The Role of Data

Leveraging Artificial Intelligence to Combat Child Sexual Abuse.

CHILDHOOD
WORLD CHILDHOOD FOUNDATION
FOUNDED BY H.M. QUEEN SILVIA OF SWEDEN

Publisher: World Childhood Foundation
Graphic design: Bon Relations
Year: 2025
www.childhood.org

Table of contents

The Role of Data – Leveraging Artificial Intelligence to Combat Child Sexual Abuse.

World Childhood Foundation

World Childhood Foundation works to prevent sexual abuse against children. With knowledge, funding and networks, we empower ideas and innovation that protect children in Sweden and internationally. We support innovative projects, contributing to long-term systemic change, while at the same time improving the lives of individual children here and now.

OUR THEMATIC AREAS

We work in the following three thematic areas: Child supportive relationships and environments, Child safety online and Child focused respons to abuse.

Child supportive relationships and environments

Child safety online

Child focused response to abuse

Within these areas, we focus on where the needs are the greatest and where our expertise and experience can make the biggest difference by:

- **Inspiring and developing new approaches,** and strengthening and disseminating proven methods to help children and families at risk.
- **Contributing to long-term systemic changes** that strengthen children's rights and protection.
- **Initiating, running and supporting strategic actions** with potential, often in partnership with grassroots organizations.
- **Investing in innovative ideas** and helping establish new organizations.
- **Creating and strengthening networks** between initiatives, organizations, and other child rights actors.
- **Shining a light on and investing** in issues and areas that few are talking about, and even fewer are working on.

Our origin

World Childhood Foundation was founded in 1999 by HM Queen Silvia and is a religiously and politically independent, private foundation. Read more about our work at childhood.org

Stella Polaris

The Children's virtual defense force

Childhood's Stella Polaris is a four-year project that aims to coordinate, encourage, and intensify AI-related initiatives to combat child sexual abuse. By bringing together actors in Sweden with different competences, we enable closer interaction between police, prosecutors, and child rights actors on the one hand and AI experts, programmers, researchers, and technology companies on the other. By doing so, we accelerate the development and utilization of useful AI solutions in the fight against child sexual abuse. Stella Polaris is funded by the Swedish Postcode Lottery.

Foreword

In 2019, we at the World Childhood Foundation began exploring how artificial intelligence could accelerate the fight against child sexual abuse. We recognized the emergence of a powerful technological shift and instinctively understood that it, too, would be exploited by perpetrators to find and harm children.

At the time, we couldn't have predicted the speed and scale of the developments that followed —the rise of AI-generated abuse, deepfakes, and nudify apps that have made this technology accessible to anyone, child or adult, for both good and harmful purposes.

What we did know was this: to understand and prevent the ways perpetrators misuse technology, we must first understand and use it ourselves. The challenge, however, is that unlike those who exploit AI to harm children, we —who aim to protect them—must operate within legal, ethical, and organizational boundaries. Sometimes these limitations are concrete. Sometimes they are merely perceived. And sometimes, technology itself can help us overcome them.

That, in essence, is why we chose to commission this report.

Artificial Intelligence begins with data. And yet, for anyone who works with or for children, we know that the data we need is too often incomplete, inaccessible, or invisible.

This report represents our attempt to change that —and the journey has proven far more complex than we first imagined.

It soon became evident how little clarity there is among stake-holders about what data exists, what is needed, or even how to define 'data' in this context. Through the Stella Polaris project, we set out to explore how AI could help prevent child sexual abuse. But two years into our four-year initiative, we realized we couldn't meaningfully discuss AI without first understanding the broader data landscape. That realization led us into the long and complex task of mapping that landscape– a process that took more than two years and was repeatedly slowed by the very challenges it set out to examine.

This report seeks to clarify what types of data exist across relevant sectors and how they might be used to develop or apply AI in efforts to prevent child sexual abuse. It also examines the practical, legal, ethical, and organizational obstacles that hinder data access and sharing

"This report is our contribution to ensuring that children are never among those we fail to see."

Britta Holmberg
Director Global Programs & Advocacy
Deputy Secretary General

—and importantly, distinguishes between those that are genuine and those rooted in assumptions, habits, or fragmentation.

By attempting to map the current landscape, the report provides a starting point for professionals across sectors to better understand what data may be available, how it can be used, and what is needed to make it actionable. It is intended to support more informed dialogue, highlight gaps, and point toward opportunities for collaboration and further research.

In the end, this report turned out to be not just about how data can fuel AI. It's also about how AI can help us navigate the fragmented, sensitive, and often overwhelming world of data related to child protection, including helping stakeholders understand what data may be relevant in the first place. Through anonymization, synthetic generation, and privacy-preserving tools, AI may help us create datasets where none exist, amplify insights where few voices are heard, and uncover patterns where children might otherwise remain invisible.

In particular, we must ask: who is represented in the data? And who is left out?

Children with disabilities, younger children, or those facing systemic exclusion are often missing—not because they do not matter, but because they are harder to reach. If we fail to see these children in our data, we risk failing to protect them in real life. So this is also a report about equity. About ethics. About integrity. It's a call to those working in child rights, research, technology, justice, and health to come together and reimagine the data landscape—not just to train algorithms, but to inform action. To identify risks earlier. To make ethical decisions smarter. And ultimately, to ensure that every child, especially the most vulnerable, is seen, heard, and protected.

Data is never neutral. It tells the story of what we choose to see— and what we choose to ignore. This report is our contribution to ensuring that children are never among those we fail to see.

Britta Holmberg
Director Global Programs & Advocacy
Deputy Secretary General

The Role of Data

Leveraging Artificial Intelligence to Combat Child Sexual Abuse.

Introduction

In 2024, the first global estimated prevalence of child sexual abuse was released, showing that one in five girls and one in seven boys have experienced some form of childhood sexual violence.[1] [2] Until now, reliable global data on this issue has been scarce, highlighting the challenges linked to data on this topic. These challenges grow even more complex with rapid technological change, such as the rise of artificial intelligence (AI) and children's increasing online presence, where prevalence figures are even less certain.

Technology has long been misused by perpetrators to sexually abuse and exploit children. However, the emergence of AI also presents new opportunities to strengthen the prevention of child sexual abuse. This report examines the potential of AI technologies to serve that specific purpose.

Access to high-quality, relevant data is essential for developing effective AI tools, for training models to more complex or tailored reasoning. However, data is also vital in other aspects, underlying decisions, policy and strategies. Yet, across Childhood's work with tech and AI, data access has repeatedly proven difficult. Many partners lack clarity on what qualifies as relevant data, and limited awareness of existing data sources makes cross-sector collaboration even harder.

To address these challenges, we commissioned this report to map available data across relevant sectors and explore how it can be leveraged for AI research, development, and training aimed at combating child sexual abuse.

The report highlights existing data sources and datasets that can inspire innovation, while also examining barriers—legal, operational, and ethical—that may prevent organizations from using them. It further introduces the Swedish data landscape, with the aim of sparking similar efforts internationally.

Furthermore, the report explores technical tools and solutions that allow data to be used without transferring or exposing sensitive information. To make this concrete, it provides examples of relevant data sources and datasets. This report is intended for a wide range of stakeholders, including academia, public authorities, civil society and the private sector.

[1] UNICEF, When Numbers Demand Action: Confronting the Global Scale of Sexual Violence against Children (New York: UNICEF, October 2024) https://data.unicef.org/wp-content/uploads/2024/10/UNICEF_When-Numbers-Demand-Action_Oct_10_2024.pdf accessed 25-05-2025

[2] United Nations Children's Fund, International Classification of Violence against Children (New York: UNICEF, 2023) https://data.unicef.org/topic/child-protection/violence/sexual-violence/#_ftnref1 accessed 25-05-2025

"Data is never neutral. It tells the story of what we choose to see—and what we choose to ignore."

Britta Holmberg
Director Global Programs & Advocacy
Deputy Secretary General

Method

The study focused on five sectors of Swedish society that were identified as either having data sources which could be relevant in combating child sexual abuse. Or, the selected sectors were found to be uniquely equipped to utilize data to a higher degree to increase insight into the issues or strengthen reliability and efficiency of the sector. The sectors analyzed consist of non-governmental organizations, criminal justice, health, academia and the private sector with a special focus on finance and the gaming industry. While there are other sectors such as education, telecom and social platforms holding relevant data for the cause, these sectors were left out due to a combination of them having high level of secrecy on their data management or having their data stored overseas, or simply not being able to partake in the study.

The insights in this report are derived from a mixed methods approach consisting of an online questionnaire distributed to pre-selected participants and focus group discussions with representatives from the sections mentioned earlier. Additional information was contributed by Childhood staff.

QUESTIONNAIRE

A survey was conducted online in October 2023 targeting participants representing diverse expertise within the focus sectors of this report. The questionnaire was supplemented with qualitative interviews and correspondence to ensure a comprehensive understanding. Several recurring responses in the surveys asked for clarifications about the term 'data' and exhibited a reluctance to share general information about dataset characteristics, often citing privacy concerns. Consequently, the questionnaire component was constrained, with the final analysis based on insights from 8 respondents (out of the 27 targeted).

The survey asked questions on the following key components:

1. **Participant Details:** Including names, roles, contact information, and affiliated organization.

2. **Data Source:** General description of the data identified within their sector, categorization, data source information, accessibility, and storage details, including data publishers, owners and contact points.

ARTIFICIAL INTELLIGENCE
Artificial intelligence (AI) refers to computer systems capable of performing tasks that normally require human intelligence, such as learning, reasoning, and problem-solving. These systems can analyze large volumes of data in order to improve logical reasoning, recognizing patterns, and make decisions with speed and scale beyond human capability [3].

3. **Dataset Information:** General dataset description, intended purpose of collected data, format, data size, accessibility, anonymization protocols, associated costs, applications in development using the data, legal framework, requests for data samples, and technical solutions.

4. **Privacy and data sharing practices:** Participants explored tools for protecting tabular data, ensuring secure data transfer, and related solutions. They also completed a multiple-choice section to assess their knowledge of different anonymization techniques.

DEVELOPMENT OF CASE STUDY EXAMPLES

Based on the questionnaire responses, case studies were developed to illustrate five key sectors in focus. The AI tool ChatGPT was used to create a tailored model to analyze the survey data and generate the desired case studies. The AI did not have access to the web or any external documentation beyond the survey responses provided.

To ensure the validity of the AI-generated content, manual review, cross-checking, and refinement were conducted by comparing the survey responses with the draft case studies. The preliminary case studies were then shared with survey respondents for fact-checking and to gather additional information. Respondents were asked to elaborate on key aspects, including:

- **Details about the** format of the data, its volume, and any other relevant characteristics.
- **The main constraints** of sharing and using the data for collaborative purposes.
- **Opportunities for** leveraging AI tools to enhance data usage and sharing.
- **If applicable:** lessons learned from sharing data and insights on improving collaboration.

The case studies were finalized based on the respondents' feedback.

[3] Britannica, "Methods and Goals in AI," Encyclopædia Britannica https://www.britannica.com/technology/artificial-intelligence/Methods-and-goals-in-AI accessed 11-04-2025.

FOCUS GROUP DISCUSSIONS

Twenty participants from five sectors (non-governmental organizations, criminal justice, private sector, academia, and health) took part in three focus group discussions held in Stockholm in April 2024. The purpose of these discussions was to gather reflections on the role of data within their organizations and explore how AI can be used and developed to combat child sexual abuse. This included identifying available data sources and datasets within specific sectors and examining legal, ethical, administrative, and technical barriers to data collection and sharing.

ANALYSIS OF FOCUS GROUP DISCUSSIONS

The discussions were transcribed using the AI tool Klang, with prior consent obtained from all participants. A manual review of the transcriptions was conducted, and the raw text data was compiled into an Excel sheet, categorized based on its relevance to the focus group questions. Additional categorizations highlighted examples of data sharing discussed during the sessions and any potential recommendations that emerged. This manual process ensured human quality control in selecting and organizing the data for further AI analysis.

ChatGPT model 4 and 4.o was then used to analyze the collated data, addressing each focus group question to identify trends and summarize responses. For instance, one prompt used was: "Answer the question: What are the barriers to allowing other actors to access the data mentioned so far? Categorize by legal, administrative, ethical or some other trends that come up. Use this information: [inserted raw text collated in the excel sheet]." The ChatGPT outputs were consolidated into a master document, providing an overall summary analysis of the three focus group discussions. This summary was then manually cross-checked against the Excel text data to ensure accuracy, relevance, and analytical rigor.

"This report represents our attempt to change that—and the journey has proven far more complex than we first imagined."

Britta Holmberg
Director Global Programs & Advocacy, Deputy Secretary General

The effectiveness of AI in combating child sexual abuse will depend not only on its technical capabilities but also on clear ethical guidelines, responsible implementation, and cross-sector collaboration to ensure both child protection and fundamental rights are upheld.

Background

ARTIFICIAL INTELLIGENCE AND CHILD SEXUAL ABUSE IN EU LEGISLATION

AI has emerged as a promising tool in the fight against child sexual abuse. At its core, AI enables machines to learn from data and make conclusions without being explicitly programmed. It can analyze vast amounts of information, including images, videos, and text. If applied on, for example, online platforms, AI can detect patterns indicative of abuse, thus making human moderation more efficient in detecting and reporting illegal activity.

However, the use of AI in such sensitive areas raises ethical, legal, and privacy concerns, prompting regulatory efforts to ensure that it is being deployed responsibly. As AI continues to evolve, finding the right balance between technological advancement and safeguarding human rights remains crucial.

In response to these challenges, the European Union (EU) introduced legislations such as the Digital Services Act and the Artificial Intelligence Act, a legislation designed to regulate AI systems within the EU to ensure they are safe, transparent, and respectful of fundamental rights[4].

