Exploring Appropriate Standards to Develop a Child Online Protection Dataset For a Deep Learning Model to Educate Parents

Jennyphar Kavikairiua, Fungai Bhunu Shava, Mercy Chitauro

NAMIBIA UNIVERSITY OF SCIENCE AND TECHNOLOGY

Introduction

Parents frequently struggle to control children's online exposure since they feel they have more control over their offline activities than their online exposure, due to their lack of Internet education. Parents can better comprehend cybersecurity if they are exposed to it early on and are aware of the risks it presents. In order to educate parents about their children's safety online, this study explores the appropriate standards to develop a dataset on children online protection. Deep learning has received a lot of attention recently in the subject of cybersecurity, which includes children's online safety. As a result, it is crucial to explore the permissible standards that can be applied while developing a dataset for children online protection for a deep learning model.

Problem Statement

- Parents frequently struggle to control their children's online exposure, due to lack of Internet knowledge.
- Lack of deep learning models developed to educate parents on protecting children safely online.
- Lack of public dataset that address child online safety.

Objective

Explore the appropriate standards to develop a dataset that is suitable for a deep learning model that can educate parents on ways to protect their children online.

Related Work

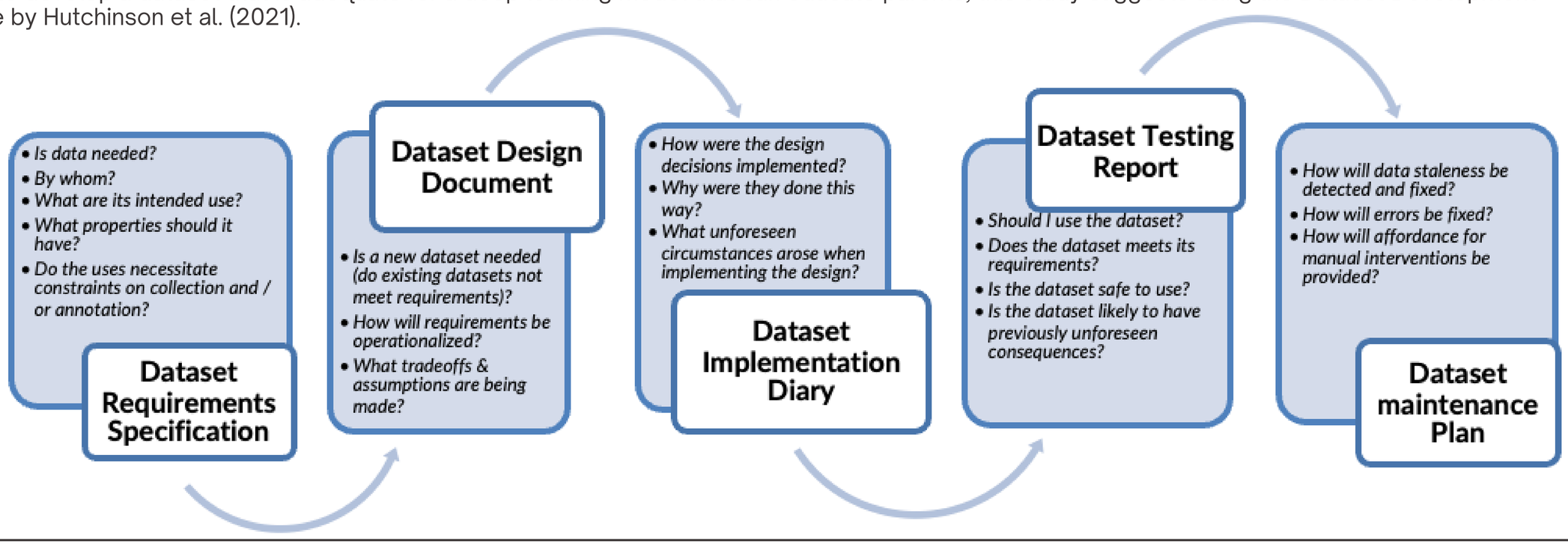
Cybersecurity exiting frameworks/models/standards that can be applied to develop the Child Online Protection Dataset to educate parents	Framework/model/ strategies focus area
NIST Cybersecurity Framework (CSF)	Identify; protect; detect; respond; and recover
Cybersecurity Strategy	Risk identification; vulnerability reduction; threat reduction; consequence mitigation; and enable cybersecurity outcomes
Guidelines for Industry on Child Online Protection	Educating children, parents, and educators about children's protection and responsible use of ICTs
Guidelines for Children on Child Online Protection	Digital etiquette; digital access; digital law; digital security (self-protection); digital communication; digital rights and responsibilities; and digital literacy, digital commerce
Guidelines for Parents, Guardians and Educators on Child Online Protection	Personal device safety and security; Law, education of parents, guardians, and teachers; education for children; and communication
Standards for measuring child online protection	Risk-prone conduct of children (activities and time spent online); cyber threats and incidents; reaction of children to such incidents; and preventive measures
Child Protection in the Online/Offline environment Strategies	Understand the social and cultural context; online / offline sexual harassment and abuse; actions of children using information and communication technology; vulnerability and harm to risks; and sources of help and assistance

