

X-by-design Ansatz: Bürde oder Fortschritt?

Zusammenfassung

Der X-by-design-Ansatz macht gerade internationale Karriere. Wie Erfahrungen zeigen, müssen bereits in der Designphase die Auswirkungen innovativer Aktivitäten mitbedacht werden: etwa als Privacy-by-design, Clean-by-design oder Ethics-by-design¹. Wiewohl die Einführung des X-by-design-Ansatzes überwiegend als Fortschritt wahrgenommen wird, dürfen möglicherweise entstehende administrative Bürden nicht ausgeblendet werden. In Österreich ist das X-by-design Thema bereits teilweise angekommen und sollte vertieft werden.

Überblick zum Thema

Der Ansatz X-by-design bedeutet, dass bestimmte Auswirkungen einer innovativen Aktivität (von Forschung über Wirtschaftsprozesse bis zur Gesetzgebung) bereits vom Start weg berücksichtigt werden. Der Gedanke des Design Thinking ist nicht neu,¹ macht aber mit Spezifikationen wie Privacy-by-design, Security-by-design, Safe-by-design, Clean-by-design, Law-by-design, Resilience-by-design, Quality-by-design oder Ethics-by-design erneut Karriere.² Im Folgenden sollen Ethics-by-design als grundlegendes gesellschaftspolitisches Thema und Privacy/Security-by-design als angewandte Themen im Zusammenhang mit Künstlicher Intelligenz (KI) beispielhaft charakterisiert und bewertet werden.

Derzeit werden vor allem sieben Prinzipien bzw. Prinzipiengruppen von Ethics-by-design diskutiert.³

*Ethics-by-design:
Allgemeine Prinzipien*

- Zum Prinzip „Freiheit und Würde“ gehören menschliche Handlungsermächtigung sowie positive und negative Freiheit.
- Das Prinzip „Robustheit und Sicherheit“ beinhaltet Resilienz, Absicherung („Plan B“), Verlässlichkeit und Reproduzierbarkeit.
- Beim Prinzip „Privatsphäre und Datenschutz“ sind der Respekt der Privatsphäre, die Datenqualität und die Datenrechte (Datenschutz-Grundverordnung DSGVO) zu nennen.
- Das Prinzip „Transparenz“ umfasst Nachverfolgbarkeit, Erklärbarkeit und Kommunikation.
- Das Prinzip „Vielfalt, Gleichbehandlung und Gerechtigkeit“ bezieht sich auf Nicht-Diskriminierung und Einbeziehung der Betroffenen.

¹ Vgl. Buchanan (1992), Sarwar & Fraser (2019).

² Vgl. OJEU (2016), EC-HLEG (2019), WEF (2020), UK-DCMS (2019/2021), WBG (2018), OECD (2020), CFDA (2019), EP (2020), EC (2021a), Brey et al. (2020), Brown et al. (2020), Rhomberg, Kienegger & Mattsson (2018), Jackson (2016), Yu et al. (2014).

³ Diese Prinzipien sind adaptiert von Brey et al. (2020) und EC-HLEG (2019). Da es ganz unterschiedliche Ethikansätze gibt, sei als Prüfstein der „kategorische Imperativ“ Kants angewandt: „Handle nur nach derjenigen Maxime, durch die du zugleich wollen kannst, dass sie ein allgemeines Gesetz werde.“ Kant (1968/1785: 421).

- Das Prinzip „Persönliches, gesellschaftliches und ökologisches Wohlergehen“ zielt auf sozialen Zusammenhalt, Demokratie und starke Institutionen ab.
- Das Prinzip „Verantwortung“ meint Nachvollziehbarkeit, Minimierung und Meldung negativer Auswirkungen, Schadenersatz und menschliche Beaufsichtigung.

Die Berücksichtigung dieser Liste ethischer Prinzipien, beziehungsweise deren Anpassung an die jeweilige Aufgabenstellung, ist angewandtes Ethics-by-design.

*KI: Ethics-by-design &
Privacy-by-design &
Security-by-design*

Im Zusammenhang mit KI folgt Privacy-by-design den oben dargestellten ethischen Prinzipien mit einem erweiterten Schwerpunkt auf Datenverarbeitung. Die EU-Kommission und das EU-Parlament haben kürzlich zwei Dokumente veröffentlicht, die sich mit den Rahmenbedingungen von KI auseinandersetzen und Ethics-by-design, Privacy-by-design und Security-by-design adressieren.

