

Digital Discovery

rsc.li/digitaldiscovery

The Royal Society of Chemistry is the world's leading chemistry community. Through our high impact journals and publications we connect the world with the chemical sciences and invest the profits back into the chemistry community.

IN THIS ISSUE

ISSN 2635-098X CODEN DDIIAI 2(6) 1633–2000 (2023)



Cover

See Frank X. Gu *et al.*, pp. 1697–1709. Image reproduced by permission of Jeffrey Watchorn, Samantha Stuart and Frank X. Gu from *Digital Discovery*, 2023, 2, 1697.



Inside cover

See Shu Huang and Jacqueline M. Cole, pp. 1710–1720. Image reproduced by permission of Shu Huang and Nan Tian, who used imagery from rawpixel.com from Freepik, from *Digital Discovery*, 2023, 2, 1710.

PERSPECTIVE

1644

What is missing in autonomous discovery: open challenges for the community

Phillip M. Maffettone,* Pascal Friederich,* Sterling G. Baird, Ben Blaiszik, Keith A. Brown, Stuart I. Campbell, Orion A. Cohen, Rebecca L. Davis, Ian T. Foster, Navid Haghmoradi, Mark Hereld, Howie Joreess, Nicole Jung, Ha-Kyung Kwon, Gabriella Pizzuto, Jacob Rintamaki, Casper Steinmann, Luca Torresi and Shijing Sun

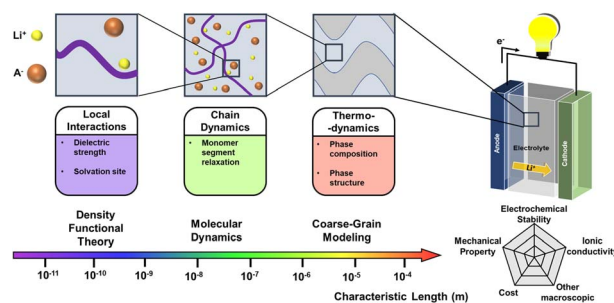


REVIEW

1660

Computational and data-driven modelling of solid polymer electrolytes

Kaiyang Wang, Haoyuan Shi, Tianjiao Li, Liming Zhao, Hanfeng Zhai, Deepa Korani and Jingjie Yeo*



Editorial Staff

Editor

Anna Rulka

Deputy Editor

Audra Taylor

Editorial Production Manager

Viktoria Titmus

Assistant Editors

Shwetha Krishna, Angelica-Jane Onyekwere, Michael Whitelaw, Alexander Whiteside

Editorial Assistant

Samantha Campos

Publishing Assistant

Brittany Hanlon

Publisher

Neil Hammond

For queries about submitted articles please contact Viktoria Titmus, Editorial Production Manager in the first instance. E-mail digitaldiscovery@rsc.org

For pre-submission queries please contact Anna Rulka, Editor.

Email digitaldiscovery-rsc@rsc.org

Digital Discovery (electronic: ISSN 2635-098X) is published 6 times a year by the Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, UK CB4 0WF.

Digital Discovery is a Gold Open Access journal and all articles are free to read. Please email orders@rsc.org to register your interest or contact Royal Society of Chemistry Order Department, Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, CB4 0WF, UK Tel +44 (0)1223 432398; E-mail: orders@rsc.org

Whilst this material has been produced with all due care, the Royal Society of Chemistry cannot be held responsible or liable for its accuracy and completeness, nor for any consequences arising from any errors or the use of the information contained in this publication. The publication of advertisements does not constitute any endorsement by the Royal Society of Chemistry or Authors of any products advertised. The views and opinions advanced by contributors do not necessarily reflect those of the Royal Society of Chemistry which shall not be liable for any resulting loss or damage arising as a result of reliance upon this material. The Royal Society of Chemistry is a charity, registered in England and Wales, Number 207890, and a company incorporated in England by Royal Charter (Registered No. RC000524), registered office: Burlington House, Piccadilly, London W1J 0BA, UK, Telephone: +44 (0) 207 4378 6556.

Advertisement sales:

Tel +44 (0) 1223 432246; Fax +44 (0) 1223 426017;

E-mail advertising@rsc.org

For marketing opportunities relating to this journal, contact marketing@rsc.org

Digital Discovery

rsc.li/digitaldiscovery

Digital Discovery is a gold open access journal publishing top research at the intersection of chemistry, materials science and biotechnology. Blurring the barriers between computation and experimentation, we focus on the integration of digital and automation tools with science, putting data first to ensure reproducibility and faster progress.

