Overall, U.S. data protection regulation is less restrictive than European regulation.

2. Research objectives and methods.

With our study, we want to gain an in-depth insight into data protection practice in the global crowdsourcing market. We have investigated the following research questions: (1) To what extent do platforms in Germany, China and the USA collect personal and sensitive data from crowd workers and other platform users?[17] (2) How is personal data collected and used

[11] For more information, *see* https://www.boeckler.de/de/suchergebnis-forschungsfoerderungsprojekte-detailseite-2732.htm?projekt=2021-130-2.

[12] *See* on empirical research approaches in law, e.g. Gelbach J.B., Klick J., *Empirical Law and Economics*, in Parisi F. (ed.), *The Oxford Handbook of Law and Economics*, Oxford University Press, Oxford, 2017, 29-59.

[13] *See* Serfling O., *Crowdworking Monitor Nr. 1. Für das Verbundprojekt "Crowdworking Monitor"*, Bundesministerium für Arbeit und Soziales, September 2018, available at: https://www.bmas.de/SharedDocs/Downloads/DE/Meldungen/2018/crowdworking-monitor.pdf?__blob=publicationFile&v=1; Zhou I., *Digital Labour Platforms and Labour Protection in China*, ILO Working Paper 11, October 2020, available at https://www.ilo.org/global/publications/working-papers/WCMS_757923/lang--en/index.htm; World Bank Group, nt. (1).

[14] For an in-depth analysis of the legal framework for data privacy on crowdsourcing platforms in Germany, China and the USA, *see* Hornuf L., Mangold S., Yang Y., *Crowdsourcing and Data Privacy - A Comparison of Selected Problems in China, Germany and the USA*, Springer International, Berlin, 2023 (forthcoming).

[15] The text of the law is available at https://digichina.stanford.edu/work/translation-personal-information-protection-law-of-the-peoples-republic-of-china-effective-nov-1-2021/.

[16] In recent years, California passed the California Consumer Privacy Act (CCPA) and the California Privacy Rights Act (CPRA), which have similarities to the GDPR standards. The CCPA has a broad definition of "consumers", which may include crowd workers.

[17] A first exploratory evaluation of the privacy statements of the platforms showed that they often do not differentiate in terms of data collection according to the user group (website visitors, clients, crowd workers).

to monitor and control work processes? (3) To what extent do platforms disclose platform users' and workers' personal data to third parties? (4) Do the platforms take own precautions that indicate an appropriate level of data protection? (5) What legal policy action is needed at the European level to increase crowd workers' privacy?

In a first step, we have identified all active crowdsourcing platforms with headquarters and/or physical locations in Germany, China and the USA. In this regard, we were able to draw on existing market studies.[18] In addition, we have conducted intensive Internet research. We have focused on platforms that mediate digital remote work. Portals that offer location-based gig work (such as taxi service provider Uber) have been excluded from the sample. In total, we have identified 47 German, 145 Chinese and 293 U.S. crowdsourcing platforms. To identify data privacy risks, we have evaluated all the data privacy statements that can be found on the websites of the portals. The privacy statements are a main source of information on the platforms' data privacy practices. We have coded the retrieved documents using a variable scheme and then analysed them using descriptive statistical methods. In terms of methodology, we have been inspired by existing empirical studies on data privacy issues in the digital market.[19] In the following sections, I will present selected key findings of our study.[20]

3. Data handling and privacy risks on platforms.

3.1 Information on the collection and use of personal data.

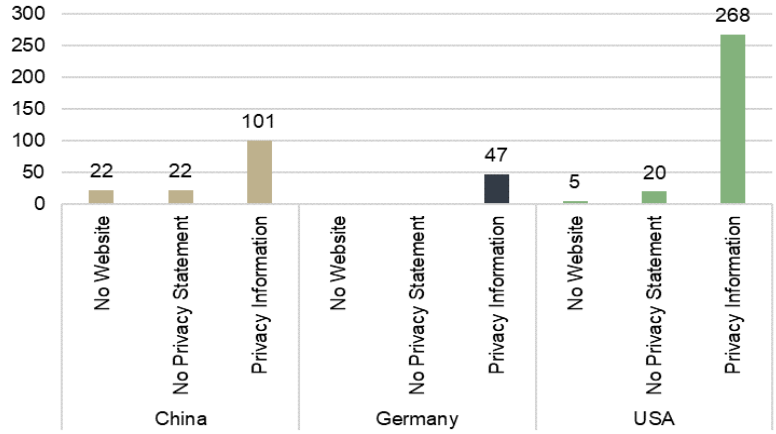
By publishing privacy statements, platforms often comply with legal obligations to inform users and workers about the processing of personal data.[21] Our study found that all German crowdsourcing platforms have published a privacy statement on their website (*see* Figure 1). The vast majority of U.S. portals also provide privacy information for users. In contrast, a significant number of Chinese platforms (44) have no website or no available privacy information. Chinese portals also frequently use geo-blocking to prevent user access to website content from Europe. This indicates that the Chinese digital platform labour market is not open to foreigners.

Our study further showed, that the privacy statements of platforms are quite long. The average reading time was more than 15 minutes. This suggests that many users and workers do not read the privacy statements at all or not completely. Standardization or alternative ways of presenting information, such as icons,[22] could be a more user-friendly solution.

