

Polymer Informatics: Current Status and Critical Next Steps

Lihua Chen^a, Ghanshyam Pilania^b, Rohit Batra^c, Tran Doan Huan^a, Chiho Kim^a, Christopher Kuenneth^a, Rampi Ramprasad^{a,*}

^a*School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA*

^b*Materials Science and Technology Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

^c*Center for Nanoscale Materials, Argonne National Laboratory, Lemont, Illinois 60439, USA*

Abstract

Artificial intelligence (AI) based approaches are beginning to impact several domains of human life, science and technology. Polymer informatics is one such domain where AI and machine learning (ML) tools are being used in the efficient development, design and discovery of polymers. Surrogate models are trained on available polymer data for instant property prediction, allowing screening of promising polymer candidates with specific target property requirements. Questions regarding synthesizability, and potential (retro)synthesis steps to create a target polymer, are being explored using statistical means. Data-driven strategies to tackle unique challenges resulting from the extraordinary chemical and physical diversity of polymers at small and large scales are being explored. Other major hurdles for polymer informatics are the lack of widespread availability of curated and organized data, and approaches to create machine-readable representations that capture not just the structure of complex polymeric situations but also synthesis and processing conditions. Methods to solve inverse problems, wherein polymer recommendations are made using advanced AI algorithms that meet application targets, are being investigated. As various parts of the burgeoning polymer informatics ecosystem mature and become integrated, efficiency improvements, accelerated discoveries and increased productivity can result. Here, we review emergent components of this polymer informatics ecosystem and discuss imminent challenges and opportunities.

Keywords: Polymer Informatics; machine learning; deep learning; polymer design and discovery; polymer synthesis

Contents

1	Introduction	2
2	Data generation, acquisition and management	4
2.1	Scientific literature	4
2.2	High-throughput and autonomous computational agents	5
2.3	Hypothetical polymers	6
3	Polymer representation	7
4	Property prediction schemes	9
4.1	Linear/Non-linear regression	9
4.2	Multi-fidelity information fusion approaches	11
4.3	Deep neural networks	12
4.4	Polymer Genome online platform	12

*Corresponding author

Email address: rampi.ramprasad@mse.gatech.edu (Rampi Ramprasad)