[4] European Commission, "Regulatory Framework on AI," Shaping Europe's Digital Future https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai accessed 11-04-2025.

The EU Artificial Intelligence Act and the Digital Services Act:

The EU AI Act, which took effect in August 2024, is the world's first comprehensive law regulating artificial intelligence. It uses a risk-based framework, classifying AI systems into four categories: minimal, limited, high, and unacceptable risk. These rules aim to reduce potential dangers from AI, such as threats to privacy, unfair treatment, or harm to society. At the same time, the law tries to support useful innovations, including those that help protect children. While it recognizes children as especially vulnerable, it doesn't establish child-specific protection. The Act will be phased in, with full enforcement expected by mid-2026, giving organizations time to comply[5].

Further regulating the digital ecosystem, the Digital Services Act (DSA)[6] stipulates the protection of children's online presence. The DSA obliges service providers to mitigate online harms that affect both physical and psychological wellbeing. It outlines the responsibilities of digital services such as social media platforms, search engines, and online marketplaces. In particular, it imposes specific obligations on large online platforms and search engines to protect minors from harmful content, including bullying, illegal material, and misinformation.

Additionally, the law requires platforms to prioritize privacy and safety by design, promoting transparency in content moderation and addressing systemic risks to users' mental and physical health. Measures include features such as age verification, parental control centers, and a complete ban on targeted advertising directed at children. The DSA was implemented in 2022.

[5] European Parliament, EU AI Act: First Regulation on Artificial Intelligence (News, 1 June 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence accessed 11-06-2025.
[6] European Commission, The Digital Services Act (DSA) explained: Measures to protect children and young people online (Publications Office of the European Union, 22 November 2023) https://op.europa.eu/en/publication-detail/-/publication/f3556a65-88ea-11ee-99ba-01aa75ed71a1 accessed 11-06-2025.

To strengthen regional efforts to prevent child sexual abuse, the EU is advancing a proposed regulation focused on improving the identification and reporting of grooming behaviors, as well as both known and previously undetected child sexual abuse material (CSAM). The proposal requires online service providers to assess the risk of their platforms being exploited for child sexual abuse and to implement effective mitigation measures. This is being proposed as the dissemination of child sexual abuse material continues to increase, with more material being generated or augmented by artificial intelligence. In July, the Internet Watch Foundation reported a 400% increase of confirmed reports of AI-augmented child sexual abuse images just in the first six months of 2025[7].

The regulation introduces mandatory steps for a more coordinated and proactive approach to prevention. It obliges certain actors to detect and remove CSAM that has not yet been identified or logged in existing databases. While this may involve the use of artificial intelligence, the regulation remains technology-neutral. An interim regulation has been extended until April 2026 in wait of a permanent legal framework[8].

While AI is not a standalone solution, it is widely recognized to have the potential to significantly enhance our efforts to detect and prevent child sexual abuse, see examples of use cases below (Table 1). By improving early detection, intervention, and investigation, AI can support both victims and professionals working in this field. It is, however, important to incorporate safeguarding principles for children and other risk mitigation principles such as having a Human in the Loop-system.

HUMAN IN THE LOOP:
The principle of Human in the Loop refers to the design of an AI system where human intervention, judgment and input is integral for the AI's output. The human intervention can be integrated through validating an output made by an AI tool, provide domain expertise or feed-back for quality assurance. Having a Human in the Loop approach helps ensure accountability and correctness in decisions which are assisted by the AI tool[9].

[7] Internet Watch Foundation, "IWF urges for 'loophole' to be closed in proposed EU laws criminalising AI child sexual abuse as synthetic videos make 'huge leaps' in sophistication", 11 July 2025, Internet Watch Foundation. Available at: https://www.iwf.org.uk/news-media/news/iwf-urges-for-loophole-to-be-closed-in-proposed-eu-laws-criminalising-ai-child-sexual-abuse-as-synthetic-videos-make-huge-leaps-in-sophistication/ accessed: 12-09-2025.
[8] European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse COM/2022/209 (Brussels, 11 May 2022) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0209 accessed 11-06-2025.

Table 1. Potential benefits of using AI tools to combat child sexual abuse

POTENTIAL BENEFITS	AI USE CASES
Early Detection of Risk Behaviors	AI can be utilized to analyze vast amounts of data from various conversations, social media interactions, and other digital footprints to identify early indicators of risk behaviors. This could include picking up patterns indicating child sexual abuse in text or search results. Early identification enables timely intervention, in some cases even before abuse occurs.
Training and Support for Professionals	AI-driven tools can aid therapists, social workers, law enforcement and other professionals by providing them with data-driven insights and training modules tailored to enhance their skills in supporting victims and identifying offenders. This could include AI simulations and interactive scenarios that help professionals better understand and react to complex situations involving sexual abuse.
Data-driven prioritization in investigations and service provision	As AI excels in managing and structuring large data quantities, it can be applied as a prioritization tool for law enforcement and service providers, helping to identify the most urgent cases.
Expanded First Contact Counselling Services	AI can be trained to help in facilitating safe and supporting conversations for victims or those at risk who might feel too ashamed to speak with adults or authorities directly. These AI chatbots can provide initial support, answer questions, and direct individuals to appropriate human services or helplines.
Enhanced Referral Opportunities	Integrating AI into social media platforms allows for monitoring and analyzing user behavior and content, flagging abusive actions or concerning search terms. This proactive approach can redirect individuals searching for harmful content towards help and rehabilitation services instead of enabling destructive behavior.
Identification of Problematic Sexual	AI can help detect early signs of problematic behaviors in young individuals, acting as an "alarm bell" for caregivers and professionals.
Behaviors in Children and Adolescents	This early intervention can be crucial in preventing the progression to more serious offenses.
Detection of Artificially Generated Abusive Content	By autonomously crawling the web, AI can detect and block distribution of child sexual abuse material and notify law enforcement for takedown. This includes content generated or augmented by AI. Using so-called classifiers, we can use AI to identify novel child sexual abuse material which has not yet been verified and rated by domain experts.
Data-Driven Therapeutic and Prevention Programs	AI can analyze large datasets to determine which interventions are most effective for specific patient profiles or at-risk groups. This can guide the development of targeted prevention programs and therapeutic approaches based on evidence and efficacy.

THE IMPORTANCE OF DATA

One of the biggest challenges facing the child protection field is the need for more relevant data. Data refers to information that can be collected, stored, and processed for various purposes, and can take many different forms, including images, text, or video. Once collected, data might be stored in databases or other structures for retrieval and analysis.

Data can consist of multiple modalities and take the form of text, figures, images, video, audio, or consist of for example tabular information, time series, or geographic details, existing in both structured and unstructured forms. In addition to having a sufficient amount of data, relevance, accuracy and quality are essential in developing reliable AI models.

- **A data source**
  is where the data comes from, essentially the origin of the data. It can be any entity that provides data that can be used for analysis, such as a database, a web API, or real-time data streams.
- **A dataset**
  is a more refined and organized collection of data, to be prepared or ready for analysis, often extracted from one or multiple data sources. A dataset could be the contents of a database table, the contents of a CSV file, or data collected from multiple sources and aggregated into a single table.
- **Data quality**
  refers to how accurate, consistent, and reliable the data is.
- **Data quantity**
  refers to the amount of data.
- **Training data**
  refers to data used to teach machine learning models how to recognize patterns and draw conclusions.
- **Data cleaning**
  is the process of correcting inconsistencies in data and removing irrelevant information - also called “noise”- from a dataset.

[9] Bahuguna, Anuj, “AI in the Loop vs Human in the Loop: A Technical Analysis of Hybrid Intelligence Systems” (IBM Community, 25 May 2025) https://community.ibm.com/community/user/blogs/anuj-bahuguna/2025/05/25/ai-in-the-loop-vs-human-in-the-loop accessed 06-05-2025.

It is important to consider potential biases in data and their effects on AI outputs. Bias in data occurs when the data reflects unfair, incomplete, or unbalanced information, often mirroring societal prejudices or stereotypes. Bias can be a result of who collects and owns the data, and how the data is used, with marginalized communities and groups including vulnerable children often being the ones missing from the data. This can lead AI models to produce discriminatory, inaccurate, or misleading results. To ensure AI systems contribute to ethical and equitable decision-making, it is necessary to mitigate bias through diverse and representative datasets, fairness audits, and continuous monitoring.

Data from children is in general more difficult to obtain than from adults, as stricter legal protection comes into play. Using data from children also requires careful consideration, weighing the potential benefits of the children's right to privacy and participation. There is a lot of data collected within the private sector, and for various uses. While some actors collect data primarily for surveillance or commercial gain, there are also significant opportunities to make greater use of private sector data to support efforts to prevent abuse, identify victims, and assess the broader impact of child sexual abuse—without necessarily relying on children's personal data.

One way to increase security in sharing sensitive data is by applying privacy-preserving techniques like encryption, anonymization, and pseudonymization. These methods offer a way to protect individual privacy while enabling collaboration and data sharing. When implemented correctly, they help overcome legal and ethical barriers to data sharing by ensuring that personal information is either protected, rendered unidentifiable or that the data cannot be accessed by unauthorized third party.

- **Encryption:** A process that transforms data into a coded format that can only be read by someone with the correct decryption key. It protects data during transmission or storage from unauthorized access. This technique is largely used in ensuring secure online communications but can be applied when sharing data to ensure it is not intercepted along the way.
- **Anonymization:** A process that removes or alters sensitive information, such as personal identifiers in a dataset so that individuals cannot be identified, directly or indirectly. Once data is anonymized, it cannot be traced back to a person, even by a data controller.
- **Pseudonymization:** A method that replaces personal identifiers with fake identifiers or pseudonyms. It allows data to be linked to the same individual across different datasets without revealing their identity.

Data from children is in general more difficult to obtain than from adults, as stricter legal protection comes into play.

The link between the pseudonym and the actual identity is kept separately and securely.

Privacy by design is a proactive approach to building systems and technologies where privacy is integrated from the outset, rather than added as an afterthought. In the context of AI, it means embedding data protection principles throughout the entire lifecycle of model development—from data collection and preprocessing to training, deployment, and monitoring.

Implementing privacy by design can support fair and reliable data management and AI development by minimizing unnecessary data use, applying techniques like anonymization, differential privacy, and federated learning, and ensuring transparency in data handling. This reduces risks of bias being amplified by AI, it protects individuals from misuse of sensitive information and fosters trust among users. By making privacy a default setting, organizations not only comply with regulations but also create AI systems that are more equitable, accountable, and resilient[10]. You can read more about these privacy-enhancing techniques in Table 6.

To establish the scope and the scale of child sexual abuse, updated and relevant data is crucial for development and monitoring of interventions, and to inform policy and decision making. Data can highlight areas in need of funding and can also be used to generate political will, catalyze action, and mobilize resources for evidence-based investments[11].

[10] Integritetsskyddsmyndigheten (IMY), "Privacy by design and privacy by default", updated 16 April 2021, https://www.imy.se/en/organisations/data-protection/this-applies-accordning-to-gdpr/privacy-by-design-and-privacy-by-default/ accessed 12-09-2025.

[11] Webinar Recap: Data for Change (2023). InHope and UNICEF. Available at: https://inhope.org/EN/articles/webinar-recap-data-for-change-listening-to-the-voices-of-children-about-their-experiences-online accessed 17-06-2024.

The data landscape in Sweden

Swedish society has systematically worked to improve children's health and well-being since the early twentieth century and is internationally recognized as a leading example in combating violence against children[12]. Sweden also has a long tradition of maintaining population and national data registers, containing information on areas ranging from public health to court cases, which are used for statistics and research (see Appendix 1.4 for more information). Ongoing efforts are being made to increase data accessibility and sharing across sectors and organizations in Sweden. One example is AI Sweden, which supports multiple initiatives aimed at improving data sharing and even provides member organizations with access to datasets for testing and research[13].

AI Sweden[14], the National Center for Applied AI, has been tasked with accelerating the adoption of artificial intelligence in Sweden. To achieve this, it collaborates with more than 130 partners from the private sector, the public sector, and academia. Its work focuses on maximizing organizations' use of AI, supporting researchers and infrastructure to enable bold innovations, and building strong AI communities both within Sweden and internationally.

[12] Laura Korhonen, Linnea Lindholm, Maria Lindersson and Ann-Charlotte Munger, The Inclusion of Children in Public Enquiries on Violence, Health and Welfare: The Example of Sweden in Roth, M., Alfandari, R. and Crous, G. Participatory Research on Child Maltreatment with Children and Adult Survivors: Emerald Studies in Child Centered Practice. (Emerald Publishing, 2023), pp. 197–213.
[13] AI Sweden. Dataset, https://www.ai.se/sv/ai-labs/technology-infrastructure/dataset accessed 12-09-2025.
[14] AI Sweden, AI Sweden – National center for applied AI https://www.ai.se/en accessed 14-06-2025.
[15] AI Sweden, My AI (AI Sweden) https://my.ai.se/ accessed 11-09-2025.

EXAMPLE 4: NETWORKING AND COLLABORATION WITH THE SWEDISH AI COMMUNITY

"My AI"[15], run by AI Sweden, is a personalized platform for the Swedish AI community—similar to a "LinkedIn for AI." It is designed to inspire, share knowledge, facilitate networking, and showcase the work of organizations. The platform is built on key principles such as personalization, openness, connectivity, datadriven insights, and support for both many-to-many and one-to-one interactions.

Its purpose is to strengthen collaboration and data sharing by connecting AI experts and organizations, enabling the exchange of relevant insights and technologies, and fostering a united effort to develop effective AI-driven solutions.