[18] *See*, e.g. Zhou I., nt. (13); Mrass V., Peters C., Leimeister J.M., *Handlungsbroschüre Crowdworking- Plattformen-Neue Organisationskonzepte für Dienstleistungen nachhaltig gestalten*, Kassel University Press, Kassel, 2018.
[19] *See* in particular, Hornuf L., Dorfleitner G., *FinTech and Data Privacy in Germany. An Empirical Analysis with Policy Recommendations*, Springer International, Berlin, 2019.
[20] The data of the study was last checked in 2021. For a comprehensive presentation of the results of the study, *see* Hornuf L., Mangold S., Yang Y., nt. (14).
[21] *See* in particular the information rights under Arts. 13–14 GDPR.
[22] Geminn C.L., Francis L., Herder K.R., *Die Informationspräsentation im Datenschutzrecht – Auf der Suche nach Lösungen*, in *ZD- Aktuell*, 05335, 2021.

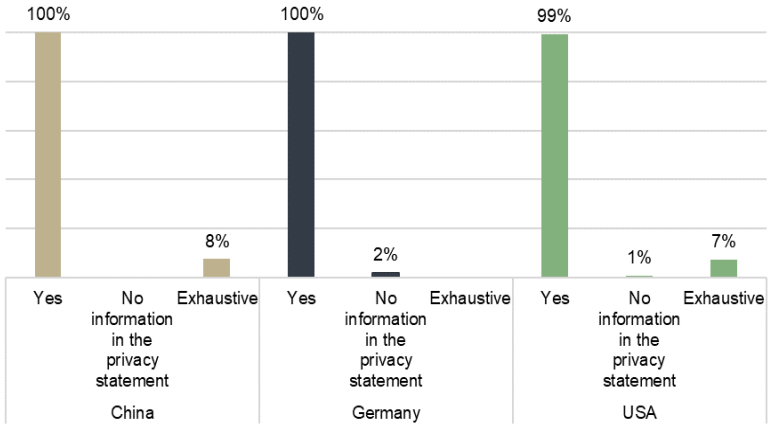
All platforms that have a privacy statement say they collect and process personal data (Figure 2). However the portals seldom state conclusively what personal data are processed. Therefore it can be assumed that platforms collect even more personal data in their practice than they publicly admit.

Figure 1. *Frequency of platforms providing a privacy statement. Distinction by country.*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

Figure 2. *Frequency of privacy statements indicating that personal information is processed. Distinction by country.*



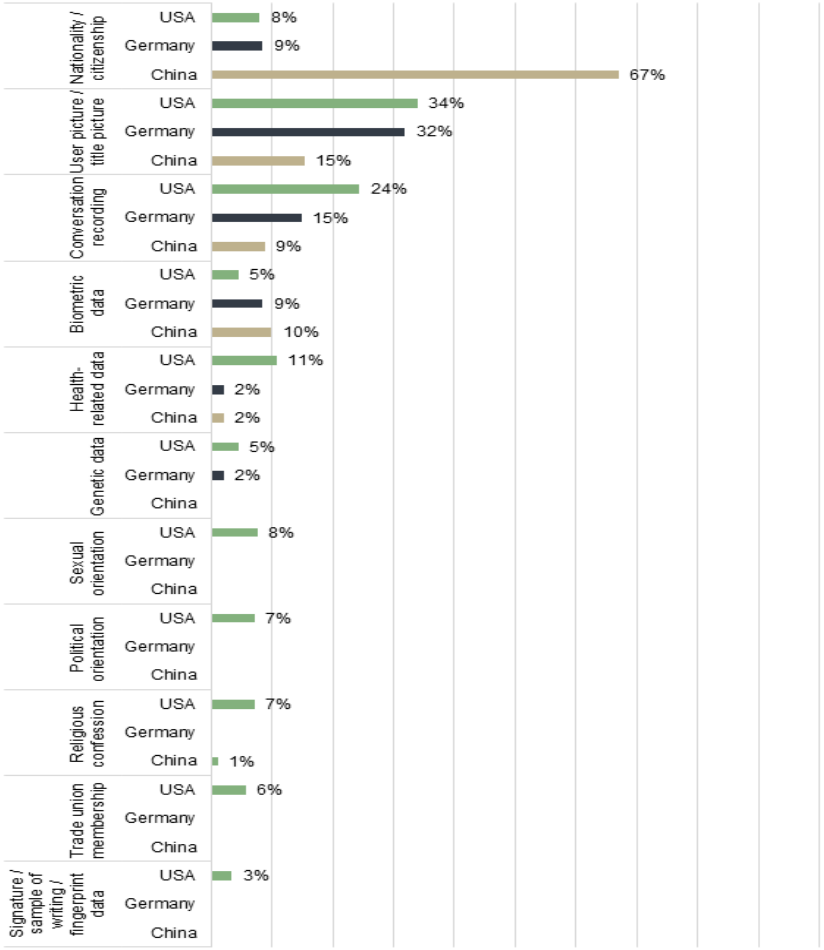
Source: Hornuf L., Mangold S., Yang Y., nt. (14).