Editorial Board

Editor in Chief

Alán Aspuru-Guzik, University of Toronto, Canada

Associate Editors

Jason E. Hein, University of British Columbia, Canada
Linda Hung, Toyota Research Institute, USA
Joshua Schrier, Fordham University, USA
Kedar Hippalgaonkar, Nanyang Technological University, Singapore
Cesar de la Fuente, University of Pennsylvania, USA

Members

Yousung Jung, KAIST, South Korea
Anat Milo, Ben-Gurion University of the Negev, Israel
Lilo D. Pozzo, University of Washington, USA
Ekaterina Skorb, ITMO University, Russia

Advisory Board

Juan Alegre, Colorado State University, USA
Silvana Botti, Friedrich Schiller University Jena, Germany
Pablo Carbonell, University of Valencia, Spain
Cecilia Clementi, Freie Universität Berlin, Germany
Conor Coley, MIT, USA

Abigail Doyle, University of California Los Angeles, USA
Ola Engkvist, AstraZeneca and Chalmers University of Technology, Sweden
Ian Foster, University of Chicago, USA
Jan Jensen, University of Copenhagen, Denmark
Heather Kulik, MIT, USA

Shuye Ping Ong, University of California San Diego, USA
Marwin Segler, Microsoft, Germany
Berend Smit, EPFL, Switzerland
Isao Tanaka, Kyoto University, Japan
Alexandre Tkatchenko, University of Luxembourg, Luxembourg
Koji Tsuda, The University of Tokyo, Japan

Information for Authors

Full details on how to submit material for publication in Digital Discovery are given in the Instructions for Authors (available from <http://www.rsc.org/authors>). Submissions should be made via the journal's homepage: rsc.li/digitaldiscovery

Authors may reproduce/republish portions of their published contribution without seeking permission from the Royal Society of Chemistry, provided that any such republication is accompanied by an acknowledgement in the form: (Original Citation)–Reproduced by permission of the Royal Society of Chemistry.

This journal is © The Royal Society of Chemistry 2023.

Apart from fair dealing for the purposes of research or private study for non-commercial purposes, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988 and the Copyright and Related Rights Regulation 2003, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the Publishers or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency in the UK. US copyright law is applicable to users in the USA.

Registered charity number: 207890

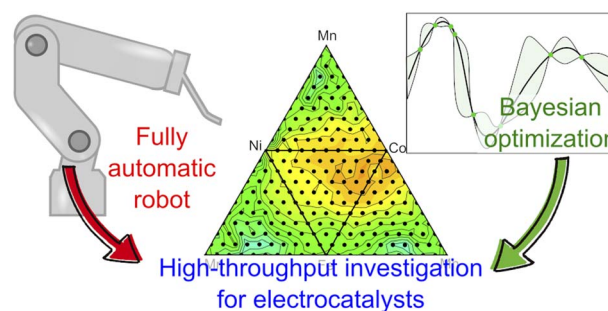


COMMUNICATIONS

1683

An automatic robot system for machine learning-assisted high-throughput screening of composite electrocatalysts

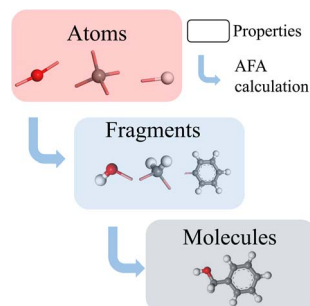
Masanori Kodera* and Kazuhiro Sayama*



1688

Atomic fragment approximation from a tensor network

Haoxiang Lin and Xi Zhu*

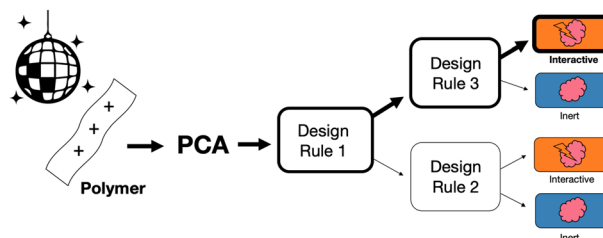


PAPERS

1697

An interpretable machine learning framework for modelling macromolecular interaction mechanisms with nuclear magnetic resonance

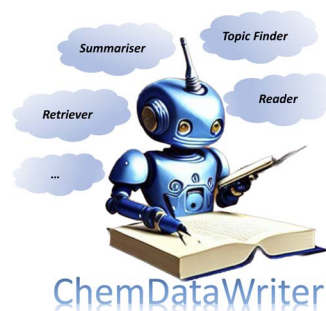
Samantha Stuart, Jeffrey Watchorn and Frank X. Gu*