Personal data processing in Sweden

Personal data refers to any information that identifies or can identify a natural person. Any action involving personal data constitutes processing-including collection, structuring, storage, modification or destruction. Examples of personal data include a person's identity number, name, and address. Images and sound recordings of individuals processed by a computer may also constitute personal data, even if no names are mentioned. Encrypted data and electronic identifiers, such as IP addresses and cookies, are considered personal data if they can be linked to individuals. Similarly, encoded, encrypted, or pseudonymized information that can be associated with a natural person using additional data is classified as personal data. The management of personal data is governed by the General Data Protection Regulation (GDPR).

For children under 18, stricter regulations apply for the handling of personal data. Most actions require parental consent. The collection and use of their data must be individually assessed to balance the benefit of access to services with the sensitivity of their data.

[16] European Parliament, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4 May 2016, 1–88 https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng accessed 11-09-2025.

The General Data Protection Regulation (GDPR) applies to all sectors in the EU, including businesses, associations, organizations, authorities, and private individuals, encompassing nearly all operations and activities. It ensures the lawful processing of personal data and informs individuals about how their data is managed. GDPR covers all automated personal data processing and certain manual processes. Compliance is overseen by the Swedish Authority for Privacy Protection (IMY)[16].

GENERAL OVERVIEW OF RELEVANT DATA SOURCES

During the focus group discussions, the conversations extended beyond the scope of each representative's organization, highlighting several sources across the Swedish data landscape which might hold relevance for prevention of child sexual abuse. General examples of significant data identified by the focus groups were:

1. **Annual or recurring Surveys on Child Experiences:** These surveys can provide detailed insights into the prevalence, nature, and aftermath of child sexual abuse and violence. Analyzing trends and patterns from these surveys can help identify risk factors and develop preventive measures.
2. **Financial Transaction Data:** Monitoring transactions, especially those involving payments to certain high-risk countries or flagged accounts, can be instrumental in identifying commercial sexual exploitation and abuse of children. Scrutinizing customer behaviors based on age, payment patterns, VPN usage, and suspicious activity reports can detect potential abusers and prevent transactions related to abuse. Similar to anti-money laundering methodologies, triaging multiple risk indicators in something called "link analysis" can uncover suspicious financial activities indicative of abuse-related transactions.
3. **Public Registers and Agencies:** Data from public registers can be used to cross-reference and verify information in child abuse cases.
4. **School Data:** Information from schools—such as absenteeism, behavioral changes, and reports from counselors—can provide early warning signs of abuse.
5. **Research Data and Peer-Reviewed Articles:** Datasets underlying scholarly articles often contain valuable, anonymized information. With ethical clearance, repurposing this data can expand knowledge bases and foster interdisciplinary research to prevent abuse.
6. **Institutional Documentation:** Records from child welfare institutions, including staff interviews and meeting notes, might reveal organizational cultures that either prevent or inadvertently allow abuse to persist.
7. **Chat Logs:** With consent and careful and ethical considerations, analyzing chat logs could identify grooming behaviors or potential threats. AI tools already used in law enforcement can be adapted to efficiently monitor and analyze large volumes of chat data.
8. **Health and Legal Databases:** Entities like the Center for Health Data (Centrum för Hälsodata) hold extensive health records that can be accessed by researchers under strict ethical guidelines. Similarly, forensic data from legal medical institutions can provide insights into the physical and psychological impacts of abuse, aiding in both investigation and prevention.

EXAMPLE 1: LEARNING FROM THE SWEDISH DATA PORTAL

Dataportalen[17] serves as Sweden's national data portal, acting as a central hub for accessing various public digital resources aimed at fostering development and innovation. It provides a comprehensive gateway for public and private organizations to share data, as well as for data producers to find support for data sharing. This makes it a "portal of portals" by linking to different sources of data and resources, rather than hosting the data itself. It is offered by the Swedish Agency for Digital Government (DIGG)[18] who support the development of digital solutions for the Swedish public sector.

Application to AI and Data in Combating Child Sexual Abuse: Using a model similar to Dataportalen, a specialized data portal could be developed to combat child sexual abuse by linking various datasets, tools, and resources related to the issue. Such a portal could facilitate the sharing of best practices, innovative AI tools, and research between organizations and researchers. Additionally, the portal could provide guidelines on how to annotate and manage metadata and code, which is crucial for developing effective AI models and ensuring the integrity and usability of the data used in combating child sexual abuse. Alternatively, the application of AI techniques to existing data within DataPortalen, such as unique tagging and use of synthetic data, could enhance the mining of existing data to combat child sexual abuse.

[17] Myndigheten för digital förvaltning (DIGG), Sveriges dataportal https://www.dataportal.se/ accessed 11-09-2025.
[18] Myndigheten för digital förvaltning (DIGG), Digg – Myndigheten för digital förvaltning https://www.digg.se/ accessed 11-09-2025.

BOX 1: FURTHER USE CASES FOR AI APPLICATION IN EXPLORING, ENHANCING AND STREAMLINING HOW TO PREVENT CHILD SEXUAL ABUSE

Our analysis highlights a range of AI techniques that hold significant potential for broader application—or more intensive use—across Sweden's public and private sectors. These tools can play a vital role in supporting efforts to combat child sexual abuse, particularly in the areas of data handling, analysis, and secure information sharing.

Privacy by design: This proactive approach integrates privacy and data protection principles directly into system and process architectures from the beginning. Rather than being an afterthought, privacy is built in by default —through measures like data minimization which ensures only essential data is being used, encryption, and stringent access controls. This methodology is critical when handling and sharing sensitive data, ensuring the protection of individuals' information, safeguarding confidentiality, and maintaining compliance with legal and regulatory standards throughout the data lifecycle.

Chat Data Aggregation and Analysis: AI-powered tools can aggregate and analyze communication data from diverse platforms to identify potential signs of abuse or exploitation. Using for example natural language processing (NLP), these systems can detect emotional tones, subtle cues, or linguistic patterns that may indicate harmful behavior or signs of victim distress—providing early warnings that could trigger intervention. This technique has been tested in multiple efforts to identify early indicators of grooming on internet platforms.

Mapping and Gap Analysis in Reports: AI can streamline the review and synthesis of existing research and reports to map out current knowledge and identify gaps. This process can inform policymakers and researchers by highlighting underexplored areas, ensuring that future efforts are focused where they are most needed.

Identification and Segmentation Tools: AI can assist in the secure identification and segmentation of service users, helping to pinpoint individuals who may be particularly vulnerable or at risk of abusing others. By analyzing usage patterns, behaviors, or interactions, especially in healthcare or social services contexts, AI tools can be a tool for developing early detection and intervention strategies.

Anonymization Tools: AI can aid in creating secure environments for sensitive data collection, such as in interviews with minors. Anonymization tools ensure that data can be used for analysis without compromising the identity of the individuals involved, minimizing risks in case of data breach and encouraging more honest feedback and participation from the person whose data is being managed.

Financial Transactions Analysis: AI can analyze financial transactions to detect patterns indicative of illegal activities such as sexual exploitation. This analysis could extend beyond direct transactions to include patterns in suspect and victim bank statements. Such insights can help uncover criminal networks and behavioral patterns associated with sexual abuse. This is being used in financial crime prevention efforts.

Integrative Data Analysis Across Multiple Entities: AI enables the integration and holistic analysis of datasets from multiple sources. When data is viewed collectively rather than in isolation, new patterns and correlations can emerge—providing deeper insights and revealing more effective strategies for preventing and responding to child sexual abuse.

Analysis

The following chapters builds on what we have learned about the data landscape in Sweden, and takeaways from discussions with focus group participants and survey responses. A more extensive deep-dive into data opportunities within each sector can be found in the appendicies.

So far in this report, we have explored a wide range of data types and sources. We have also shared how data comes in different forms, such as text, numbers and photos, and that some sectors have more or less of these different types. Importantly, we highlight how much of this data from different sectors is crucial in the fight against child sexual abuse. However, it is worth noting that a significant amount of valuable data remains unused.

The focus groups noted the differences in data quality and quantities—like billions of financial records, thousands of survey answers, and individual health records. They also examined the various barriers that exist in Sweden to enabling effective data sharing and access with the purpose of combating child sexual abuse.

For AI tools to be effective in preventing and responding to child sexual abuse, the data they learn from must be both abundant and high-quality. Quantity enables machine learning models to recognize complex patterns, while quality ensures those patterns reflect reality accurately and ethically.

DATA QUANTITY

Some domains offer rich opportunities for data collection. For example, online communications—such as public chat logs, messaging platforms, and forums—can be analyzed to detect grooming behaviors, patterns of coercion, or signs of online enticement. Similarly, case records from child protection services, court decisions, and anonymized helpline transcripts provide structured, incident-level data that can support training of AI systems.

However, there are also critical data deserts where information is scarce, underutilized, or structurally inaccessible. Among others, these include:

- **Youth Sports and Extra-curricular Activities:** Data on children's interactions with coaches, volunteers, and peers—contexts where abuse of power can occur—is rarely collected in a systematic or analyzable way.
- **Pre-Teen Digital Behavior:** Because children under 13 are officially excluded from many platforms (due to age restrictions), there is a lack of visibility into how they interact online, leading to underrepresentation in risk modeling and safety interventions.
- **Community and Faith-Based Settings:** Information about children's participation in religious institutions, informal education programs, or cultural groups is often decentralized or anecdotal, making AI analysis nearly impossible.
- **Chat Metadata Limitations:** Even when chat data is available, the surrounding metadata (timestamps, user info, geolocation, platform type) is often inconsistently captured or poorly formatted, reducing its analytical value.

DATA QUALITY

In terms of quality, academic research and criminal justice records are often held in high regard due to their rigorous documentation and review processes. Peer-reviewed studies, government inquiries, and official reports typically offer well-structured, transparent, and validated data sources. These kinds of datasets help ensure that AI models make sound inferences and avoid harmful bias.

Yet, Sweden—like many countries—faces significant challenges when it comes to ensuring consistency, clarity, and ethical standards across datasets:

- **Institutional Variability:** Data quality can vary greatly between municipalities, agencies, or service providers, leading to issues when datasets are combined or compared.
- **Lack of Standardization:** Differences in how data fields are labeled, for example in how abuse types are coded, or how outcomes are tracked can undermine interoperability and lead to unrepresentative conclusions.
- **Inconsistent Redaction and Privacy Practices:** Efforts to anonymize or redact data to protect children's identities are crucial—but when done inconsistently, they can inadvertently obscure important patterns or introduce bias. For example, unnecessary redaction may erase context necessary for understanding abuse dynamics, while not sufficient redaction risks exposing identifiable information.

Barriers for data access and sharing

Focus group discussions revealed several barriers (Table 5) that pose significant challenges to data sharing across sectors. These obstacles limit opportunities for collaborative research, informed policy-making, and the enhancement of services. The complex interplay of legal, ethical, and operational constraints demands careful navigation to improve data accessibility, while also safeguarding individual rights and maintaining public trust. Later in the report, we present potential solutions to addressing the identified barriers.

Table 5. Barriers in data access and sharing in Sweden

	TYPE OF BARRIER	DISCRIPTION OF BARRIER
Legal Barriers	Data Protection and Privacy Laws	Strict data protection regulations, such as GDPR, govern how personal data is collected, stored, and shared. In Sweden, legal interpretations tend to be conservative to minimize the risk of penalties, which restricts data sharing and aggregation, particularly in sensitive sectors like healthcare and child protection.
	Sector-specific Privacy Laws	Certain professions, particularly within healthcare and municipal services, are subject to specific legal restrictions (sekretess) that prevent data sharing, even when the data has been de-identified. These sector-specific rules further limit access to data across different sectors.
	Legal Interpretation and Fear of Repercussion	While many organizations adopt cautious data-handling practices due to unclear legal guidelines, there is a risk that a reflexive avoidance of uncertain legal territory may prevent them from even trying to explore potential opportunities. The fear of legal repercussions—combined with the complexity and time-consuming nature of obtaining legal clarity—serves as a significant barrier to data sharing.
Operational Barriers	Fragmented Systems	Incompatible systems and labeling of data lead to difficulties in data sharing and aggregation. For example, between various online chat support services using different systems, or between medical care providers in different cities or regions. Difficulties encountered are compounded when sharing data between different sectors.
	Inconsistent Data Handling	Lack of standardization in data handling and labeling across different providers and sectors complicates data sharing and aggregation.
	Lack of Inter-agency Collaboration	There are significant challenges in coordinating between different public agencies and between the public, private, and academic sectors due to different administrative practices and data handling protocols.
	Permit and Regulation Compliance	Numerous permits and regulations required at various administrative levels impede the efficient use of data for research and operational purposes.
	Resource Limitations	Non-profits and other organizations often lack the financial resources to invest in necessary technology and expertise for proper data management and sharing.
	Encryption needs	While data encryption is essential for safeguarding privacy, it also presents significant challenges to data accessibility and analysis. For some strong encryptions, the decrypting process is time-consuming and requires increased computational power. Striking a balance between strong encryption and minimizing bottlenecks is a challenge for sectors that rely on swift data processing.
Ethical Barriers	Sensitive Nature of Data	Data related to topics such as child sexual abuse is particularly sensitive, carrying a high degree of stigma and requiring careful ethical consideration. This sensitivity leads to restrictive data-sharing practices aimed at protecting the dignity and privacy of affected individuals.
	Consent Concerns	Even when data is anonymized and consent is obtained, there often remains a reluctance to share it due to ethical considerations around potential misuse.
	Patient Confidentiality in Healthcare	Maintaining patient confidentiality is a core principle in healthcare, which poses a significant barrier to data sharing—even when such sharing could improve patient outcomes or support research initiatives.

