Data Scraping

What is data scraping?

Data scraping is the process of extracting data from online sources, including social media platforms and websites that host publicly accessible personal information. While the information may be publicly available in that the individual has chosen to post their own information online, organisations that use data scraping are responsible for the data that they collect from these sites, as they are onwards processing this information for their own purposes. The obligations under the UK GDPR and the Data Protection Act 2018 ("Data Protection Legislation") remain the same in respect of personal data, regardless of whether the information is publicly accessible or not.

Examples of data scraping may include social media monitoring for the purpose of measuring the impact of marketing, market research, using publicly available information for research.

There has been an increase in reported instances of data scraping, which have led to targeted cyberattacks, identity fraud, monitoring and profiling, unauthorised political and intelligence gathering purposes and unwanted direct marketing or spam. However, more broadly, individuals are losing control over their personal information, when it is being processed without their knowledge or out with their expectations.

What do you need to consider if your data processing involves data scraping of personal data?

The Information Commissioner's Office has provided best practice recommendations as follows in its Joint statement on data scraping and data protection | ICO.:

- ✓ **Transparency** – UK GDPR requires organisations that scrape data from publicly accessible sources to provide affected individuals with a privacy notice containing specific information prescribed in Article 14 of UK GDPR, which details what actions an organisation should take where personal data has not been obtained directly from the data subject, including providing privacy information and advice on how an individual can exercise their rights under UK GDPR.
- ✓ **Invisible Processing** – it is advisable for organisations that scrape personal data to only do so if they are able and willing to contact affected individuals directly to inform them about it to avoid "invisible processing', which is widely regarded to constitute a high-risk data processing activity. If invisible processing of personal data is happening, a Data Protection Impact Assessment may be required. Data Protection | StaffNet | The University of Aberdeen
- ✓ **Data Minimisation and Limitation** – the purpose limitation principle requires organisations to only collect and process personal data to achieve specified, explicit and legitimate purposes and not engage in further processing unless it is

compatible with the original purpose for which the data are collected. The data minimisation principle requires organisations to only collect and process personal data that are relevant, necessary and adequate to accomplish the purposes for which the data have been collected.

- ✓ **Satisfying a lawful basis** – organisations must satisfy one of the lawful bases available under Article 6 of UK GDPR in relation to any data scraping activities that they carry out. For more information on lawful basis, please visit Data Protection | StaffNet | The University of Aberdeen
- ✓ **Engagement of third parties** – if an organisation engages a third party service provider to carry out data scraping on its behalf, it will still be responsible for complying with all of the obligations set out above in respect of the data scraping and will also need to ensure that it has put in place a data processing agreement with the provider which addresses the requirements under Article 28 of UK GDPR.

You must contact the Information Governance Team at dpa@abdn.ac.uk if you are undertaking Data Scraping. The Team can assist you to ensure that you are processing data in compliance with Data Protection legislation.

Approval/Review History

Version	Date	Action
1.0	Prepared by Assistant Director (Information Governance & Security) February 2025	