Biography: Niva Elkin-Koren is a Professor of Law at the University of Haifa, Faculty of Law and a Faculty Associate at the Berkman Klein Center at Harvard University. She is the Founding Director of the Haifa Center for Law & Technology (HCLT), a Co-Director of the Center for Cyber, Law and Policy. During 2009-2012 she served as Dean of the Faculty of Law at the University of Haifa.

Her research focuses on innovation policy and access to knowledge, digital governance, online intermediaries, and the legal implications of AI and big data. Prof. Elkin-Koren has been a Visiting Professor of Law at Harvard University, Columbia Law School, UCLA, NYU, George Washington University and Villanova University School of Law. She is the Chair of the Scientific Advisory Council, of the Alexander von Humboldt Institute for Internet and Society in Berlin, a member of the Executive Committee of Association for the Advancement of Teaching and Research in Intellectual Property (ATRIP), and a board member of the MIPLC Scientific Advisory Board of the Munich IP Law Center at the Max Planck Institute for Innovation and Competition. She is also a member of the editorial boards of the Journal of the Copyright Society (since 2009) the Journal of Information Policy (since 2010) and the Internet Policy Review (since 2016). Prof. Elkin-Koren received her LL.B from Tel-Aviv University Faculty of Law in 1989, her LL.M from Harvard Law School in 1991, and her S.J.D from Stanford Law School in 1995.

Abstract: Social media platforms are responsible for mediating public discourse online. Increasingly platforms are using Artificial Intelligence and machine learning (AI) to perform content moderation (e.g., matching users and content, adjudicating conflicting claims, or detecting unwarranted content). In a digital ecosystem governed by AI, we currently lack sufficient safeguards against the blocking of legitimate content. Moreover, we lack a space for negotiating meaning and for deliberating the legitimacy of particular speech. In this presentation, Elkin-Koren proposes to address AI-based content moderation by introducing an adversarial procedure. Algorithmic content moderation often seeks to optimize a single goal, such as removing copyright infringing materials as defined by right holders, or blocking hate speech. Meanwhile, other values of the public interest, such as fair use, or free speech, are often neglected. Contesting Algorithms introduces an adversarial design, which reflects conflicting interests, and thereby, offers a check on dominant removal systems. The presentation will introduce the strategy of Contesting Algorithms, discuss its promises and limitations, and demonstrate how regulatory measures could promote the development and implementation of this strategy in online content moderation.