



Unlocking the Potential of Internet Search Data for Health & Social Science Research

# SESSION PROMPTS

12 June 2025 Roundtable | London, UK

# Session 1: Discussion Prompts

- What are the most promising health/social science research questions that can be answered using internet search data?
- What are the main barriers preventing researchers from using this data currently?
- How can internet search data uniquely contribute to improving health outcomes compared to other data sources?
- Why are researchers not using these data more extensively?

# Session 2: Discussion Prompts

- What are the current limitations and capabilities of data portability options in the EU and UK?
- How can Google's data portability and API tools be more effectively used by researchers?
- What are the key ethical and legal concerns associated with using individual internet search data for research?
- What governance frameworks need to be in place to ensure responsible use of donated personal data?

# Session 3: Discussion Prompts

- What are the major challenges in integrating internet search data with other health and social data types?
- What technical solutions are needed to facilitate the secure and efficient linkage of diverse datasets?
- How can we ensure the privacy and security of consumer-held data during the integration process?
- What practical steps can be taken to encourage researchers to utilize consumer-held data, including Google data?

# Session 4: Discussion Prompts

- How can we improve public awareness and understanding of research using internet search data to facilitate informed consent?
- What strategies can be used to effectively recruit and consent participants for studies using internet search data?
- How can we best integrate internet search data into existing population cohort studies?
- What infrastructure, governance is needed?
- What are the most effective study designs (prospective vs. retrospective) for research using internet search data?

# Session 5 Prompts

## **Bridging Data Science, Law, and Public Health**

- What do we not know in health? Could search data help us to fill knowledge voids?
- How can data trusts or similar legal frameworks be adapted to facilitate responsible data sharing and portability for health research using internet search data in the EU and UK?
- From a data science perspective, what are the major challenges in combining internet search data with other health and social data types, and how can legal frameworks support these challenges?
- What innovative funding models or partnerships could incentivize cross-disciplinary research in this area, aligning legal, ethical, and data science needs?
- How do we ensure public awareness campaigns about research using internet search data?

# Session 5 Prompts

## Exploring AI, Epidemiology, and Data Governance

- What existing studies have successfully utilized individual-controlled internet search data, and what were their key findings, particularly where AI has led to significant epidemiological insights or public health improvements?
- How can Google's data portability and API tools be more effectively used by researchers and further optimized to facilitate AI-driven analysis of internet search data while maintaining strong data governance and privacy?
- What best practices or emerging technologies, including technical solutions for secure and efficient linkage of diverse datasets, can ensure the security and integrity of linked datasets analyzed by AI algorithms?
- How can we develop robust, transparent, and ethically sound frameworks for obtaining informed consent for AI-powered research using personal internet search data, and what strategies can be used to effectively recruit and consent participants?
- What specific policy changes or international standards are needed to govern the ethical use of AI in health research involving internet search data and promote data portability and the use of donated personal data for research?
-

# Session 5 Prompts

## **Integrating Social Data, Cancer Informatics, and Research Strategy**

- What are the main barriers preventing researchers from using this data currently, and how can social data science techniques enhance cancer informatics research, especially when combined with internet search data?
- What are the key and unique ethical and legal concerns associated with using individual internet search data for research, particularly when linking social data (including social media and internet searches) with sensitive cancer data?
- How can we ensure the privacy and security of consumer-held data during the integration process, and what data governance models are most effective for protecting privacy while allowing for rich, linked data analysis in cancer research?
- How can we best integrate internet search data into existing population cohort studies, particularly those focused on cancer outcomes and behavior?
- How can individuals with lived experience, especially cancer patients and their support networks, be meaningfully involved in shaping the direction of this research and ensuring research findings are relevant and accessible?

# Session 5 Prompts

## Addressing Data Portability, Health Data Epidemiology, and Clinical Research

- How can internet search data uniquely contribute to improving public health outcomes compared to other data sources, and what are the real-world challenges and successes of data portability initiatives in health research, particularly regarding individual-controlled internet search data?
- What governance frameworks need to be in place to ensure responsible use of donated personal data, and how can we combine data portability solutions with health data epidemiology techniques to analyze large-scale datasets effectively and responsibly?
- What legal and regulatory solutions are most effective for ensuring ethical data sharing and consent in the context of data portability and clinical research, and what practical steps can be taken to encourage researchers to utilize consumer-held data, including Google data?
- How do we balance the need for detailed clinical data with the availability of less structured but potentially insightful internet search data, and what funding opportunities exist or need to be created to support research using individual-controlled internet search data?
- What are the most effective study designs (prospective vs. retrospective) for research using internet search data, and how can data donation initiatives be designed to ethically and effectively contribute to health data epidemiology and clinical research studies?

# Session 5: Prompts

## Examining Internet Geography, Research Challenges, and Funding

- Why are researchers not using these data more extensively, and how can insights from internet geography and web behavior be used to refine health research questions and study designs?
- How can we build public trust in research using internet search data, particularly given concerns about privacy and potential misuse?
- What specific legal barriers or uncertainties are hindering research using internet search data, and how can they be addressed?
- What technical solutions are needed to improve data linkage, portability, and AI-driven analysis of internet search data, and what innovative funding mechanisms can be developed to support interdisciplinary research that combines internet geography, public health, law, and data science?
- What infrastructure and governance models are necessary to support responsible and effective research using internet search data on a large scale?