Dataset Development

- Yang et al. (2023): The research framework of cement production life cycle inventory dataset for China.
- Khan et al. (2023): A Framework for Dataset Accountability
- Paullada et al. (2021): Dataset design and Development
- Hutchinson et al. (2021): Dataset Development Lifecycle
- Shaukat et al. (2018): A Dataset for Software Requirements Risk Prediction
- Xu et al. (2020): Contents of the MBD Dataset
- Pushkarna et al. (2022): Purposeful and Transparent Dataset Documentation for Responsible AI
- Göbel et al. (2022): A holistic forensic data set synthesis framework
- Baumgartner et al. (2020): The pushshift telegram dataset

Proposed Approach

In order to develop a dataset that is adequate for a deep learning model that can educate parents, this study suggests using the Dataset Development Lifecycle by Hutchinson et al. (2021).



Conclusion

We explore the necessary standards for developing a dataset suited for a deep learning model that can educate parents in order to protect their children online. The development of the child online protection dataset can be supported by related literature on cybersecurity that highlights risk identification awareness and training, creating a more secure and age-appropriate online environment, children's risk-prone behaviour, online threats and incidents, and children's reactions to those incidents as well as preventive measures. In addtion, the Dataset Development Lifecycle was suggested to develop the dataset.

References

Altarturi, H. H. M., & Anuar, N. B. (2020). A preliminary study of cyber parental control and its methods. 2020 IEEE Conference on Application, Information and Network Security, AINS 2020, 53–57. https://doi.org/10.1109/AINS50155.2020.9315134

Broadband Commission ITU UNESCO. (2019). Child Online Safety: Minimizing the Risk of Violence, Abuse and Exploitation Online.

Cecillon, N., Labatut, V., Dufour, R., Linares, G., Cecillon, N., Labatut, V., Dufour, R., & Linares, G. (2021). Graph embeddings for Abusive Language Detection To cite this version : HAL Id : hal-03042171.

Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. Methods in Ecology and Evolution, 10(10), 1632–1644. https://doi.org/10.1111/2041-210X.13256

Hutchinson, B., Smart, A., Hanna, A., Denton, E., Greer, C., Kjartansson, O., Barnes, P., & Mitchell, M. (2021). Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. FAccT 2021 - Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 560–575. https://doi.org/10.1145/3442188.3445918

UNICEF. (2023). Child Protection in Digital Education.

Upadhyay, A., Chaudhari, A., Arunesh, Ghale, S., & Pawar, S. S. (2017). Detection and prevention measures for cyberbullying and online grooming. Proceedings of the International Conference on Inventive Systems and Control, ICISC 2017, 1–4. https://doi.org/10.1109/ICISC.2017.8068605