A closer look showed that platforms in the three countries collect many different kinds of personal identifiable information. According to the privacy statements, most platforms process the email address, name, address and phone number of their users. The platforms also frequently collect bank data from clients for payment processing purposes. Most of the German and U.S. portals (91 %) state that they store the IP address.[23] In contrast, only a few Chinese platforms mention the IP address. Chinese crowdsourcing providers process passwords (91%), age (67 %) and passport data (65 %) of users comparatively frequently.

[23] According to the Court of Justice of the European Union (CJEU), a dynamic IP address with additional information can constitute personal data; *see* CJEU, 19 October 2016, C-582/14, Breyer.

In addition, crowdsourcing portals collect information about their users that can be classified as highly sensitive. EU law, in Art. 9 GDPR, contains specific requirements as additional safeguards to protect sensitive data such as racial or ethnic origin, political opinion, sexual orientation, biometric data or health data. Processing of such data is admissible only in exceptional cases. The Chinese PIPL also imposes higher legal requirements on the processing of sensitive information. In U.S. law, on the other hand, there is no overarching principle according to which sensitive data is specially protected. As can be seen in Figure 3, crowdsourcing platforms from all three countries collect sensitive information. Particularly frequently stored are photos of users, conversation recordings,[24] biometric data and data on nationality. German platforms are quite restrained in comparison when it comes to collecting sensitive data. Chinese portals process information on the nationality with striking frequency. U.S. platforms collect a particularly wide range of sensitive information, including data on political opinion, sexual orientation or trade union membership. This could be due to the fact that U.S. privacy laws generally permit the processing of sensitive information.

Figure 3. *Sensitive data processed according to the privacy statements.*



Source: H Hornuf L., Mangold S., Yang Y., nt. (14).

[24] User photos or voice recordings can be seen as "biometric data" if they are processed for the unambiguous identification of a natural person, *see* also GDPR recital (51).

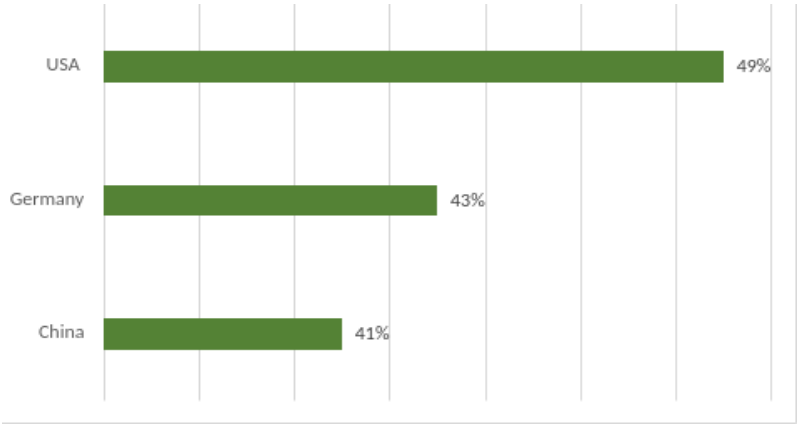
We also wanted to know whether and to what extent crowdsourcing platforms use customer profiling. The creation and analysis of user and personality profiles for commercial purposes is a major challenge for data protection in the digital economy. Web-based services often collect personal information from users and enrich the data from other sources in order to construct comprehensive profiles.[25] This usually involves Big Data techniques. On the one hand, profiling can help provide better services. On the other hand, it can be used for targeted advertising and marketing purposes. There is also serious concern that personal information contained in user profiles is sold to third parties for profit.[26] The EU GDPR imposes conditions and restrictions on automated profiling evaluating personal characteristics (*see*, in particular, Art. 22 GDPR). As Figure 4 shows, user profiling is relatively common on crowdsourcing platforms. Around half of the U.S. platforms say they construct customer profiles to improve the offer or for commercial purposes. Only slightly fewer German and Chinese privacy statements also mention the establishment of user profiles. A number of platforms emphasize that they only build user profiles on anonymous or pseudonymous bases. However, the privacy statements also contain indications of questionable privacy practices. U.S. platforms in particular sometimes create extensive customer profiles by combining collected personal data with data sets from numerous third-party sources. Several U.S. platforms state that they enrich profile data about crowd workers with information from business partners, credit agencies, employers, public agencies or social media networks. A U.S. marketplace platform states that it receives personal data from numerous third party sources, including data brokers in order to provide personalised advertising and to tailor its services to individual preferences. The practice of platforms to combine and analyse masses of personal data about users and workers from different sources seems problematic from the perspective of the principle of data minimisation.

Moreover, the evaluation of the privacy statements revealed that German and U.S. platforms make extensive use of cookies to track user behaviour.

[25] *See* on the problem early on Ladeur K.H., *Datenschutz und Datenverarbeitung bei neuartigen Programmführern in "virtuellen Videotheken". Zur Zulässigkeit der Erstellung von Nutzerprofilen*, in *Multimedia und Recht (MMR)*, 2000, 715-721.

[26] *See*, e.g. Hasan O., Habegger B., Brunie L., Bennani N., Damiani E., *A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case*, 2013, available at: https://perso.liris.cnrs.fr/omar.hasan/publications/hasan_2013_bigdata.pdf.