While data encryption is essential for safeguarding privacy, it also presents significant challenges to data accessibility and analysis. Encrypted data is not readily available for analytical use, which can impede operations that require swift data processing—such as those in the criminal justice system.

Even when decryption is legally authorized, the process often demands substantial computational resources and time, resulting in delays to critical investigations and decision-making. Striking a balance between strong encryption and practical data usability remains a central challenge for sectors that rely on timely data analysis.

Using AI to Combat Child Sexual Abuse

While many barriers to data sharing have been identified within the Swedish context, there are also numerous AI-driven techniques and solutions that can enable secure and collaborative, data-based action to combat child sexual abuse. Practical and effective approaches exist for using and sharing sensitive data responsibly. These range from implementing "privacy by design" principles to adopting advanced technical methods that allow insights to be drawn from data without moving, disclosing, or sharing the actual datasets (see Box 1).

For instance, federated learning enables AI models to be trained across decentralized datasets without the data ever leaving its source. This approach allows for the development of robust, fine-tuned models while maintaining data confidentiality. Similarly, blockchain technologies offer secure, decentralized platforms for sharing sensitive information—such as health-related interventions—among stakeholders with transparency and trust.

While AI systems, particularly those based on machine learning, typically require large volumes of data to perform well, there are techniques designed to mitigate this challenge. One such technique is the use of synthetic data—artificially generated datasets that closely mimic real-world data. This method allows for effective model training in situations where there is data scarcity, or access to actual data is restricted due to privacy concerns.
Each of these approaches enhances AI's capacity to analyze data, detect early warning signs, and potentially prevent harmful behaviors. However, deploying AI in such sensitive contexts must be accompanied by a strong ethical framework. Privacy safeguards, transparency, and a keen awareness of potential unintended consequences are essential to ensure that these technologies are used responsibly and effectively.

With the EU AI Act formally adopted, Europe has entered a new era of AI regulation aimed at ensuring ethical, transparent, and safe use of artificial intelligence across sectors. This legislation has significant implications for how AI can be applied in the fight against child sexual abuse. By imposing strict requirements on high-risk AI systems—including transparency, human oversight, and data quality—the AI Act aims to enhance the safety, reliability, and ethical integrity of these technologies.

In the context of child protection, these regulations may improve trust in AI tools and help ensure they are used responsibly. However, the new compliance obligations also introduce practical challenges. Organizations developing or deploying AI for detecting and preventing child sexual abuse may face increased costs, technical demands, and administrative hurdles. These could potentially slow down innovation or hinder the adoption of AI tools in time-sensitive scenarios.

Striking a balance between regulatory safeguards and the urgent need for effective technological solutions will be key. It will require collaborative efforts among policymakers, developers, civil society, and child protection experts to ensure that AI tools remain both compliant and impactful in protecting children.

With the EU AI Act formally adopted, Europe has entered a new era of AI regulation aimed at ensuring ethical, transparent, and safe use of artificial intelligence across sectors.

USING AI TECHNIQUES TO OVERCOME IDENTIFIED BARRIERS

Overcoming the barriers identified in this study requires a combination of advanced technical and strategic AI approaches that emphasize ethical standards and privacy protection. A carefully designed strategy is essential—one that incorporates a legal awareness and ensures strong technological implementation.
To maintain compliance and operational effectiveness, organizations must regularly update these systems in line with evolving legal frameworks and technological progress. Several technical solutions have been identified to address the specific challenges of accessing and sharing data in Sweden (see Table 6).

Table 6. Technical AI Solutions To Overcome Data Access And Sharing challenges

	TYPE OF BARRIER	POTENTIAL AI SOLUTIONS
Legal Barriers	Data Protection and Privacy Laws	**Homomorphic Encryption:** A way to do computations on encrypted data without ever needing to decrypt it. **Differential Privacy:** A technique that adds controlled randomness to data so individual information cannot be identified. **Zero-Knowledge Proofs (ZKPs):** A method that lets someone prove they know or have something without revealing the actual information.
	Sector-specific Privacy Laws	**Data Anonymization and Pseudonymization:** Techniques that remove or replace personal details in data to protect people's identities. **Federated Learning:** A system where AI models are trained across multiple devices or locations without moving the raw data.
	Ambiguity in Legal Regulations	**AI-based Legal Advisors and LLM-powered policy summarization tools:** AI tools that help interpret complex laws or summarize policy documents in plain language. **Automated Compliance Systems:** Software that checks whether data use follows laws and regulations automatically.
	Legal Interpretation and Fear of Repercussion	**AI-driven Simulation and Risk Assessment Tools:** AI systems that test different scenarios to predict potential risks before they happen. **AI-driven Continuous Compliance Monitoring:** AI that constantly checks data processes to make sure they stay within legal and ethical rules.
Operational Barriers	Fragmented Systems	**Data Integration Platforms:** Tools that combine data from different sources into one consistent system. **Ontology-based Systems:** Structures that define and organize terms so data from different systems can be understood the same way. **Knowledge Graphs:** Databases that connect information through relationships between concepts and entities.
	Inconsistent Data Handling	**Machine Learning Models for Data Standardization:** Algorithms that clean and reformat data into a common structure. **Natural Language Processing (NLP):** AI that enables computers to understand, process, and generate human language.
	Lack of Inter-agency Collaboration	**Collaborative AI Platforms:** Shared systems where multiple groups can work together on AI projects securely. **AI-powered Workflow Automation:** AI that performs repetitive tasks in processes without needing human input. **Federated Analytics:** A way to analyze data stored in different locations without combining it into one place, similar to federated learning.
	Permit and Regulation Compliance	**AI-designed Regulatory Compliance Tools:** AI programs that help design and implement systems aligned with legal standards. **AI-automated Document Analysis and Management Systems:** Tools that use AI to read, sort, and organize large sets of documents. **Explainable AI:** AI systems designed to show how and why they make decisions in a clear way.
	Resource Limitations	**AI as a Service (AIaaS):** Cloud-based platforms that let people use AI tools without building them from scratch. **Open Source AI Tools:** AI software that is freely available for anyone to use, modify, and share.
Ethical Barriers	Sensitive Nature of Data	**Secure Multi-party Computation (SMPC):** A method where multiple parties can work together on data without revealing their own inputs. **Synthetic Data Generation:** Creating artificial datasets that mimic real data while protecting sensitive information. **Trusted Execution Environments (TEEs):** Secure parts of computer hardware where sensitive data can be processed safely.
	Consent Concerns	**Blockchain for Data Tracking and Audit Trails:** A system that records and verifies data transactions in a tamper-proof way. **Explainable Consent Management:** Tools that make it clear to users how their data will be used and allow them to give or withdraw permission. **Dynamic Consent Platforms:** Systems that let people update and manage their data-sharing preferences at any time.
	Confidentiality	**Federated Learning and Analytics:** Combining training of AI models and data analysis across different sources without centralizing the raw data. **Differential Privacy:** A privacy technique that hides individual data points by adding noise while keeping overall patterns useful

Building on lessons learned

Several key lessons about data sharing and collaboration emerged during the focus group discussions, particularly related to research and regulatory environments and overarching suggestions for initiatives and future developments. To effectively use AI techniques in combating and preventing child sexual abuse, organizations must navigate a complex landscape of ethical considerations, data privacy regulations, and technological challenges. Here we present some tailored recommendations based on the focus group discussions:

1. **Utilization of Research Datasets:** Allowing researchers who have previously received funding to apply to use the same data again underscores the value of iterative research. This approach not only extends the utility of the data but also encourages the reuse of existing data, optimizing research investments and uncovering new opportunities for discovery.
2. **Regulatory Sandboxes**[19]**:** The concept of regulatory sandboxes highlights the importance of flexible regulatory frameworks that adapt over time. These frameworks facilitate innovation by allowing temporary relaxation of regulations to test new ideas while still ensuring compliance. This method supports sustainable innovation within legal boundaries, encouraging stakeholders to interpret and engage with regulations creatively but responsibly.
3. **Collaboration:** The Swedish culture stresses the importance of exploring diverse solutions to problems collectively. This approach fosters a collaborative mindset, crucial for organizational and technological breakthroughs. Sharing challenges and solutions in a communal setting can lead to unexpected insights and stronger partnerships. Furthermore, AI can facilitate better collaboration across sectors by standardizing data formats and sharing protocols, similar to the Safer Internet Centers[20] in the EU which integrate resources aross multiple platforms for compre-hensive analysis and response.
4. **Daring to be Brave:** Successful cross-sectoral initiatives often begin with a few stakeholders advocating for data sharing. leading to broader collaboration. It often requires a leader, or an organization, to take the first brave step to then impact leveraging collective resources and capabilities.
5. **Understanding and Interpreting the Law:** Understanding the legal implications of data sharing is crucial. This includes knowing what data you have, its sensitivity, and ensuring compliance with relevant laws and regulations. Understanding the legal ground-work is essential to enable you to make brave decisions while safeguarding privacy and fulfilling ethical obligations.
6. **Open Access and Data Sharing:** Work towards dismantling barriers that prevent the sharing of data across different regions, cities, and sectors. Promote the development of open public registers and databases that are accessible to all stakeholders, thus enhancing the potential for effective AI-driven trend analysis and intervention strategies.
7. **Standardization of Metadata Structures:** Establish a uniform metadata tagging framework across all platforms and institu-tions that deal with data related to child protection. This standar-dization will facilitate the aggre-gation, mapping, and analysis of data, making it more actionable for AI applications.
8. **AI Education and Awareness:** Increase general knowledge of AI capabilities and ethical implications across all levels of organizations—from top management to frontline workers like youth coaches. Comprehensive understanding and acceptance of AI technologies are crucial for their successful implementation.
9. **Develop Clear Risk Indicators for the Financial Sector:** Provide clear and accessible risk indicators for the financial where non-experts need to be able to flag suspicious activities related to funding or transactions that could be linked to child exploitation.
10. **Utilization of Common Technical Solutions and Data Protocols in Swedish Governmental and Legal Bodies:** Encourage the adoption of common software systems and databases among local municipalities and institutions. This approach not only enhances data compatibility and ease of integration but also supports a more unified response to child protection issues. Facilitate better understanding and cooperation among various governmental and legal bodies to harmonize the data they collect and share. This includes reconciling different legal and data management standards and privacy regulations to enable more effective data sharing and joint initiatives.

[19] "A regulatory sandbox is a tool that can be used to give actors the opportunity to test and experiment with, for example, new innovative products or services that would otherwise have been prohibited because they conflict with laws or other regulatory frameworks that actors have committed to follow." RISE, Regulatoriskt växthus, sandlåda eller försöksverksamhet (RISE) https://www.ri.se/sv/expertisomraden/expertiser/regulatoriskt-vaxthus accessed 14-04-2025

[20] European Commission, "Safer Internet Centres," Shaping Europe's Digital Future https://digital-strategy.ec.europa.eu/en/policies/safer-internet-centres accessed 17-06-2025.

Understanding the legal groundwork is essential to enable you to make brave decisions while safeguarding privacy and fulfilling ethical obligations.

Areas for further exploration

Focus group participants clearly indicated a need for improved collaboration and better implemented technical solutions in order to work more data-driven. Multiple groups also underscored the need for clear legal guidelines to ease in navigating data and AI management, as legal ambiguity often hinders the testing of new approaches or tools. As actors working "within legal boundaries face stricter ethical and legal constraints than perpetrators exploiting technology, participants expressed a need to distinguish between real barriers and merely perceived ones.
Without this clarity, the concern was that innovation risks being stifled by an unnecessary sense of self-limitation.

There were some outstanding questions (Table 7) reflecting a broader concern about how to best balance the need for maximizing the use of data with the requirements for privacy, security, and legal compliance in highly sensitive areas.

In conclusion, this report underscores the vital role of collecting, utilizing, and sharing data to advance the use of AI in combating child sexual abuse. Improved access to high-quality data empowers AI to more effectively analyze public records, revealing patterns and trends in abuse cases that might otherwise remain hidden.

Table 7. Summary compilation of questions for further exploration

OUTSTANDING QUESTIONS	
Data Utilization	• How can different types of data be integrated and analyzed together to generate deeper insights into child sexual abuse prevention?
Data Quality	• How can we improve the quality of existing data that is used in studies and analysis? • What standards should be established for data labeling and metadata to ensure consistency and usability across different platforms and research efforts? • What strategies can be employed to manage and utilize large volumes of data effectively, especially when dealing with encrypted or low-quality data?
Legal Challenges	• How can legal barriers, such as privacy laws and data protection regulations, be navigated to facilitate better collaboration and data access? • What are the legal boundaries for sharing sensitive data between child protection actors and law enforcement?
Data Sharing	• Are there standardized procedures that can facilitate easier data sharing across agencies and countries?
AI Solutions	• How can federated learning be effectively implemented to protect data privacy while still allowing for comprehensive data analysis? • What is the potential of synthetic data in training AI models, and how can it be utilized without accessing actual sensitive data? • How can we distinguish AI-Generated content[21] from other content, especially in augmented images? • What are the ethical implications of using AI to identify risk behaviors and trends in sensitive areas? • What are the ethical implications of not using AI to identify risk behaviors and trends in sensitive areas?
Funding	• How can non-profits and other organizations ensure sustainable financing for data-intensive initiatives?

[21] Internet Watch Foundation, How AI Is Being Abused to Create Child Sexual Abuse Imagery (Research Report, IWF) https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/ accessed 10-03-2025.