1710

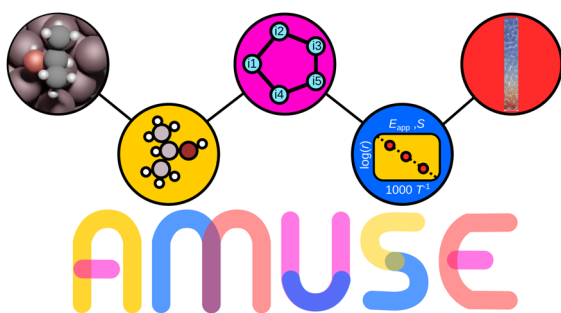
ChemDataWriter: a transformer-based toolkit for auto-generating books that summarise research

Shu Huang and Jacqueline M. Cole*



PAPERS

1721

**Automated Multiscale simulation environment**

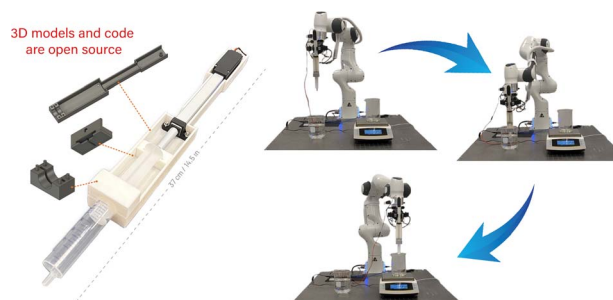
Albert Sabadell-Rendón,^{*} Kamila Kaźmierczak, Santiago Morandi, Florian Euzenat, Daniel Curulla-Ferré and Núria López^{*}

1733

**Autonomous biomimetic solid dispensing using a dual-arm robotic manipulator**

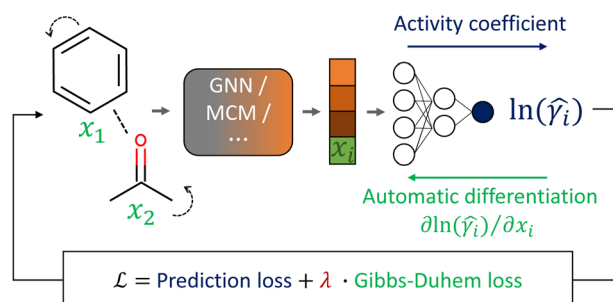
Ying Jiang, Hatem Fakhroldeen, Gabriella Pizzuto, Louis Longley, Ai He, Tianwei Dai, Rob Clowes, Nicola Rankin and Andrew I. Cooper^{*}

1745

**Digital pipette: open hardware for liquid transfer in self-driving laboratories**

Naruki Yoshikawa,^{*} Kourosh Darvish, Mohammad Ghazi Vakili, Animesh Garg^{*} and Alán Aspuru-Guzik^{*}

1752

**Gibbs–Duhem-informed neural networks for binary activity coefficient prediction**

Jan G. Rittig, Kobi C. Felton, Alexei A. Lapkin and Alexander Mitsos^{*}

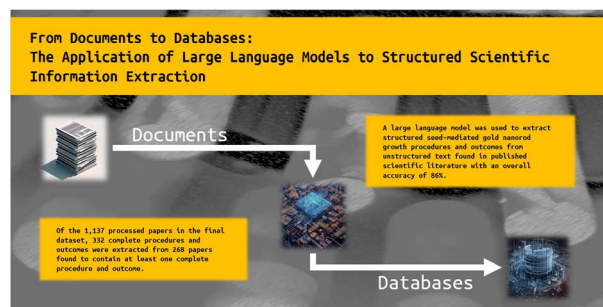


PAPERS

1768

Extracting structured seed-mediated gold nanorod growth procedures from scientific text with LLMs

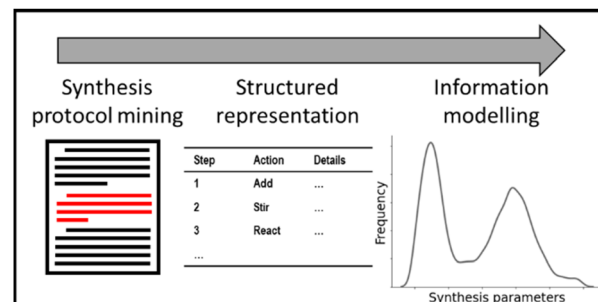
Nicholas Walker,^{*} Sanghoon Lee, John Dagdelen, Kevin Cruse, Samuel Gleason, Alexander Dunn, Gerbrand Ceder, A. Paul Alivisatos, Kristin A. Persson and Anubhav Jain^{*}