Figure 4. *Do the privacy statements of the platforms contain information that user profiles are created, e.g. for marketing purposes or to improve the offer?*

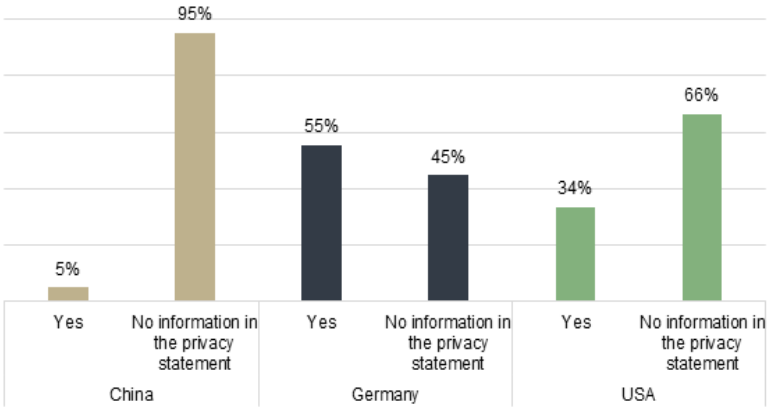


Source: Hornuf L., Mangold S., Yang Y., nt. (14).

3.2 Processing of personal data of crowd workers.

A further central aim of our study was to find out how platforms in Germany, China and the USA process personal data, especially of crowd workers. As Figure 5 shows, many platforms do not differentiate in their privacy statements between the data collected from crowd workers and other user groups. German platforms are comparatively transparent in this regard. 55% of German portals provide specific information on the processing of workers' personal data. After all, 34% of U.S. platforms provide differentiated information. In contrast, only 5% of Chinese privacy statements contain information on how personal data of crowd workers are processed.

Figure 5. *Frequency of privacy statements that differentiate between the data collection from crowd workers and other groups (clients, website visitors).*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

A closer look showed, that numerous platforms that provide differentiated information carry out extensive data collection and analysis in selection and work processes. Most of the

German and U.S. platforms state that they collect applicant data (CV, education, qualifications, certificates, etc.) from crowd workers. Some platforms request additional information about social media profiles. Several privacy statements also say that identification data (passport, driver's license, etc.) are processed for identity verification purposes. A larger number of U.S. platforms state that they conduct background checks on crowd workers. Unlike in Europe, background checks are common in the work context in the USA and are also frequently permitted by law.[27]As part of the background check, U.S. platforms sometimes ask about criminal records. Occasionally, crowd workers are even required to pass drug tests. Furthermore, a few platforms state that they conduct personality tests or demand information from crowd workers about their hobbies and private lives.

A greater number of German and U.S. platforms say they store work history and individual work performance. Often, the duration of job completion, availability and rejection rates are recorded. Some platforms mention extensive controls and monitoring of work behaviour. For example, a market-leading U.S. freelance platform admits to automatically gathering information such as number of mouseclicks, keyboard strokes and regularly taken screenshots as part of a "work diary". A privacy statement from a German testing platform similarly states that data on mouse movements, screenshots and microphone input are stored and transmitted. Several platforms say they control work processes using GPS tracking. Such practices of detailed monitoring and surveillance significantly encroach on the personal sphere of crowd workers and are problematic from a data protection point of view.

Furthermore, numerous German and U.S. platforms state that they collect and analyse feedbacks, ratings and review data on the performance of individual crowd workers. If customer ratings are kept by the platforms for a long time and used for digital reputation systems, this can be questionable under the principles of data minimisation and storage limitation according to Article 5 GDPR.[28]

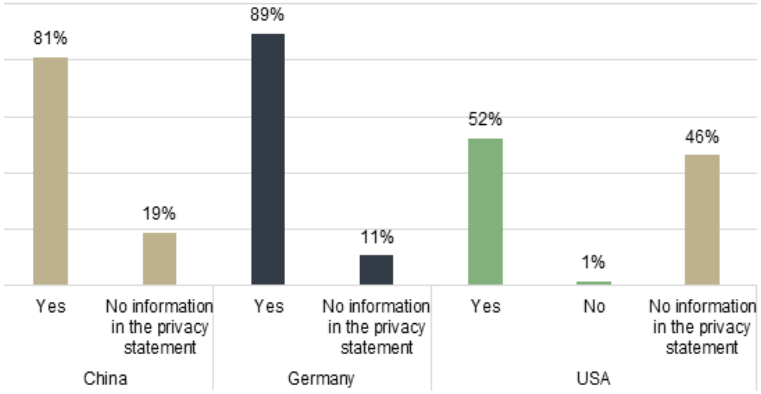
A limited number of privacy statements contain information on the use of automated or semi-automated decision-making systems in work processes. For example, a U.S. freelancer platform states that it uses AI to search the resumes of crowdworkers for efficient task matching. A U.S. survey portal asks for sensitive data such as political opinion, union affiliation or sexual orientation in order to assign tasks with the help of automated systems. Occasionally, platforms mention the right of crowd workers to object to automated decisions. Several German and U.S. platforms say they do not use automated decision-making at all. For example, a German freelancer portal emphasizes in its privacy statement: "As a responsible company, we do not use automatic decision-making or profiling."

[27] Schemmel F., *Background Checks aus internationaler Perspektive*, in *Zeitschrift für Datenschutz*, 2022, 541-545.
[28] *See* also Gogola M., nt. (10), 173-188.