Sweden is well-positioned to lead global efforts in leveraging data for child protection. Continued investment in AI development and implementation is critical, particularly when paired with strong collaboration between AI experts, social services, and the private sector. Such cooperation is essential to harnessing technology's full potential to safeguard children. By continuing efforts to support cross-sectoral sharing and collaboration around data, we can strengthen efforts to prevent child sexual abuse.

This report calls for the integration of advanced AI technologies into both preventive and responsive child protection strategies. By increasing data literacy, ensuring secure data access, and expanding the use of AI-driven solutions, Sweden can further strengthen its child protection efforts. In doing so, it can also serve as a global model, demonstrating how data and technology can be responsibly used to create safer environments for children everywhere.

Appendices

The following appendices provide detailed deep-dives into the selected sectors analyzed in this report. Each chapter offers insights into the nature of the data, its sources, and the sector-specific context in which it is utilized. The findings presented in the appendicies are primarily identified by the focus groups and gathered from surveys to relevant stakeholders in each sector.

NON-GOVERNMENTAL SECTOR

Non-governmental organizations (NGOs) working with children play a critical role in preventing and responding to child sexual abuse through advocacy, education, direct support services. They are often able to reach children through different avenues than for example, school and health services. Many organizations have extensive access to data such as surveys and reports and have a high capital of trust from the public.

NGOS IN SWEDEN

While an extensive public sector takes overall responsibility for social services and welfare provision, NGOs are increasingly participating as service providers. NGOs play a vital societal role by often compensating for governmental and municipal deficiencies. By creating new programs and models for support, they also can test innovative solutions to existing and new challenges. Additionally, they are central in holding authorities accountable and ensuring that children's voices are heard in policymaking. However, a significant barrier within the NGO sector is the lack of funding and a comprehensive network structure. While there are some well-established collaborative initiatives, there is often a lack of strong coordination and communication between the different organizations.

The NGO sector in Sweden is crucial for strengthening the rights and protection of children and young people, often reaching the most vulnerable groups of society. NGOs offer a wide range of services and support both online and offline, for example:

- **Psychological and emotional support** – Through support groups, counseling and helplines, young people can receive support for their mental health and well-being and a sense of belonging.
- **Legal help** – Assistance with legal issues, including migration and family law.
- **Education and Recreational Activities** – Organization of workshops, sports activities, and other programs that promote physical health and personal development.
- **Advocacy and lobbying** – Efforts to influence political decisions that affect children, through lobbying and campaign work.

DATASETS & DATA SOURCES

Swedish NGOs have access to a wide array of data sources that can significantly strengthen efforts to combat child sexual abuse. These sources include case management systems used in service delivery, such as counseling session notes and records of service usage, as well as survey data collected directly from children, families, and communities.

Support services—such as telephone helplines, chat platforms, and online forums—also generate valuable data, including call logs, chat transcripts, and issue tracking information.

In addition, government data from agencies such as the National Board of Health and Welfare (Socialstyrelsen), local social services, and the judicial system—including court records and statistics from the Swedish Crime Victim Authority (Brå)—provide critical insights that help shape effective responses.

Furthermore, participants from the NGO sector in Sweden who contributed to focus groups identified several specific data sources they consider particularly relevant (see Table 2).

Table 2. Examples of data sources that the swedish ngo sector uses – identified by focus groups

TYPE OF DATA	EXAMPLES
Survey Data	• Recurring or standalone surveys across demographics or countries on child sexual abuse and violence, children's voices, or vulnerability
Chat Data	• Chat logs from support chats with professionals or volunteers • Real-time streams of chats, with harmful behaviour indicated by AI tool
Financial Data	• Transactions suspected to be related to using child sexual abuse material • Financial indicators related to payments for sexual abuse collected by financial and tech coalitions
Personal Data	• Interviews • Hotline reports regarding suspected child sexual abuse • Target group data including sensitive, health-related information • Public data/records from incident reports to the Health and Social Care Inspectorate (IVO)
Usage & Meta Data	• Data from Google Analytics and AI transcription services • Metadata from support chats – data that has been coded or tagged by administrators

DATA ACCESS

The type of data managed within a non-governmental organization (NGO) significantly influences its openness and accessibility. Generally, focus group participants from the NGO sector offered limited reflections on data sharing. One participant noted, "Everyone wants to use other actors' data, but no one wants to share their own." In most cases, access to data is granted only within the framework of a joint project.

DATA RELEVANCE

There are numerous examples of NGOs leveraging data to identify high-risk behaviors in children, both with and without the assistance of AI. These efforts help organizations stay informed about emerging trends, produce reports to disseminate knowledge, and implement earlier intervention measures. Additionally, AI-driven chatbots and automated systems enable support services to provide immediate assistance and efficiently refer cases to human operators for further support (see Case Study 1). This technology not only broadens the reach and accessibility of support services but also improves their responsiveness and overall impact in protecting children from abuse.

Case study 1

EXAMPLE OF SWEDISH NGO USE OF DATA AND AI TECHNIQUES

Overview
The Support Cluster against sexualized violence, or "the Cluster" , consists of five Swedish child rights organizations, which have in common that they all provide online chat support services to children and young people. Together they meet nearly 20 000 children and young people annually. The Cluster was created with the aim of coordinating nationwide low-threshold support for children and youths through support lines and chats run by employees and volunteers. Topics span partner violence, sexual exploitation, and various aspects of youth well-being. Currently, the cluster is exploring the usage of AI in analysis and evaluation of the online chats, in collaboration with the AI-company TenFifty.

Data Location & Collection
The data is generated from anonymous online chat services provided to children coming to these organizations for advice and help. Text and metadata are extracted from these chats. Any personal identifiable information is automatically redacted before the information is analyzed, such as phone numbers, school or city names. The intended use of the datasets is to:

- **Automatically fill out** evaluations after completing chats with children and youths, tagging conversations, and generating chat summaries in text.
- **Gather and segment** the estimated ages and age categories of children into a common taxonomy, instead of working with fragmented age labels such as "13-17" or "middle schooler" in the same dataset.
- **Understand,** segment, and make concepts/synonyms searchable in the database.
- **Capture new words** and concepts from chat conversations to enable early detection of trends and patterns in the vulnerability of children and youths.
- **Provide sentiment analysis,** a field of natural language processing that involves determining the emotional tone behind a body of text, to determine the young child's attitude towards a particular topic.

Data Constraints & Barriers
The data from chats are only shared in aggregated form between the organizations. The main constraints and barriers in sharing and using this data for collaborative purposes regard ownership, privacy and child protection laws.

- **Conditional Access:** The data is owned by the Cluster organizations and is only available under certain conditions or for specific purposes (e.g. academic research).
- **Privacy:** the data created from chats with volunteers may include private sensitive information categories such as personal (for example: gender, age, name, address, social security number, etc.) and identity (for example: voting behavior, religion, sexual orientation). It cannot therefore be shared before anonymization.
- **Data concerns children** (under 18) who are subject to stricter data protection.

Opportunities
The support seekers consent to data being collected and used for insights upon accessing the online chat service. However, the datasets contain both sensitive data and children's data, meaning that access to this data and its processing would most likely require both a legal basis and ethical review. One solution the data source is currently exploring is using anonymization or synthetic data generation techniques to create access to the data, whilst providing robust protection. Both solutions vary in terms of feasibility when factoring in cost and time, however, establishing direct integrations with third-party chat systems might streamline continuous data flow to improve secure access to the data.

CRIMINAL JUSTICE

The criminal justice system plays a pivotal role in preventing and addressing child sexual abuse within society. Through law enforcement agencies, courts, and corrections facilities, the criminal justice system works to investigate allegations of abuse, prosecute offenders, and provide justice for victims. Additionally, the system implements preventive measures such as educational programs, community outreach initiatives, and offender monitoring to reduce the occurrence of child sexual abuse.

As a key stakeholder, the criminal justice system provides a framework for addressing cases of abuse, holding perpetrators accountable, and safeguarding vulnerable children. Its involvement in research on AI-driven solutions is crucial, as it can offer valuable insights into patterns of abuse, the effectiveness of interventions, and emerging trends in criminal behavior. By leveraging its expertise and access to relevant data, the criminal justice system can inform the development of AI tools and strategies to enhance prevention efforts and protect children from harm.

THE CRIMINAL JUSTICE SYSTEM IN SWEDEN

The criminal justice sector holds vital data on perpetrators and potential crimes, yet access to such data remains governed by various legal frameworks. The Principle of Public Access grants the public access to certain judicial documents in Sweden, including legal rulings, court records, statistical reports, and research commissioned by public authorities. However, access to personal data is restricted under data protection laws like GDPR.

DATASETS & DATA SOURCES

The criminal justice data landscape in Sweden comes from many different sources. For example:

- Legal rulings and records can be found in legal databases. Entities like the Swedish National Courts Administration (Sveriges Domstolar), subscription-based services like Juno, JP Infonet, and InfoTorg Juridik, alongside university libraries, offer access.
- Statistical data and reports often accessible through websites or via special requests. For example, the Swedish Crime Survey (Nationella trygghetsundersökningen) is an annual survey of the attitudes and experiences of the general population of Sweden regarding victimization, fear of crime and public confidence in the justice system. Organizations like the Swedish National Council for Crime Prevention (Brå) manage and publish much of this data. See Case Study 2 for more information on the Swedish Crime Survey.
- Data related to prison activities, probation, and personal data processing within the criminal justice system is held by the Swedish Prison and Probation Service (Kriminalvården) and the Swedish Prosecution Authority (Åklagarmyndigheten).
- In addition, focus group participants working within the Swedish criminal justice systems identified data sources that they work with most frequently (Table 3).

Table 3. Examples of data sources that the swedish NGO sector uses – identified by focus groups

TYPE OF DATA	EXAMPLES
Interview and metadata	• Investigation data - including interrogation, evidence, metadata from forensic investigations • Interviews with suspects (written and recorded) • Material in the form of text, mail addresses, IP, nicknames, geolocation, media platforms • Administrative data, structured form, no text, only code
Legal data	• Reports from National Center for Missing and Exploited Children (NCMEC) • Internal police large language model doing transcriptions and sum-ups of texts • Judicial systems records • Crime reports records
Video, photo, audio	• Images and video connected to evidence • Legal interception (wire taps) • Audio files
Chat data	• Text chats (mainly from suspects)
Financial data	• Financial statements from suspects • Datasets for training AI (specifically made available to special ops group) • Financial transaction data

DATA ACCESS

Many types of data processed within the criminal justice sector are illegal to possess or use unless handled within the framework of criminal investigations, research, or other professional purposes. Personal data in law enforcement is subject to special regulations beyond the GDPR.

The Criminal Data Act[22] applies to personal data processing within law enforcement activities at authorities such as the Swedish Police Authority and the Swedish Prosecution Authority. The term law enforcement activities refer to all work carried out with the purpose of preventing, investigating, detecting or prosecuting crimes, as well as executing sentences. The Criminal Data Act stipulates that authorities may only process personal data necessary for carrying out their duties, executing sentences, or maintaining public order and security. Data must be erased when no longer needed, and access is restricted to authorized personnel. Data protection officers consult the Swedish Authority for Privacy Protection (IMY) on how they collect and use personal data and report certain personal data breaches to the agency.

Additionally, despite public accessibility, certain documents undergo redaction or face restrictions to safeguard sensitive information, such as information related to national security or ongoing investigations. Obtaining data digitally might be cumbersome, often requiring physical visits or formal requests.

DATA RELEVANCE

By analyzing large-scale datasets, AI can provide decision recommendations with a scale, speed and depth of detail unmatched by human analysts[23]. For example, by analyzing data from various digital sources such as conversations, social media interactions, and search terms, AI can detect early indicators of high-risk behaviors by identifying patterns in chat activity, suspicious contacts with minors, and specific search behaviors on forums. AI tools enhance this process by identifying grooming behaviors and potential threats through deep analysis of digital footprints and chat logs, thereby enabling early intervention.

[22] Integritetsskyddsmyndigheten (IMY), Data Protection within Different Areas https://www.imy.se/en/organisations/data-protection/data-protection-within-different-areas/ accessed 17-06-2024.
[23] Bracket Foundation, Artificial Intelligence: Combatting Sexual Abuse of Children (2019) https://static1.squarespace.com/static/5d7cd3b6974889646fce45c1/t/632f37b896470d1340136fc9/1664038845748/AI.pdf accessed 17-06-2024.

Furthermore, AI can be integrated directly into social media platforms and online forums to monitor user behavior in real-time. This proactive approach not only flags potential abusive actions but also helps in redirecting individuals exhibiting risky behaviors towards rehabilitation services, thereby preventing potential offenses. AI is already being used globally to prevent, detect and prosecute crimes against children[24].

AI can also significantly streamline criminal investigations by enhancing the efficiency and effectiveness of various operational processes. It can accelerate the analysis of seized data, such as images, videos, and documents, by employing machine learning and pattern recognition to sift through vast volumes of information more rapidly than manual methods. This capability not only speeds up investigations but also allows law enforcement professionals to focus on other critical tasks.

Through the analysis of vast datasets, including historical behavior and social media activity, AI algorithms can assess threat levels and prioritize individuals posing the highest risk. This enables law enforcement to allocate resources more effectively, focusing on the most dangerous or relevant cases first.

Additionally, AI can automate the sorting and filtering of reports and tips, eliminating duplicates and irrelevant information. This improves the efficiency of early-stage investigations and reduces the psychological strain on law enforcement professionals who must review harmful or violent content.

In Sweden, several AI tools are used by law enforcement for perpetrator prevention and prosecution. For example, Paliscope[25] software for data analysis and management allows users to conduct online investigations and quickly and easily collect open source data for review and analysis using third party services to correlate and find more data on the Clearnet and Darknet. Meanwhile, Griffeye[26] is used for image and video detection, classification as well as data analysis and management.