Die High-Level Expert Group KI (HLEG) der Kommission hält in ihrem Bericht fest, dass es enorm wichtig ist, die Vertrauenswürdigkeit von KI-Systemen sicherzustellen (EC-HLEG 2019): einerseits durch die Berücksichtigung der allgemeinen Ethics-by-design Prinzipien, die in diesem Dokument eng mit Privacy-by-design Prinzipien verbunden sind; andererseits durch Prinzipien, die eher Security-by-design zuzuordnen sind, wie:

- Hohe Sicherheit in KI-Prozessen, Daten und Ergebnissen.
- Robustheit von KI-Systemen gegenüber feindlichen Daten-Attacken.
- Gewährleistung von sicherem Abschalten und Wiedereinschalten, etwa nach einer Attacke.

Die Vorstellung der HLEG ist dabei, dass entsprechende Normen beim Design von KI-Systemen berücksichtigt werden sollen. Dies würde nicht nur der Bevölkerung dienen, sondern es würden auch Unternehmen profitieren, weil sie für die Auswirkungen ihrer Produkte und Dienstleistungen verantwortlich sind und durch den By-design-Ansatz negativen Auswirkungen vorbeugen können.

In diesem Zusammenhang hat die EU-Kommission am 21.4.2021 den Entwurf einer KI-Verordnung vorgelegt (EC 2021b). In diesem wird nicht die Technologie selbst, sondern die Verwendung von KI geregelt und es werden Designkriterien explizit angesprochen. Zwar wird der By-design-Ansatz nicht erwähnt, jedoch eine Reihe harmonisierter Regeln, die auf das Design, die Entwicklung und die Nutzung von (hoch-riskanten) KI-Systemen angewandt werden können (EC 2021b: 4). Des Weiteren kommen alle oben genannten ethischen Prinzipien vor (Freiheit und Würde, Robustheit und Sicherheit, Privatsphäre und Datenschutz etc.; den Prinzipien Transparenz und Kommunikation ist sogar ein eigener Abschnitt gewidmet). Interessanterweise werden diese Prinzipien nicht als „ethisch“ ausgewiesen. Allerdings wird bereits am Beginn des Textes auf den Bericht „Ethics Guidelines for trustworthy AI Artificial Intelligence“ (EC-HLEG 2019) hingewiesen und dessen Relevanz betont.

Darüber hinaus hat sich das EU-Parlament damit auseinandergesetzt, wie sich die Datenschutz-Grundverordnung (DSGVO) auf KI auswirkt (EP 2020). Dabei schreibt die DSGVO vor, dass Verantwortliche intern technische und organisatorische Vorsorge treffen müssen, um den Grundsätzen des Datenschutzes durch Technik (data protection by design) und durch datenschutzfreundliche Voreinstellungen (data protection by default) Genüge zu tun (EC 2021b). Das Parlament stellt nun fest, dass die DSGVO-Maßnahmen, die Privacy-by-design betreffen, die Entwicklung von KI nicht behindern, solange sie korrekt implementiert werden – allerdings möglicherweise mit zusätzlichen Kosten verbunden sind.

By-design-Ansätze haben auf europäischer Ebene als Thema eindeutig ihren Platz gefunden (EC-HLEG 2019, EP 2020, EC 2021b). Auch auf globaler Ebene werden solche Ansätze zum Beispiel durch das World Economic Forum und die World Bank Group promotet (WEF 2020, WBG 2018). Auf beiden Ebenen wird der By-design-Ansatz durchgängig als Fortschritt dargestellt. In kritischen Diskussionen in Forschung und Industrie wird allerdings auch die Befürchtung geäußert, dass mit verpflichtenden By-design-Implementierungen weitere administrative Bürden (z. B. bei Forschungsanträgen) hinzukommen.

*By-design-Ansätze:
Bürde oder Fortschritt?*

Relevanz des Themas für das Parlament und für Österreich

By-design-Ansätze sind derzeit auf globaler und europäischer Ebene ein prominent diskutiertes Thema. Auch in Österreich ist dieses Thema angekommen.⁴ Österreich als EU-Mitgliedsland sollte eine eigenständige Position zum maßvollen Umgang mit By-design-Ansätzen erarbeiten (d. h. By-design-Ansätze gezielt fördern unter Minimierung administrativer Bürden). Das Parlament sollte speziell eine Positionierung in Bezug auf die rechtlichen Entwicklungen vornehmen, da dem Vorschlag zur EU KI-Verordnung weitere technologiespezifische Verordnungen folgen dürften, die ebenfalls By-design-Ansätze beinhalten.