3.3 Disclosure of personal information.

In our study, we also wanted to know whether crowdsourcing platforms transfer users' personal data to third parties such as public authorities or other companies. Data disclosure and sharing are associated with a loss of control over data and are therefore a major privacy concern in the digital economy. According to Articles 13 and 14 GDPR, platforms must inform their users about disclosures and to whom the data have been or will be disclosed. Article 23 of China's PIPL also stipulates information obligations with regard to the sharing of data. As Figure 6 shows, 89% of German platforms and 81% of Chinese platforms state that they transfer users' personal information to other parties, mostly with their consent. More than half of the U.S. privacy statements also mention data disclosures.

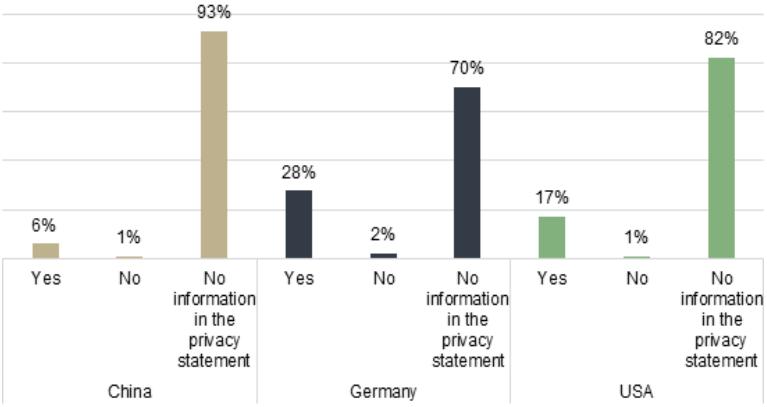
Figure 6. *Do platforms transfer personal data to other parties?*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

In the next step, we investigated to what extent platforms transfer personal data of crowd workers to crowdsourcers or other recipients. As can be seen in Figure 7, almost one third of the German platforms state that they disclose personal data of crowd workers. Only 17% of the U.S. and 6% of the Chinese privacy statements provide such information. The information practice of the German platforms is thus comparatively transparent. Some platforms state that they forward CVs and resumes of crowd workers to potential clients. Several platforms say they transmit reviews and ratings about crowd workers. A U.S. freelancer platform states it shares profile data including work diary and work history with clients and other platforms. Another U.S. portal says that it gives clients the first name and location of the crowd worker for every job. The privacy statement of a German testing platform states that screen videos of work processes are made accessible to clients. On the other hand, a few crowdsourcing platforms in the three countries emphasize that they do not share personal data from crowd workers at all or only in anonymous form.

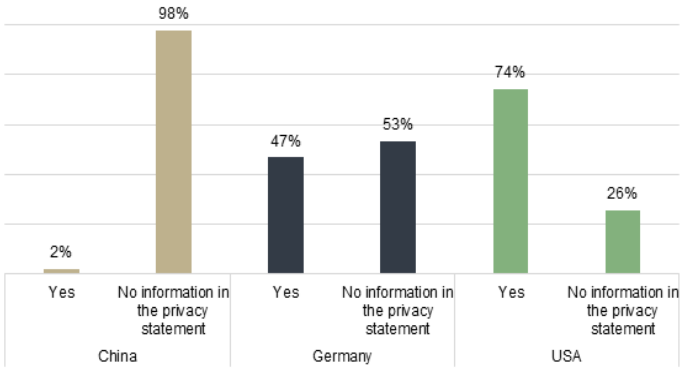
Figure 7. *Do platforms share crowdworkers' personal data with other parties?*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

Our study further showed that many platforms do not specify to whom the data is transferred. Legal transparency obligations are thus not met.[29] As can be seen in Figure 8, almost all Chinese privacy statements and more than half of German platforms are silent about the recipients of the data. 26% of U.S. platforms also do not specify to whom they transfer the data.

Figure 8. *Do privacy statements indicate to whom personal data is disclosed?*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

3.4. Data privacy measures of the platforms.

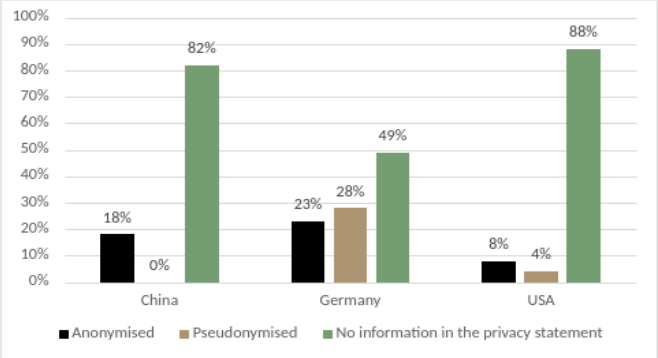
In our study, we also investigated whether and how platforms take precautions to protect the privacy of users and crowd workers. Self-regulatory initiatives for fair privacy practices can strengthen customer trust and increase the credibility of the crowdsourcing business model.[30] Two central tools for protecting the privacy of users are anonymisation and

[29] Under the GDPR (Articles 13 and 14), platforms must provide information about recipients or categories of recipients when they disclose personal user data to others. Similar requirements can be found in Chinese data protection law.