[24] Brackett Foundation, Artificial Intelligence: Combatting Sexual Abuse of Children (2019) https://static1.squarespace.com/static/5d7cd3b6974889646fce45c1/t/632f37b89647013401360fc9/1664038845748/AI.pdf accessed 17-06-2024, 24 fig 4:1.
[25] Paliscope, Paliscope https://www.paliscope.com/ accessed 14-02-2025.
[26] Magnet Forensics, Magnet Griffeye https://www.magnetforensics.com/products/magnet-griffeye/ accessed 30-04-2025.

Case study 2

ENHANCING COMMUNITY SAFETY THROUGH THE SWEDISH CRIME SURVEY

Overview
The Swedish Crime Survey (SCS) (Nationella trygghets-undersökningen) collects comprehensive annual data on public exposure to crime, perceptions of safety, and confidence in the criminal justice system. The survey targets a demographically representative sample of 200,000 individuals aged 16–84 across Sweden, using web- and postal-based questionnaires to gather responses.

The dataset from the SCS is essential for understanding societal trends in crime perception and victimization, particularly among vulnerable populations including children. The insights drawn from this data are pivotal in shaping public policy and enhancing community safety strategies.

Data Location & Collection
SCS data covers a broad spectrum of topics including crime exposure, fear of crime, and interaction with the justice system. Data collected is enriched with demographic information such as age, education, and residency, sourced directly from national registries.

Each year, the survey receives responses from approximately 64,000 participants. After collection, the data undergoes rigorous anonymization to ensure privacy before being securely stored on-premises.Transfers are conducted using encrypted methods to maintain confidentiality.

For data security, Brå employs various technical solutions, including data masking, pseudonymization, and encryption, to protect sensitive information.

Data Constraints & Barriers
While the survey's extensive reach and detailed questionnaire generate extensive data, its use and sharing are subject to certain constraints:

- **Restricted Access:** Access to detailed datasets is restricted to ensure privacy, limiting broader academic and public utility. Microdata can be provided to researchers for specific research projects after a confidentiality review. An prerequisite is that the project has been approved by the Swedish Ethical Review Authority.
- **Anonymity and Privacy Concerns:**
A significant constraint is the limitation in data sharing due to privacy concerns and the sensitivity of the subject matter. Anonymization of the data, while crucial for ethical reasons, can reduce the granularity of the data, limiting the depth of analysis that can be performed.

Opportunities
Utilizing AI and advanced analytics on the SCS dataset presents new opportunities for combating child sexual abuse, including pattern recognition and predictive analysis. Advanced data analytics and machine learning models can uncover hidden patterns and trends, providing deeper insights into risk factors. These insights can also support the early development of preventive strategies, enhancing efforts to protect vulnerable individuals.

HEALTH

Health professionals, including doctors, nurses, psychologists, and social workers, are often the first point of contact for children who have experienced abuse. They play a crucial role in identifying signs of abuse, providing medical care and treatment, and connecting victims and their families with appropriate support services such as counseling and legal assistance. Moreover, training programs for healthcare providers on recognizing and responding to abuse and advocating for policies aimed at protecting children from harm are essential for prevention.
By addressing the physical and psychological consequences of abuse and promoting prevention strategies, the health system contributes significantly to safeguarding the well-being of children and reducing the prevalence of child sexual abuse.
Health institutions also hold valuable data related to cases of abuse, medical records, and patterns of healthcare utilization. This data can provide critical insights into the prevalence and characteristics of child sexual abuse, supporting research, policy development, and intervention strategies.

THE HEALTH SECTOR IN SWEDEN[27]

Public and private healthcare providers in Sweden collectively manage extensive records, spanning outpatient care (Öppenvård), inpatient care (Slutenvård), and private practices. In Sweden, both healthcare and dental care providers are legally required to immediately report to the social services if they become aware of or suspect that a child is being harmed. According to a 2019 survey by the National Board of Health and Welfare (Socialstyrelsen), healthcare and dental care providers account for 17% of all the reports subjected to social services in Sweden.

DATASETS & DATA SOURCES

The term "health data" lacks a legal definition but encompasses a wide range of information, including patient records, quality registries, research data, and health information collected through apps. In Sweden, health data is stored across various journal systems, managed by multiple entities contributing to a diverse data landscape.
Key sources of health data in Sweden include:

- **The Public Health Agency of Sweden** (Folkhälsomyndigheten): A government agency responsible for public health, producing statistics, and collecting data from national surveys, such as the annual public health survey.
- **The National Board of Health and Welfare (Socialstyrelsen):** A government agency responsible for managing national health data and producing statistics and registers.
- **The Swedish eHealth Agency (eHälsomyndigheten):** Manages electronic health records, e-prescriptions, and other digital health data.
- **Public & private healthcare providers:** Various healthcare providers collect data on patient medical history, diagnoses, treatment plans, and medications.
- **Patient Registers:** National registers like the National Patient Register (Patientregistret) record hospital admissions, diagnoses, and treatments.
- **Swedish Quality Registers (Svenska Kvalitetsregister):** They collect data on specific diseases or medical conditions to monitor and improve healthcare quality, including real-time data generation.
- **Research Institutions:** Universities and institutions conducting medical research and collecting health data for research purposes in Sweden.

Across regions and healthcare providers, health data is stored in various journal systems. In Stockholm alone, multiple systems, including TakeCare, the National Patient Overview (Nationella Patientregistret), Obstetrix, and Frapp, are in use —illustrating the complexity and diversity of health data sources.

The sophisticated health data infrastructure is exemplified by the VAL database (VAL-databaserna)[28], administered by Region Stockholm. This database aggregates data from 3,000 healthcare providers and records all healthcare contacts reimbursed by Region Stockholm. While the VAL database covers most healthcare interactions, certain exceptions exist, including some privately funded healthcare providers, basic municipal home care, and healthcare services in special housing. The database contains detailed records of each care contact, information on care providers, and de-identified patient data.

[27] Emma Börjesson and Loisa Cedergren, An Introduction to the Field of Health Data in Detecting Child Sexual Abuse (Health Data Sweden, 2023) prepared for World Childhood Foundation.

[28] VAL-databaserna – Region Stockholm, VAL-databaserna – Folkhälsokollen https://www.folkhalsokollen.se/datakallor/val-databaserna/ accessed 04-03-2024.

DATA ACCESS

Beyond GDPR regulations, the Patient Data Act (Patientdatalagen) governs personal data processing in healthcare and dental care, imposing additional obligations on both private and public healthcare providers. Sensitive patient data is strictly regulated and may only be processed under exceptional circumstances. The Ethical Review Act (Lag om Etikprövning) requires prior ethical approval from the Ethical Review Authority (Etikprövningsmyndigheten) for any research involving such data.

The Ethical Review Act
The Swedish Ethical Review Act governs research involving humans, biological material from humans, or sensitive personal data. Its main goal is to protect individuals' integrity and ensure research is conducted ethically. The ethical review act allows health registry data to be accessed for research following a special review process, which typically takes two to four months and incurs fees. In order to be approved, the research need to be ethically approved, the data anonymized or pseudonymized where possible and the use of the data clearly justified.

Public entities managing health data are currently restricted to using it within their jurisdictions, making data sharing between different authorities challenging.

Legal barriers also hinder the secondary use of data for research and innovation. Despite advancements in privacy-enhancing technologies such as anonymization and encryption, there are still insufficient legal, organizational, or technical frameworks for effective secondary use of health data. Several national and European initiatives aim to improve health data utilization. In June, 2025, the Swedish government launched an inquiry into increasing the sharing of relevant data between government agencies and municipalities which will safeguard sensitive information while still enhancing the opportunities for AI development. The inquiry is set to present in June 2026. Sweden's national life science strategy prioritizes the use of health data for research and innovation, while the national data strategy focuses on increasing data access for artificial intelligence and digital innovation. Both strategies emphasize the importance of leveraging existing data more effectively. Additionally, in May 2022, the Swedish government appointed a committee to explore and enhance the potential for secondary use of health data to strengthen healthcare. In January of 2025, the EU adopted the European Health Data Space (EHDS), described as an infrastructure of regulations and practices to create a common

[29] European Health Data Space, Home | European Health Data Space https://www.european-health-data-space.com/ accessed 11-09-2025.

The Swedish Ethical Review Act governs research involving humans, biological material from humans, or sensitive personal data. Its main goal is to protect individuals' integrity and ensure research is conducted ethically.

infrastructure for managing and sharing of health data across the EU. The EHDS is prioritizing data protection and patient safety while still enabling secondary use of data for informing policies, research and innovation[29].

DATA RELEVANCE

Adult health data remains relevant in understanding child sexual abuse, as adults may disclose past abuse to healthcare providers. Such disclosures can help identify long-term consequences, including mental disorders, reduced psychological well-being, and other health issues linked to childhood sexual abuse, ultimately improving treatment and support.

In line with this approach, research using the Swedish VAL health database has demonstrated how health data can be leveraged to combat child sexual abuse. By identifying patterns of healthcare utilization among adolescent girls who have experienced abuse, the study highlighted opportunities for early intervention and the development of tailored treatment plans[29].

[30] Gita Rajan, Sanna Syding, Gunnar Ljunggren and others, 'Healthcare Consumption and Psychiatric Diagnoses among Adolescent Girls 1 and 2 Years after a First-Time Registered Child Sexual Abuse Experience: A Cohort Study in the Stockholm Region' (2021) 30 European Child & Adolescent Psychiatry 1803 https://doi.org/10.1007/s00787-020-01670-w accessed 06-10-2024.

Case study 4

PREVENTING CHILD SEXUAL ABUSE WITH PROJECT BRIDGE

Overview
Karolinska Institutet, in collaboration with Region Stockholm and European academic partners, administers Project Bridge with the purpose to develop tailored healthcare and evidence- based services for individuals at risk of committing child sexual abuse. Based on the individual's assessed risk level and cultural factors, it screens the individual's need for either in-person treatment or self-help programs if deemed sufficient. Initial online chats help identify potential offenders, followed by therapist-led interviews on sexual thoughts, feelings, and behaviors involving children. Participants provide chat data and complete online questionnaires before, during, and after the clinical trials.

Data Location & Collection
Demographic and behavioural data on adult participants at risk of sexually offending is systematically collected through digital means and stored on a platform managed by Linköping University.

The collected data comprises of anonymized text datasets which include motivational interviewing chat logs, written home assignments, and self-report questionnaires from participants.

Data Constraints & Barriers
There are some constraints related to the use and sharing of this data:

- **Conditional access:** Access to the research data is strictly regulated, requires ethical approval and is available only under specific conditions which limits broader academic and clinical application.
- **Anonymity and Privacy Concerns:** A significant constraint is the limitation in data sharing due to privacy concerns and the sensitivity of the subject matter. Anonymization of the data, while crucial for ethical reasons, can reduce the granularity of the data, limiting the depth of analysis that can be performed.

Opportunities
The data within Project Bridge presents several opportunities for employing AI to support therapists in screening individuals applying for support. Within the project, the evidence-based treatment model Motivational Interviewing (MI) has been tested on specific groups—a model in which each session is thoroughly evaluated using quantitative measures to ensure that the therapist adheres to the protocol. These evaluations are currently conducted manually and limited to 20-minute segments to maintain high-quality assessment, which may create a bottleneck.

The structured protocols make the model theoretically well-suited for applying AI—both in the evaluation process and as a complement to current chat-based sessions through chatbot-led alternatives. One hypothesis is that the anonymity and absence of human judgment in such interactions may help reduce feelings of stigma and shame among participants, potentially encouraging more open dialogue and greater engagement from individuals dealing with sensitive issues.

ACADEMIA

By bridging interdisciplinary research, academia is advancing our understanding of the causes, effects, and prevention of child sexual abuse. Experts from psychology, sociology, criminology, law, and related fields are coming together to explore the issue from multiple perspectives. Through empirical studies, researchers are uncovering the prevalence and risk factors of child sexual abuse and identifying effective intervention strategies.

In advancing expertise in data analysis and computational modeling, academia can lead efforts to harness the potential of AI for identifying patterns, detecting early warning signs, and developing predictive models related to child sexual abuse. By strengthening collaboration with AI experts, the social sector, and private industries, academia can bridge the gap between research and practical application, driving innovative solutions to combat child sexual abuse. Academia's access to diverse datasets and case studies offers valuable insights into available data types and their effective use in shaping AI-driven approaches to this critical issue.

ACADEMIA IN SWEDEN

Academic institutions in Sweden play a crucial role in training professionals in child protection, law enforcement, healthcare, and education, providing them with the knowledge and skills needed to identify, address, and prevent instances of child sexual abuse. Sweden's open access and open data movement seeks to make scientific research more transparent and widely accessible. Led by organizations such as the Swedish Research Council (Vetenskapsrådet), this initiative emphasizes that research data funded by public resources should be freely available online, adhering to the principle of "as open as possible, as closed as necessary." This ensures data is accessible unless valid privacy or security concerns warrant restrictions[31]. For instance, Forte[32], the Swedish Research Council for Health, Working Life and Welfare, introduced a call for proposals specifically to promote the reuse of existing data. This effort highlights the value of maximizing data utility by enabling researchers to build on previous findings rather than duplicating data collection, fostering both innovation and efficiency in research. Overall, Sweden's shift towards open access and open data represents a broader strategy to enhance the quality and impact of scientific research through greater transparency and collaboration.

[31] Gita Rajan, Sanna Syding, Gunnar Ljunggren and others, 'Healthcare Consumption and Psychiatric Diagnoses among Adolescent Girls 1 and 2 Years after a First-Time Registered Child Sexual Abuse Experience: A Cohort Study in the Stockholm Region' (2021) 30 European Child & Adolescent Psychiatry 1803 https://doi.org/10.1007/s00787-020-01670-w accessed 06-10-2024.