Vorschlag weiteres Vorgehen

Der X-by-design-Ansatz hat gesellschafts- und technologiepolitisches Potential, das in Österreich erst teilweise wahrgenommen wird. Als nächster Schritt könnte eine temporäre Plattform eingerichtet werden, in der sich Stakeholder aus unterschiedlichen Bereichen (Forschung & Forschungsförderung, Industrie, Ressorts, Sozialpartner) den Stand der Diskussion zum X-by-design-Ansatz aneignen, etwaige Handlungsbedarfe diskutieren und ein (kurzes) Positionspapier zu erstellen. Zur Unterstützung dieser Plattform sollte eine vergleichende Studie zu den unterschiedlichen X-by-design-Ansätzen durchgeführt werden, die die jeweiligen konzeptionellen und methodischen Hintergründe herausarbeitet. Auf dieser Basis kann dann die weitere Vorgehensweise zur praktischen Umsetzung entschieden werden.

⁴ Zum Beispiel im Zusammenhang mit Nanotechnologie (Rhombert, Kienegger & Mattsson 2018).

Zitierte Literatur

- Brey, P., et al. (2020) Guidelines for the ethical development of AI and big data systems: An ethics by design approach. SHERPA Consortium. project-sherpa.eu/deliverables/.
- Brown, B., Buchanan, R., DiSalvo, C., Doordan, D., Lee, K. & Mazé, R. (Eds.) (2020) The rise of legal design: Design Issues. MIT Press. doi.org/10.1162/DESI_e_00243.
- Buchanan, R. (1992) Wicked problems in design thinking. Design Issues, 8(2), 5-21. <https://doi.org/10.2307/1511637>
- CFDA (2019) Guide to sustainable strategies. Council of Fashion Designers of America. New York.
- EC (2021a) EU grants: How to complete your ethics self-assessment. European Commission. Brussels.
- EC (2021b) Proposal for a regulation laying down harmonized rules on artificial intelligence. European Commission. Brussels.
- EC-HLEG (2019) Ethics guidelines for trustworthy AI Artificial Intelligence. European Commission High-Level Expert Group. Brussels.
- EP (2020) The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. European Parliament EPRS STOA. Brussels.
- Jackson, S. (2016) Principles for resilient design: A guide for understanding and implementation. In: Florin, M.-V. & Linkov, I. (Eds.), IRGC resource guide on resilience Vol. 1. Lausanne: EPFL International Risk Governance Center (IRGC), 129-135.
- Kant, I. (1968/1785) Grundlegung zur Metaphysik der Sitten (Kant Werke Akademie Textausgabe Bd. IV). Berlin: Gruyter.
- OECD (2020) Moving towards a safe(r) innovation approach (SIA) for more sustainable nanomaterials and nano-enabled products. Organisation for Economic Co-operation and Development. Paris.
- OJEU (2016) General data protection regulation GDPR, 2016/679. *OJ L 119*, 4.5.2016, p. 1–88, eur-lex.europa.eu/eli/reg/2016/679/oj.
- Rhomberg, W., Kienegger, M. & Mattsson, M.-O. (2018) Stärken- und Schwächenanalyse NANOREG Safe-by-design. AIT - BMVIT - BMNT. Wien.
- Sarwar, A. & Fraser, P. T. (2019) Explanations in design thinking: New directions for an obfuscated field. *The Journal of Design, Economics, and Innovation*, 5(4), 343-355.
- UK-DCMS (2019/2021) Secure by design: Improving the cyber security of consumer internet of things report. UK Department for Digital, Culture, Media & Sport. London.
- WBG (2018) Privacy by design: Current practices in Estonia, India and Austria. World Bank Group. Washington.
- WEF (2020) Ethics by design: An organizational approach to responsible use of technology. World Economic Forum. Geneva.
- Yu, L. X., et al. (2014) Understanding pharmaceutical quality by design. *The AAPS Journal*, 16(4), 771-783.