1783

Unveiling the synthesis patterns of nanomaterials: a text mining and meta-analysis approach with ZIF-8 as a case study

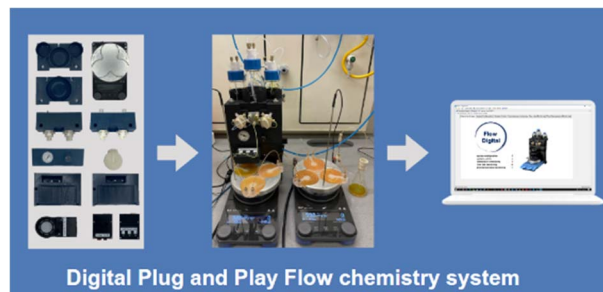
Joseph R. H. Manning^{*} and Lev Sarkisov^{*}



1797

Digitisation of a modular plug and play 3D printed continuous flow system for chemical synthesis

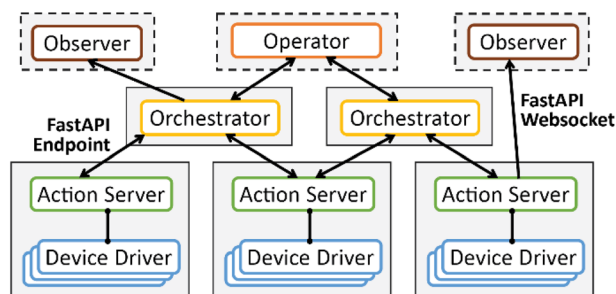
Mireia Benito Montaner, Matthew R. Penny and Stephen T. Hilton^{*}



1806

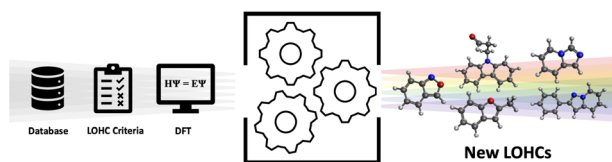
Orchestrating nimble experiments across interconnected labs

Dan Guevarra,^{*} Kevin Kan, Yungchieh Lai, Ryan J. R. Jones, Lan Zhou, Phillip Donnelly, Matthias Richter, Helge S. Stein and John M. Gregoire^{*}



PAPERS

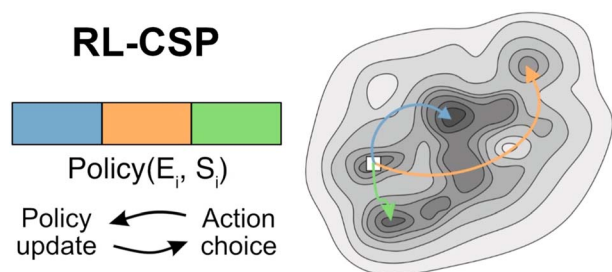
1813



Uncovering novel liquid organic hydrogen carriers: a systematic exploration of chemical compound space using cheminformatics and quantum chemical methods

Hassan Harb, Sarah N. Elliott, Logan Ward, Ian T. Foster, Stephen J. Klippenstein, Larry A. Curtiss and Rajeev Surendran Assary*

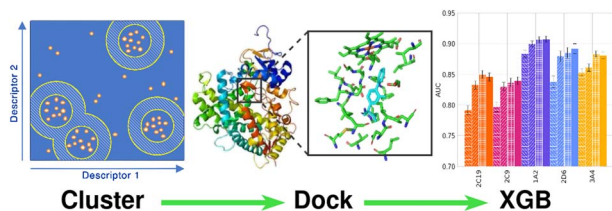
1831



Reinforcement learning in crystal structure prediction

Elena Zamaraeva, Christopher M. Collins, Dmytro Antypov, Vladimir V. Gusev, Rahul Savani,* Matthew S. Dyer, George R. Darling, Igor Potapov, Matthew J. Rosseinsky* and Paul G. Spirakis