[32] Forte, About Us – Forte, the Swedish Research Council for Health, Working Life and Welfare https://forte.se/en/about-us accessed 11-09-2025.

EXAMPLE 3: IMPROVING ACCESS TO EXISTING DATA

Improved access to and understanding of existing datasets can support more comprehensive and informed research, ultimately advancing prevention and intervention strategies.

Registerforskning.se, managed by the Swedish Research Council (Vetenskapsrådet), provides researchers with comprehensive information and support for utilizing existing registers in their studies.

A key feature of this platform is the Data Guide (Dataguiden) compiling register- and health data from municipalities, authorities, research project, among others. With this, they have also launched the Register Utiliser Tool (RUT), a new metadata tool that enables advanced searches and comparisons across Swedish registers. By offering structured and standardized metadata, RUT enhances researchers' ability to efficiently locate and access relevant data.

This tool is particularly valuable for researchers working on child sexual abuse, as it streamlines the process of identifying, requesting, and utilizing pertinent data.

DATASETS & DATA SOURCES

Universities maintain robust repositories including publication databases, journals, interdisciplinary databases, and libraries housing extensive research materials. In addition, focus group participants shared data sources that they use (Table 4).

Table 4. Examples of data sources that swedish academic sector uses – identified by focus groups and surveys

TYPE OF DATA	EXAMPLES
Personal Data	• Self-reported [behavioral] info as part of research projects • Work sheets from interventions • Interviews as part of research projects • Medical data such as from the VAL-databases which contain information on each care contact, the care provider and de-identified patient data for the Stockholm region.
Reports	• Official documents from state authorities • Statistical data (such as from Brå)
Chat data	• Chat based data as part of research projects

DATA ACCESS

Accessing and sharing data within the Swedish academic sector involves careful adherence to institutional and legal frameworks. Engaging with academic institutions often requires Data Processing Agreements (Personuppgiftsbiträdesavtal) to ensure compliance with data protection regulations. Researchers must also follow institution-specific guidelines and existing agreements, making Data Management Plans[33] essential. They outline how data will be managed during and after a research project. The Swedish National Data Service (Svensk Nationell Datatjänst) supports researchers by providing a detailed checklist[34] and digital tools for creating and maintaining these plans.

Moreover, most Swedish academic institutions appoint designated data officers responsible for overseeing data access and ensuring compliance with ethical and legal standards. Academic institutions usually have information available to guide the process of data access and sharing[35].

All research in the academic sector managing personal or sensitive data is subject to an ethical approval process, which involves submitting a detailed research plan, the responsible researcher's resume, information for research subjects, and other project-specific details. This process ensures that research meets ethical guidelines and typically concludes with a decision within 60 days after receiving a complete application.

Furthermore, at the national level, Registerforskning[36].se , operated by the Swedish Research Council (Vetenskapsrådet), provides researchers with comprehensive information and support for register-based research. It details the processes involved in identifying, requesting, and using data from various registers. Some of Sweden's reliable data sources include national authority registers, health and medical quality registers, biobanks, and researcher-generated data. The Nordic countries, with their established population registers and biobanks, offer unique opportunities for impactful research, leveraging the combined data of over 26 million people to draw significant conclusions, even for rare events or diseases.

DATA RELEVANCE

Across disciplines such as psychology, sociology, education, and criminology, researchers gather data that deepens our understanding of child development, learning, and behavior. This research also provides critical insights into the characteristics, motivations, and methods of both potential and convicted perpetrators.
By identifying behavioral patterns and risk factors associated with abuse, academic studies help reveal early warning signs. For example, psychological and criminological data contribute to profiling potential perpetrators, which can inform the development of risk assessment tools for early intervention and monitoring. Additionally, research in sociology and criminology sheds light on the broader social and environmental conditions that may foster abusive behavior.

[34] Swedish National Data Service, Checklist for Data Management Swedish National Data Service, Checklist for Data Management Plan (version 12, 1 July 2021) https://snd.se/en/resources/checklist-data-management-plans accessed 04-04-2025. Plan (version 12, 1 July 2021) https://snd.se/en/resources/checklist-data-management-plans accessed 04-04-2025.
[33] Swedish National Data Service, Data Management Plan https://researchdata.se/en/manage-data/organize-data/data-management-plan accessed 04-04-2025.

[35] For example Chambers Univervisty: https://www.lib.chalmers.se/en/publish-and-analyse/open-access/research-data/
[36] Swedish Research Council, Registerforskning.se https://registerforskning.se/en/ accessed 04-04-2025.

Case study 5

BARNSÄKERHET – ENHANCING CHILD SAFETY THROUGH DATA

Overview
BarnSäkert, a research project managed by the REACH research group at Uppsala University, has adapted and developed a working method in child health care in collaboration with social services to collect data from early health check-ups. When a parent or parents attend a health check-up at child health care facility, they are also presented with a questionnaire on certain topics pertaining to the child, their parents and their environment- to assess potential risk factors and allow for recommendations in care. The project leverages structured methods to identify and analyze psychosocial risk factors in families of young children, with a specific focus on preventing child maltreatment and fostering optimal child development.

Data Location & Collection
The primary data source is a digital questionnaire completed by parents during child health visits. The questionnaire gathers data on multiple risk factors including child safety, parental stress, and intimate partner violence. Data is collected anonymously through web-based questionnaires filled out on electronic tablets by parents, which are automatically uploaded to a secure database hosted by Uppsala University.

The data encompasses responses to 14 key questions related to psychosocial risk factors, scored and analyzed to track trends over time. The project has accumulated around 60,000 responses, providing a substantial dataset of text and tabular data for analysis.

Data Constraints & Barriers
he main constraints and barriers in sharing and using this data for collaborative purposes are:

- **Restricted Access:** Due to the sensitive nature of the information, access to the data is restricted to authorized personnel at the county, child health center or central child health services, which limits the potential for broader research collaboration.
- **Data Aggregation:** Aggregating data from various counties, some of which still use paper-based methods, presents challenges in standardization and digital data collection. This poses challenges to using data collaboratively.

Opportunities
AI techniques applied to data in the BarnSäkert project presents a number of opportunities to enhance collaboration and improve models such as:

- **Predictive Models:** The data collected could be used for training models related to prediction and identification of risk. Utilizing AI, BarnSäkert could develop predictive analytics models to identify children at high risk more effectively.
- **Pattern Recognition and Trend Analysis:** AI could be used to detect patterns and trends in the data collected, helping to identify common risk factors among cases of child maltreatment.
- **Enhancing privacy:** Implementing AI privacy- preserving technologies could facilitate safer data sharing, enhancing collaborative opportunities across regions and enabling broader, impact-driven research.

Case study 6

ALLMÄNNA BARNHUSET'S APPROACH TO RESEARCH ON CHILD SAFETY ONLINE

Overview
Stiftelsen Allmänna Barnhuset, a children's welfare foundation in Sweden, focuses on improving child protection through research and development projects. One of their initiatives involves conducting sensitive yet vital anonymous surveys to collect data on children's experiences with sex and violence on the internet through surveys titled "Unga, sex och internet" (Youth, Sex, and the Internet) and "Våld mot barn" (Violence Against Children). The surveys are structured to gather nationally representative insights, aiding significant research output widely cited in academic and policy-making arenas.

Data Location & Collection
The data comprises responses from children and high school students. This data is highly sensitive, containing personal information despite its anonymous format.

It contains demographic data, health data and behavioural data in text and tabular form.

Data Constraints & Barriers
The main constraint related to the use and sharing of this data is:

- **Conditional Access:** The data's sensitive nature necessitates stringent measures to ensure privacy and security, restricting broader data sharing and requiring permissions for specific research uses.

Permissions are strictly controlled by Stiftelsen Allmänna Barnhuset, and data access is granted only for specific approved research purposes to prevent misuse.

Opportunities
Utilizing AI technology presents significant opportunities to enhance the impact of this data. For instance:

- **Pattern Recognition:** AI algorithms can analyze responses to identify common patterns or anomalies in the experiences reported by children, which could indicate prevalent risks or new emerging threats in online environments.
- **Predictive Analysis:** Machine learning models can be developed to predict potential future trends in online child safety, helping policy-makers and educators to stay ahead of potential threats.
- **Text analysis:** AI-driven text analysis can be employed to interpret free-text responses in the surveys, which can reveal deeper insights into the children's experiences and the context of their interactions online.

PRIVATE SECTOR

The private sector collects various types of data, often to improve their products and services. In this report, we have chosen to emphasize two examples of private sector. The financial sector, indirectly reaching children and providing key infrastructure, and the gaming industry, facilitating online platforms for children and adults to communicate, and play. Other private actors, such as social media and telecommunication, carry relevant insights into internet use and factors contributing to risk environments online. As these actors are often part of multilateral companies, data from these sources are often stored in efficient data centers overseas, limiting not only insight into possible data gathered, but highly impacts the level of access for third-party actors. Therefore, they were not chosen for this report. All of these businesses use AI to analyze the data they collect, helping them make smarter decisions and create innovative solutions. By analyzing this data with AI, businesses can personalize marketing, improve customer service, develop better products, and make informed decisions, ultimately enhancing customer satisfaction and operational efficiency. For example, in the gaming industry, companies gather data on how players interact with games and other players and in the finance sector, banks and financial institutions collect data on spending habits.

FINANCE

With a rise in sexual financial extortion, also called sextortion, the role of the financial sector becomes more central in identifying and disrupting financial transactions linked to sexual exploitation of children[37]. A recent report published by Childhood investigated the possibility of utilizing financial transaction data from banks, remittance companies and other financial technology companies in identifying perpetrators paying for livestreamed sexual exploitation of children in the Philippines. The report showed clear opportunities for utilizing AI in advancing detection based on this data. Through continuous monitoring of transactions combined with customer knowledge, suspicious activities from customers can be flagged and reported to authorities, leading to the identification of perpetrators. Activities, which if coordinated between banks, financial technology companies and authorities would provide a more holistic view of the crime and potential interventions.

[37] ECPAT Sverige, "Då tog 'hon' en screen och allt började" (2023) https://ecpat.se/wp-content/uploads/2020/12/Da-tog-hon-en-screen-och-allt-borjade_2023.pdf accessed 09-10-2024.

However, the lack of collaboration and sharing of both knowledge and data proved to be a major hurdle in exploring the use of AI in this context[38].

THE FINANCIAL SECTOR IN SWEDEN

The Financial Services Authority is the principal governmental entity responsible for supervising financial regulations, markets, and stakeholders in Sweden. The Swedish banking sector includes key entities such as public banking institutions and agencies like the Swedish Central Bank (Sveriges Riksbank) and the Financial Supervisory Authority, as well as private sector banks such as Handelsbanken and Swedbank. Additionally, payment service providers like Swish and Klarna play pivotal roles in this sector. Swedish financial institutions monitor diverse activities including banking transactions, e-invoices, and cash withdrawals.

DATASETS & DATA SOURCES

In the financial sector, various types of data are utilized for a multitude of purposes, from investment decision-making to risk management and regulatory compliance. Fundamental data comprises financial statements like balance sheets, income statements, and cash flow statements, along with earnings reports and important financial ratios and metrics. Transaction data involves details of individual trades, including price, volume, and timestamp. Alternative data offers insights from non-traditional sources such as social media sentiment analysis, geospatial data from satellite images and consumer data from credit card transactions, and foot traffic analysis.

When examining larger data sources from the Swedish financial sector, several key institutions provide comprehensive information. The Swedish Central Bank (Sveriges Riksbank) offers extensive financial statistics covering national accounts, the balance of payments, and financial market data. It also publishes detailed reports on the financial health and operations of Swedish banks.

In addition, the Financial Supervisory Authority (Finansinspektionen) supplies data on financial stability, regulatory compliance, and the overall condition of financial institutions in Sweden.

Finance Sweden (Svenska Bankföreningen) contributes further insights through reports and datasets on the banking sector's performance, including profitability, lending activities, and consumer behavior trends.

[38] World Childhood Foundation, 'Follow the Money: Payments for Live Streaming of Child Sexual Abuse Online in the Philippines' (2025) https://childhood.se/wp-content/uploads/2025/03/child004-rapport-follow-the-money-digital-110325-2.pdf accessed 09-05-2025.

DATA RELEVANCE

As custodians of vast amounts of transactional data, financial institutions possess valuable insights that can contribute to understanding the dynamics of child exploitation. Financial data relevant for preventing and combating child sexual abuse includes reports of suspicious activities, containing flagged transaction records, account activity, and patterns of spending that may indicate illegal activities. For example, unusually frequent small transactions in odd numbers, payments to known suspicious entities, or the purchase of services linked to exploitative activities can signal potential abuse. Law enforcement and financial institutions can use this data to flag and investigate accounts, trace funds, and uncover networks involved in child exploitation, leading to the identification and prosecution of offenders and the disruption of criminal operations. By showcasing the diverse forms of data available within the financial sector and illustrating how AI techniques can be applied to analyze and interpret this data effectively, stakeholders can better comprehend the opportunities for intervention and prevention.

Moreover, the financial sector's involvement can extend to supporting initiatives aimed at raising awareness and funding programs dedicated to combating child sexual abuse. For example, in Sweden the Financial Coalition (Finanskoalitionen)[39] includes representatives from the NGO ECPAT, police, banks, and other financial entities focused on combating child sexual abuse and exploitation. The coalition was started in 2007 to stop payments for abuse materials, focusing on stopping purchases through regular bank cards. Their efforts have made it hard to pay for such materials with standard cards today. They continue their work with new projects to prevent the use of financial services for child sexual exploitation.