[30] *See* also Cherry M.A., Poster W.R., *Crowdwork, Corporate Social Responsibility, and Fair Labor Practices*, in Olleros X.F., Zhegu M. (eds.), *Research Handbook on Digital Trasformations*, Edward Elgar Publishing, Northampton, 2016.

pseudonymisation.[31] Under the GDPR, platforms can fulfill the principle of privacy by design (Article 25 GDPR) and data security obligations (Article 32 GDPR) by pseudonymising user data. As Figure 9 illustrates, that more than half of the German platforms mention that the personal data of users and crowd workers are anonymised or pseudonymised. Significantly fewer Chinese and U.S. privacy statements refer to such measures. The data protection practice of German platforms thus stands out positively, which could be related to the comparatively high legal standards of the GDPR. The closer document analysis showed that numerous platforms allow customers and crowd workers to use services under pseudonyms, e.g., under a freely chosen name. Rarely, the privacy statements even mention the possibility to act anonymously after registration.

Figure 9. *Are personal user data anonymised and/or pseudonymised by the platforms?*



Source: Hornuf L., Mangold S., Yang Y., nt. (14).

Furthermore, many platforms advertise on their websites that they have undergone privacy audits or have obtained privacy seals. For example, a number of U.S. portals mention that they have participated in the BBBonline or TRUSTe seal programmes[32] and thus have official approval for fair data protection practices. Several platforms state that they have globally recognised ISO 27001 certification[33] for data security. A German portal emphasizes that it participates in the IAB Europe Transparency and Consent Framework (TCF),[34] which is designed to ensure GDPR compliance. Several German privacy statements mention the voluntary Code of Conduct of the German crowdsourcing industry,[35] which contains a principle of data protection and confidentiality of personal information of crowd workers. Our analysis also revealed that some platforms are taking further measures to protect customers' privacy such as the use of encryption or staff training.

[31] Successful pseudonymisation makes it difficult to assign personal data to individuals. Anonymisation of data guarantees even greater privacy protection, since re-identification is in principle impossible. *See* in particular recital (26) GDPR for the definition of anonymisation and pseudonymisation. Anonymous data fall outside of data protection law.
[32] *See* https://trustarc.com/consumer-information/privacy-certification-standards/.
[33] More information available at: https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/Zertifizierung-und-Anerkennung/Zertifizierung-von-Managementsystemen/ISO-27001-Basis-IT-Grundschutz/iso-27001-basis-it-grundschutz_node.html.
[34] More information can be found at: https://iabeurope.eu/transparency-consent-framework/.
[35] *See* https://crowdsourcing-code.com/index-en.php.

4. Summary, conclusions and outlook.

Our study illustrates that German, Chinese and U.S. crowdsourcing platforms gather large amounts of personal information from users and crowd workers. Digital crowdsourcing offers new, increased opportunities to use and monetise data. First, platforms collect and exploit data in the same way as other online services. For example, as we have shown, they make extensive use of cookie tracking or customer profiling for targeted marketing purposes. In addition, they collect and process large amounts of registration, performance and reputation data from crowd workers. In this way, platforms harvest and retain large sensitive data sets from various sources. As our analysis showed, some platforms engage in detailed monitoring and surveillance of crowd worker activities that are questionable from a data protection perspective. The privacy concerns associated with digital crowdsourcing mentioned at the beginning are confirmed by our results in this respect. On the other hand, our study revealed positive approaches by platforms to actively protect the privacy of users and workers. For example, portals take technical protection measures such as anonymisation, pseudonymisation or encryption. Moreover, some platforms use certificates, privacy seals or audits.

The comparison of different country contexts showed that German platforms are relatively transparent in their handling of user and crowd worker data. German platforms also stand out positively in terms of protective measures such as anonymisation or pseudonymisation. This could be due to the comparatively high legal standards of the GDPR. However, we also identified gaps in the information provided by German portals, particularly with regard to data disclosures. Chinese platforms often lack transparency about their data handling. It remains to be seen whether the PIPL, which had only recently been passed at the time of our study, will lead to more openness of Chinese companies in the future. U.S. platforms collect and exploit personal and sensitive user data comparatively extensively. The U.S. privacy statements also showed some evidence of excessive control and monitoring of crowd workers. This finding could be related to the fragmented, lax data protection laws in the U.S.[36] Overall it can be concluded from our study, that transparency about data processing on platforms should be increased. Excessive intrusion into the privacy of crowd workers should be prevented. Legal grey areas, but also ambiguities regarding data collection in the new crowdsourcing business model, should be closed in the future. Further legal policy action thus seems necessary.

At the European level, several legislative initiatives have recently been taken that address aspects of privacy in the context of work platforms. In 2021, the European Commission launched a proposal for a Directive to improve the working conditions of platform workers.[37] The proposal requires, among others, that platforms must inform workers about automated monitoring and decision-making which are used to supervise or evaluate the work performance (Art. 6 I). As our analysis has revealed, platforms rarely provide information

[36] On limitations of U.S. data protection law in the employment context, *see* e.g. Kim P., *Data mining and the challenges of protecting employee privacy under U.S. law*, in *Comparative Labor Law & Policy Journal*, 40, 3, 2019, 405-420.
[37] *See* European Commission, *Proposal of 9 December 2021 for a Directive of the European Parliament and the Council on improving working conditions in platform work*, COM(2021) 762 final.

