World Academy of Science, Engineering and Technology
International Journal of Industrial and Systems Engineering
Vol:16, No:06, 2022

Towards a Conscious Design in AI by Overcoming Dark Patterns

Authors : Ayse Arslan

Abstract : One of the important elements underpinning a conscious design is the degree of toxicity in communication. This study explores the mechanisms and strategies for identifying toxic content by avoiding dark patterns. Given the breadth of hate and harassment attacks, this study explores a threat model and taxonomy to assist in reasoning about strategies for detection, prevention, mitigation, and recovery. In addition to identifying some relevant techniques such as nudges, automatic detection, or human-ranking, the study suggests the use of major metrics such as the overhead and friction of solutions on platforms and users or balancing false positives (e.g., incorrectly penalizing legitimate users) against false negatives (e.g., users exposed to hate and harassment) to maintain a conscious design towards fairness.

Keywords : AI, ML, algorithms, policy, system design
Conference Title : ICSMC 2022 : International Conference on Systems, Man and Cybernetics
Conference Location : New York, United States
Conference Dates : June 02-03, 2022