DATA ACCESS

In the finance sector, banks and financial institutions use data to assess risk, detect fraud, and offer tailored financial products. Issues arise regarding the sharing of sensitive personal data, notably due to legislative barriers such as bank secrecy laws. Ensuring compliance with privacy laws while attempting to gather meaningful data requires careful navigation of legal frameworks, often limiting the scope and availability of valuable information. Access to such data is typically restricted unless banks identify suspicious activities, hindering broader access for investigative or research purposes.

[39] ECPAT Sverige, Ecpat Sveriges Finanskoalition – Finanskoalitionen mot sexuell exploatering av barn https://ecpat.se/vadvigor/finanskoalitionen/ accessed 05-06-2025.

Globally, the standard for regulation of secrecy and risk management in banks is largely influenced by the Basel Committee on Banking Supervision (BCBS), and the Financial Action Task Force (FATF). The latter overseeing the management of sensitive data pertaining to anti-money laundering and counter-terrorism activities. In March of 2025, FATF released a comprehensive report on the role of financial intelligence to support in detecting, disrupting and investigating online child sexual exploitation, specifically the livestreaming of sexual abuse and financial sexual extortion of children. Among other conclusions, it argues that the private sector has a better basis for identifying ongoing abuse, given the amount of data the sector holds and the ability to quickly process it[40].

Additionally, the European Commission's draft proposal for a Regulation on Financial Data Access[41] proposed in June 2023 could significantly influence data access within this sector. They are putting forward proposals to bring payments and the wider financial sector into the digital age. These rules will further improve consumer protection and competition in electronic payments and will empower consumers to share their data in a secure way so that they can get a wider range of better and cheaper financial products and services.

The regulation will enter into force 24 months after its adopted while financial data-sharing schemes, which are integral to its implementation, will need to be established within 18 months of the regulation's adoption.

GAMING INDUSTRY

Vast numbers of both children and adults engage in online gaming. This mix of young and mature players presents both significant risks and opportunities. Children are particularly vulnerable in gaming environments, as they frequently interact with strangers through chat functions and online communities—spaces that can be exploited by offenders for grooming and sexual exploitation. The rapid growth of online gaming, coupled with the anonymity it provides, makes it an attractive platform for predators. Games with social interaction, such as multiplayer and livestreamed games, can become hotspots for grooming, where offenders build trust with young players to manipulate and exploit them. Many offenders move across multiple platforms to avoid detection, making it difficult to track their activities and hold them accountable.

[40] Financial Action Task Force (FATF), Detecting, Disrupting and Investigating Online Child Sexual Exploitation https://www.fatf-gafi.org/content/fatf-gafi/en/publications/Fatfgeneral/Online-child-sexual-exploitation.html accessed 05-06-2025.

[41] European Commission, Financial Data Access and Payments Package https://finance.ec.europa.eu/publications/financial-data-access-and-payments-package_en accessed 05-05-2025

THE GAMING INDUSTRY IN SWEDEN

The gaming industry operates within a global landscape, where data management involves multinational actors handling vast volumes of information. Our ambition with this report was to engage Swedish actors in the gaming industry, but it's very rarely the case that a game developer or gaming platform is not set in a global context, with global actors handling data at large volumes. We also found it difficult to engage Swedish actors in the gaming industry to contribute to the report.

DATASETS & DATA SOURCES

In the gaming industry, data on player interactions, in-game purchases, and preferences is used to create personalized gaming experiences and balance in-game economies. Gaming companies and platforms collect diverse datasets, including behavioral, social, and even biometric data. The use of game telemetry enables tracking of specific in-game actions and interactions, providing valuable behavioral insights. Despite the vast amounts of data collected, anonymization techniques are frequently applied to safeguard user privacy while still allowing for the analysis of behavioral patterns over time. Examples of data collected in games include:

- **Demographic Data:** Information on players' age, gender, location, and other characteristics.
- **Behavioral Data:** Insights into in-game behavior, such as play frequency, level completion, and purchasing patterns.
- **Financial Data:** Information on in-game spending habits, including purchases and subscriptions.

DATA RELEVANCE

The gaming sector is highly relevant due to the widespread engagement of children on gaming platforms, many of which incorporate social elements. While these platforms serve as recreational spaces, they can also inadvertently expose children to online and offline risks of child sexual abuse. To enhance safety, many gaming companies implement measures such as age verification systems and monitoring tools to detect and report suspicious activities or inappropriate behavior. Gaming companies bring valuable expertise in digital technologies and online communities, making them key stakeholders in discussions on online safety and abuse prevention. Additionally, they hold vast amounts of data on user interactions and behaviors, offering critical insights into patterns of abuse and potential intervention strategies.

Several factors affect the relevance and utilization of gaming data in child protection efforts, particularly platform switching and age verification challenges.

- **Cross-Platform Interactions:** Both children and adults frequently switch between gaming platforms, exchanging usernames across different games and communication channels like Discord. This fragmented engagement makes it difficult to conduct comprehensive data analysis and identify patterns of abuse. To mitigate risks, cross-collaboration and data sharing among gaming companies are essential.
- **Age Verification Limitations:** Many gaming platforms lack stringent age verification mechanisms, affecting the accuracy and reliability of collected data. Without distinct age-related markers, such as verifiable age information within text exchanges, it becomes challenging to differentiate between adult and child users, impacting the validity of behavioral insights.

Addressing these challenges requires industry-wide collaboration, improved verification systems, and enhanced data-sharing frameworks to strengthen child protection efforts within gaming environments.

DATA ACCESS

Gaming companies often maintain strict confidentiality around the data they collect, making it difficult to access comprehensive information. Their reluctance to share data stems from extensive user agreements that permit them to track various behavioral and transactional details, creating challenges in obtaining relevant datasets—even in cases where abuse has occurred through their platforms. Additionally, the unique characteristics of gaming data present inherent challenges for AI applications such as Natural Language Processing (NLP). The brevity and speed of communication in massive multiplayer online games make it difficult to analyze in-game speech and written interactions accurately. Furthermore, distinguishing genuine concerns from trolling or misleading language adds another layer of complexity, making large volumes of data inaccessible or unreliable for meaningful analysis.

Reference list

AI Sweden. "Dataset." *AI Sweden*, https://www.ai.se/sv/ai-labs/technology-infrastructure/dataset accessed 12-09-2025.

AI Sweden, My AI (AI Sweden) https://my.ai.se/ accessed 11-09-2025.

AI Sweden, National center for applied AI https://www.ai.se/en accessed 14-06-2025.

Bahuguna, Anuj, "AI in the Loop vs Human in the Loop: A Technical Analysis of Hybrid Intelligence Systems" (IBM Community, 25 May 2025) https://community.ibm.com/community/user/blogs/anuj-bahuguna/2025/05/25/ai-in-the-loop-vs-human-in-the-loop accessed 06-05-2025.

Börjesson E and Cedergren L, An Introduction to the Field of Health Data in Detecting Child Sexual Abuse (Health Data Sweden, 2023) prepared for World Childhood Foundation.

Bracket Foundation, Artificial Intelligence: Combatting Sexual Abuse of Children (2019) https://static1.squarespace.com/static/5d7cd36974889646ce45c1/t/632f37b896470d1340136fc9/1664038845748/AI.pdf accessed 17-06-2024.

ECPAT Sverige, 'Då tog "hon" en screen och allt började' (2023) https://ecpat.se/wp-content/uploads/2020/12/Da-tog-hon-en-screen-och-allt-borjade_2023.pdf accessed 09-10-2024.

ECPAT Sverige, Ecpat Sveriges Finanskoalition – Finanskoalitionen mot sexuell exploatering av barn https://ecpat.se/vadvigor/finanskoalitionen/ accessed 05-06-2025.

Encyclopædia Britannica, Methods and Goals in AI, https://www.britannica.com/technology/artificial-intelligence/Methods-and-goals-in-AI accessed 11 September 2025.

European Commission, Financial Data Access and Payments Package https://finance.ec.europa.eu/publications/financial-data-access-and-payments-package_en accessed 05-05-2025.

European Commission, Proposal for a Regulation of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse COM/2022/209 (Brussels, 11 May 2022) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0209 accessed 11-06-2025.

European Commission, "Regulatory Framework on AI," Shaping Europe's Digital Future https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai accessed 11-04-2025.

European Commission, "Safer Internet Centres," Shaping Europe's Digital Future https://digital-strategy.ec.europa.eu/en/policies/safer-internet-centres accessed 17-06-2025.

European Commission, The Digital Services Act (DSA) explained: Measures to protect children and young people online (Publications Office of the European Union, 22 November 2023) https://op.europa.eu/en/publication-detail/-/publication/f3556a65-88ea-11ee-99ba-01aa75ed71a1 accessed 11-06-2025.

European Health Data Space, European Health Data Space https://www.european-health-data-space.com/ accessed 11-09-2025.

European Parliament, EU AI Act: First Regulation on Artificial Intelligence (News, 1 June 2023) https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence accessed 11-06-2025.

European Parliament, Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) OJ L 119, 4 May 2016, 1-88 https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng accessed 11-09-2025.

Financial Action Task Force (FATF), Detecting, Disrupting and Investigating Online Child Sexual Exploitation https://www.fatf-gafi.org/content/fatf-gafi/en/publications/Fatfgeneral/Online-child-sexual-exploitation.html accessed 05-06-2025.

Forte, About Us – the Swedish Research Council for Health, Working Life and Welfare https://forte.se/en/about-us accessed 11-09-2025.

Integritetsskyddsmyndigheten (IMY), Data Protection within Different Areas https://www.imy.se/en/organisations/data-protection/data-protection-within-different-areas/ accessed 17-06-2024.

Integritetsskyddsmyndigheten (IMY), "Privacy by design and privacy by default" https://www.imy.se/en/organisations/data-protection/this-applies-accordning-to-gdpr/privacy-by-design-and-privacy-by-default/ accessed 12-09-2025.

Internet Watch Foundation, "IWF urges for 'loophole' to be closed in proposed EU laws criminalising AI child sexual abuse as synthetic videos make 'huge leaps' in sophistication", 11 July 2025, Internet Watch Foundation. Available at: https://www.iwf.org.uk/news-media/news/iwf-urges-for-loophole-to-be-closed-in-proposed-eu-laws-criminalising-ai-child-sexual-abuse-as-synthetic-vid os-make-huge-leaps-in-sophistication/ accessed: 12-09-2025.

Internet Watch Foundation, How AI Is Being Abused to Create Child Sexual Abuse Imagery (Research Report, IWF) https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/ accessed 10-03-2025.

Korhonen, L; Lindholm, L; Lindersson, M; and Munger, A-C. The Inclusion of Children in Public Enquiries on Violence, Health and Welfare: The Example of Sweden in Roth, M., Alfandari, R. and Crous, G. Participatory Research on Child Maltreatment with Children and Adult Survivors: Emerald Studies in Child Centered Practice. (Emerald Publishing, 2023), pp. 197-213.

Magnet Forensics, Magnet Griffeye https://www.magnetforensics.com/products/magnet-griffeye/ accessed 30-04-2025.

Myndigheten för digital förvaltning (DIGG), Digg – Myndigheten för digital förvaltning https://www.digg.se/ accessed 11-09-2025.

Myndigheten för digital förvaltning (DIGG), Sveriges dataportal https://www.dataportal.se/ accessed 11-09-2025.

Paliscope, Paliscope https://www.paliscope.com/ accessed 14-02-2025.

Rajan G, Syding S, Ljunggren G and others, 'Healthcare Consumption and Psychiatric Diagnoses among Adolescent Girls 1 and 2 Years after a First-Time Registered Child Sexual Abuse Experience: A Cohort Study in the Stockholm Region' (2021) 30 European Child & Adolescent Psychiatry 1803 https://doi.org/10.1007/s00787-020-01670-w accessed 06-10-2024.

Region Stockholm, VAL-databaserna – Folkhälsokollen https://www.folkhalsokollen.se/datakallor/val-databaserna/ accessed 04-03-2024.

Vetenskapsrådet, Registerforskning.se https://registerforskning.se/en/ accessed 04-04-2025.

RISE, Regulatoriskt växthus, sandlåda eller försöksverksamhet (RISE) https://www.ri.se/sv/expertisomraden/expertiser/regulatoriskt-vaxthus accessed 14-04-2025.

Swedish National Data Service, Checklist for Data Management Plan (version 12, 1 July 2021) https://snd.se/en/resources/checklist-data-management-plans accessed 04-04-2025.

Swedish National Data Service, Data Management Plan https://researchdata.se/en/manage-data/organize-data/data-management-plan accessed 04-04-2025.

UNICEF, When Numbers Demand Action: Confronting the Global Scale of Sexual Violence against Children (New York: UNICEF, October 2024) https://data.unicef.org/wp-content/uploads/2024/10/UNICEF_When-Numbers-Demand-Action_Oct_10_2024.pdf accessed 25-05-2025.

UNICEF and INHOPE, Data for Change: Listening to the Voices of Children about Their Experiences Online (webinar, 2023) https://inhope.org/EN/articles/webinar-recap-data-for-change-listening-to-the-voices-of-children-about-their-experiences-online accessed 11 September 2025.

United Nations Children's Fund, International Classification of Violence against Children (New York: UNICEF, 2023) https://data.unicef.org/topic/child-protection/violence/sexual-violence/#_ftnref1 accessed 25-05-2025.

World Childhood Foundation, 'Follow the Money: Payments for Live Streaming of Child Sexual Abuse Online in the Philippines' (2025) https://childhood.se/wp-content/uploads/2025/03/child004-rapport-follow-the-money-digital-110325-2.pdf accessed 09-05-2025.