about the use of automated monitoring and decision-making in their privacy policies. The planned information rights can therefore contribute to greater transparency. They can also promote the exercise of platform workers' rights to review and object to automated monitoring and decision-making. The proposed Directive also protects against excessive data processing that is not strictly necessary for the performance of the contract (Art. 6 V). In particular, it contains restrictions on the processing of certain kinds of information such as data on private conversations, the psychological state or the health state of platform workers (Art. 6 V (a)-(c)). The proposal also stipulates that platforms shall not collect any personal data while the platform worker is not offering or performing platform work (Art. 6 V (d)). As our evaluation of the privacy statements proves, platforms process sensitive data such as health data or psychological information in recruitment and work processes. Our analysis further revealed cases of excessive monitoring and surveillance of platform workers. The planned regulations can help to prevent these legally dubious practices. Furthermore, the legislative proposal fosters social dialogue on algorithmic management by introducing collective information and consultation rights (Art. 9). The transparency requirements and restrictions on data collection of the draft directive apply irrespective of the employment status of the platform workers (Art. 10). Another relevant instrument is the EU institutions' proposal for an AI Regulation of 2021[38] which deals with AI-based data processing on work platforms. The 36th recital of the proposal holds that AI systems used in employment, workers management and access to self-employment should be considered as high-risk. Digital work platforms as users of such high- risk AI systems must comply with specific obligations and requirements on transparency. Also worth mentioning in the context of data collection through platforms is the "Platform-to-Business-Regulation" (P2B Regulation).[39] The P2B Regulation includes provisions on transparency of terms and conditions and ranking mechanisms for self- employed platform users (Art. 3, Art. 5). These provisions can be relevant for worker privacy, as platforms sometimes insert data-related clauses and phrases that require to consent to data collection into their general terms and conditions.[40]The above-mentioned provisions can also reduce information asymmetries in favour of the platforms with regard to the creation of data-based rankings and ratings. However, it must be noted that the personal scope of the P2B Regulation is limited to platform workers offering services to consumers. Another relevant instrument is the Directive on transparent and predictable working conditions,[41] which contains rules on transparency with regard to essential working conditions and names platform workers as possible addressees.[42]

[38] European Commission, *Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts*, COM(2021) 206 final.
[39] Regulation (EU) 2019/1150 *on promoting fairness and transparency for business users of online intermediation services*, in OJ L186/57.
[40] *See* e.g. the general terms and conditions of the German product testing platform `Mama- REPORTER´, which contain a consent clause for the collection and transfer of data from crowdworkers, available at https://www.mama-reporter.de/home/agb.
[41] Directive (EU) 2019/1152 of the European Parliament and of the Council *on transparent and predictable working conditions in the European Union*, in OJ L186/105.
[42] *See* Maul-Sartori M., *Die neue Arbeitsbedingungenrichtlinie: EU - Arbeitsrecht weiter auf dem Vormarsch,* in *NZA*, 2019, 1161-1167.

In the light of our findings, it is to be welcomed that the EU initiatives are aimed at improving the transparency of data processing on platforms. The planned extension of privacy rights to self-employed platform workers in the proposed Directive on platform work is appropriate to the socio-economic reality.[43] Strengthening social dialogue on algorithmic management can make privacy rights enforcement more effective.[44] In addition, in view of the data protection gaps that we have identified, further regulatory measures seem adequate which will only be touched upon here. Data is a significant source of revenue for platforms, but the workers and users who generate the data do not benefit from it. Given this, alternative ideas such as "data profit sharing"[45] should be considered. Through profit sharing systems easy to understand digital workers could receive appropriate rewards from the platforms. Moreover, boundaries of lawful and unlawful data collection on work platforms should be further legally clarified. Future regulation should also take into account that due to the information asymmetries in favour of the platforms, a voluntary, valid consent of the workers to the data processing often doesn't exist.[46]

Furthermore, self-regulatory activities of the platforms for data protection should be promoted and supported. As we have shown, there are already initiatives by crowdsourcing portals to better protect the privacy of users and workers. For example, platforms use anonymisation and pseudonymisation techniques or have gone through internationally recognised privacy audit and certificate programs. Self-regulatory approaches can rely on the knowledge and self-interest of relevant market players. Against the background that online platform markets are highly globalised, industry self-initiatives on data protection can provide adequate cross-border responses.

Our study on data privacy and crowdsourcing is not without limitations. We relied on the self-reports of platforms about their data handling. Platforms that provide local gig work were excluded from our sample. Through further research on data protection and gig work in Germany, China and the USA as well as participating observations, we will supplement and deepen our findings in the future.

[43] *See* on future developments in data protection law in the work context, e.g. Hendrickx F., *Privacy 4.0 at work: regulating employment, technology and automation,* in *Comparative Labor Law & Policy Journal*, 41, 1, 2019, 147-172.
[44] *See* also De Stefano V. and Taes S., *Algorithmic management and collective bargaining*, Foresight Brief 10, ETUI, Brussels, 2021, available at https://www.etui.org/publications/algorithmic-management-and-collective-bargaining.
[45] *See*, e.g. Huang Y., Zeng Y., Ye F., Yang Y., *Profit Sharing for Data Producers and Intermediate Parties in Data Trading over Pervasive Edge Computing Environments*, in *IEEE Transactions on Mobile Computing*, 22, 1, 2023, 429-442, available at https://ieeexplore.ieee.org/abstract/document/9405445/authors#authors.
[46] *See* also Policy Input by IT for Change, nt. (8).

