

Economic and infrastructural risks to digital archives posed by AI-driven content extraction

Information Sheet

You are being invited to take part in a research study. Before you decide if you wish to participate, it is important for you to read this leaflet so you understand why the study is being carried out and what it will involve.

Reading this leaflet, discussing it with others or asking any questions you might have will help you decide whether or not you would like to take part.

What is the purpose of this study?

This research study is to understand the risks that archival bodies and digital collections are facing from AI-driven content extraction tools, or other digital extraction tools. We want to understand the impacts these archives are seeing from the advent of AI scraping and other digital extraction tools, how they conceptualise current and future risks, as well as potential benefits, from these extractive technologies, and what mitigation strategies they are currently using or planning to use.

The research is funded by Ekip, the European Cultural and Creative Sectors and Industries Policy Platform, an interdisciplinary four-year Horizon Europe project running from September 2023-August 2027.

Why do we want to talk to you?

We're looking to speak with people in the culture and creative industries and digital heritage sector - including people who manage online collections and archives within the UK and around the EU - about their experiences and thoughts around AI scraping of the digital commons and online archives and repositories.

What will you be asked to do?

We are running a workshop called an Open Innovation Factory (OIF), as part of the ekip programme. Each workshop is centered around a policy area or a speculative technology or scenario, which is introduced by invited experts before a workshop/interactive activities. This OIF is focused on the role of the digital commons.

We will use information gathered from the event to help to shape future policymaking around the digital commons for the CCIs and heritage institutions. To do this we are keen to document and learn from the events, which include capturing photographs or screen shots of workshop activities, materials, discussion boards, notes taken during breakout groups and discussions, and a recording of talks (in which participants may be included). All data collected from these events will be stored securely and appropriately.

Taking part in this study is entirely voluntary. We do not anticipate any specific risks to you taking part in this study. If you don't want to talk about or participate in something just say, and we will always respect your wishes. You can pause, or withdraw entirely from the study at any time, without giving any reason.

Confidentiality and data

'Your data' refers to the consent form and the data gathered as part of the event, specifically images, recordings, screenshots, or transcripts. The data will be used in our own reflection and reporting process, and in developing and publishing the findings of the research, with your express consent. Your data may be used and quoted in developing and publishing the findings of the research for ekip. It will be stored on a secure shared server data management portal at the University of Edinburgh for up to ten years past the end of the project.

All data we collect will be handled according to our Data Management Plan, which is available in full on request from the researchers.

Unless you state otherwise and we have your express consent, you will not be named in any public presentation of the research, your participation will be entirely anonymous, and your identity will not be revealed at any point or in any subsequent publications.

Use of results

The information gathered as a result of this study will be used to support the wider Ekip project. Please note, your participation (or not) in any Ekip research project will have no impact on your participation in other funded Ekip opportunities.

Will you be able to change your mind about taking part?

Yes, you can withdraw at any point. If you feel that you no longer wish to take part at any time just say so and we will stop, and destroy all identifiable records of your participation. You may also contact us after the study and ask us to withdraw your data from the study if it has not yet been anonymised.

Who are we?

This study is part of the larger Ekip project (<https://ekipengine.eu>), funded by the European Union's research and innovation programme Horizon Europe, administered through UK Research and Innovation. The programme is designed to improve open innovation for the Cultural and Creative Industries by increasing awareness, educating stakeholders, gather feedback, facilitate dialogue, influencing decision-making, and empowering stakeholders to participate in policy-making for positive societal change. This project is a partnership between the University of Edinburgh, Lund University (Sweden), Delft University of Technology (The Netherlands), Politecnico de Milano (Italy), Humak University of Applied Sciences (Finland), Technopolis Group (Belgium), Idea Consult Group (Belgium), Next Atlas (Italy), Creativity Lab (Estonia), Creative Business Network (Denmark), MSCOMM Group (Greece), New Moment network (South-Eastern Europe), Netherlands Institute for Sound & Vision (The Netherlands), City of Rotterdam (The Netherlands), Metropolitan Institute of Bratislava (Slovakia), Creative Industry Košice, (Slovakia) and Cité du design (France). This study is led by the University of Edinburgh.

What happens next?

If you are interested in taking part, please read carefully and sign the consent form when it is provided to you, and return this to Morgan Currie. If you have any other questions about the please contact Morgan Currie. Thank you for reading this information.

Morgan.currie@ed.ac.uk