Bibliography

Boudreau K.J., Lakhani, K.R., *Using the Crowd as an Innovation Partner*, in *Harvard Business Review*, April 2013;
Cherry M.A., Poster W.R., *Crowdwork, Corporate Social Responsibility, and Fair Labor Practices,* in Olleros X.F., Zhegu M. (eds.), *Research Handbook on Digital Trasformations,* Edward Elgar Publishing, Northampton, 2016;
De Stefano V., Taes S., *Algorithmic management and collective bargaining*, Foresight Brief 10, ETUI, 2021;
Gelbach J.B., Klick J., *Empirical Law and Economics*, in Parisi F. (ed.), *The Oxford Handbook of Law and Economics*, Oxford University Press, Oxford, 2017;
Geminn C.L., Francis L., Herder K.R., *Die Informationspräsentation im Datenschutzrecht – Auf der Suche nach Lösungen*, in *ZD- Aktuell*, 05335, 2021;
German Federal Ministry of Labour and Social Affairs, *Fair work in the platform economy*, 2020;
Gogola M., *Digitale Ratings als rechtliche Herausforderung*, in Bader V., Kaiser S. (eds.), *Arbeit in der Data Society,* Zukunftsfähige Unternehmensführung in Forschung und Praxis, Springer Nature, Berlin, 2020;
Gyulavári T., Menegatti E. (eds.), *Decent Work in the Digital Age. European and Comparative Perspectives*, Hart Publishing, Kemp House, 2022;
Hasan O., Habegger B., Brunie L., Bennani N., Damiani E., *A Discussion of Privacy Challenges in User Profiling with Big Data Techniques: The EEXCESS Use Case*, 2013;
Hendrickx F., *Privacy 4.0 at work: regulating employment, technology and automation,* in *Comparative Labor Law & Policy Journal*, 41, 1, 2019;
Hornuf L., Dorfleitner G., *FinTech and Data Privacy in Germany. An Empirical Analysis with Policy Recommendations*, Springer International, Berlin, 2019;
Hornuf L., Mangold S., Yang Y., *Crowdsourcing and Data Privacy - A Comparison of Selected Problems in China, Germany and the USA*, Springer International, Berlin, 2023 (forthcoming);
Howe J., *The Rise of Crowdsourcing*, in *Wired*, 1 June 2006;
Huang Y., Zeng Y., Ye F., Yang Y., *Profit Sharing for Data Producers and Intermediate Parties in Data Trading over Pervasive Edge Computing Environments*, in *IEEE Transactions on Mobile Computing*, 22, 1, 2023;
Kandappu T., Friedman A., Sivaraman V., Boreli R., *Privacy in Crowdsourced Platforms*, in Zeadally S., Badra M. (eds.), *Privacy in a Digital, Networked World. Computer Communications and Networks*, Springer, Cham, 2015;
Kim P., *Data mining and the challenges of protecting employee privacy under U.S. law*, in *Comparative Labor Law & Policy Journal*, 40, 3, 2019;
Ladeur K.H., *Datenschutz und Datenverarbeitung bei neuartigen Programmführern in "virtuellen Videotheken". Zur Zulässigkeit der Erstellung von Nutzerprofilen*, in *Multimedia und Recht (MMR)*, 2000;
Lease M., Hullman J., Bigham J., Bernstein M., Kim J., Lasecki W., Bakhshi S., Mitra T. and Miller R., *Mechanical Turk is Not Anonymous*, in *Social Science Research Network*, March 6, 2013;
Leimeister J.M., Zogaj S., *Neue Arbeitsorganisation durch Crowdsourcing – Eine Literaturstudie*, Arbeitspapier Nr. 287, Hans-Böckler-Stiftung, Düsseldorf, 2013;

Leist D., Hießl C., Schlachter M., *Plattformökonomie - eine Literaturauswertung*, Bundesministerium für Arbeit und Soziales - Universität Trier, Institut für Arbeitsrecht und Arbeitsbeziehungen in der Europäischen Union (IAAEU), Berlin, June 2017;
Maul-Sartori M., *Die neue Arbeitsbedingungenrichtlinie: EU- Arbeitsrecht weiter auf dem Vormarsch,* in *Neue Zeitschrift für Arbeitsrecht (NZA)*, 2019;
Mrass V., Peters C., Leimeister J.M., *Handlungsbroschüre Crowdworking- Plattformen- Neue Organisationskonzepte für Dienstleistungen nachhaltig gestalten*, Kassel University Press, Kassel, 2018;
Policy Input by IT for Change, *Decent Work vis-a-vis Workers' Data Rights and Social Security Concerns in an Algorithmified Workplace. For the ILO Meeting of Experts on Decent Work in the Platform Economy*, October 2022;
Schemmel F., *Background Checks aus internationaler Perspektive*, in *Zeitschrift für Datenschutz*, 2022;
Serfling O., *Crowdworking Monitor Nr. 1. Für das Verbundprojekt "Crowdworking Monitor",* Bundesministerium für Arbeit und Soziales, September 2018;
World Bank Group, *World Development Report. The Changing Nature of Work*, Washington, 2019;
Zhou I., *Digital Labour Platforms and Labour Protection in China*, ILO Working Paper 11, October 2020.