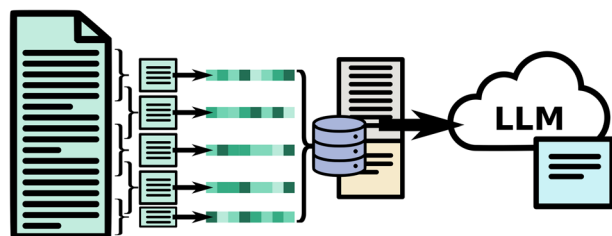
1841



Machine learning-augmented docking. 1. CYP inhibition prediction

Benjamin Weiser,* Jérôme Genzling, Mihai Burai-Patrasu, Ophélie Rostaing and Nicolas Moitessier*

1850



Domain-specific chatbots for science using embeddings

Kevin G. Yager*

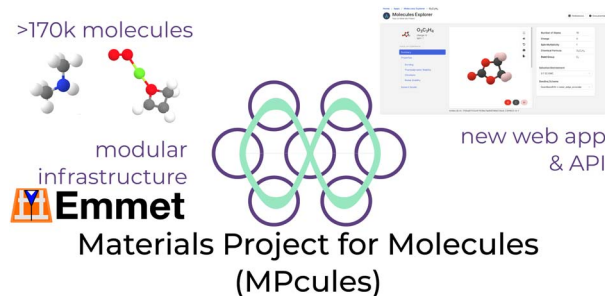


PAPERS

1862

A database of molecular properties integrated in the Materials Project

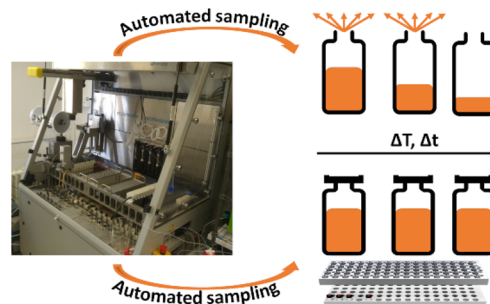
Evan Walter Clark Spotte-Smith,^{*} Orion Archer Cohen, Samuel M. Blau, Jason M. Munro, Ruoxi Yang, Rishabh D. Guha, Hetal D. Patel, Sudarshan Vijay, Patrick Huck, Ryan Kingsbury, Matthew K. Horton and Kristin A. Persson^{*}



1883

Best practice for sampling in automated parallel synthesizers

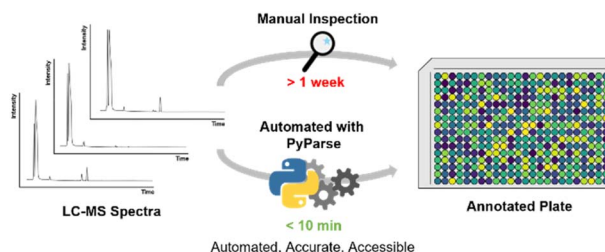
Michael Ringleb, Timo Schuett, Stefan Zechel and Ulrich S. Schubert^{*}



1894

Automated LC-MS analysis and data extraction for high-throughput chemistry

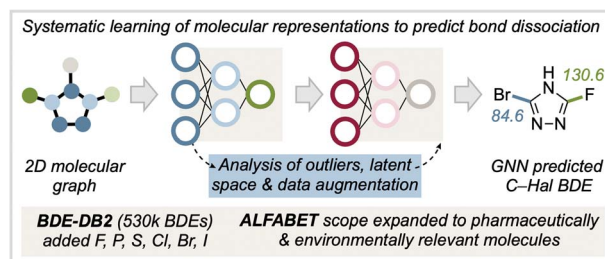
Joseph Mason,^{*} Harry Wilders, David J. Fallon, Ross P. Thomas, Jacob T. Bush, Nicholas C. O. Tomkinson and Francesco Rianjongdee



1900

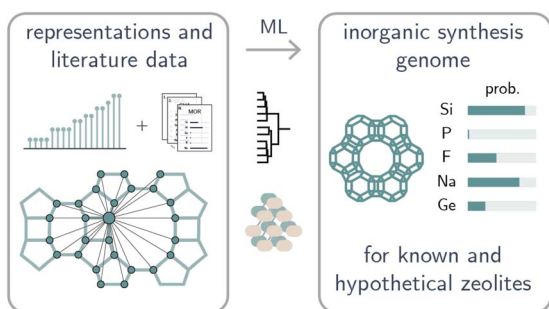
Expansion of bond dissociation prediction with machine learning to medicinally and environmentally relevant chemical space

Shree Sowndarya S. V., Yeonjoon Kim, Seonah Kim,^{*} Peter C. St. John^{*} and Robert S. Paton^{*}



PAPERS

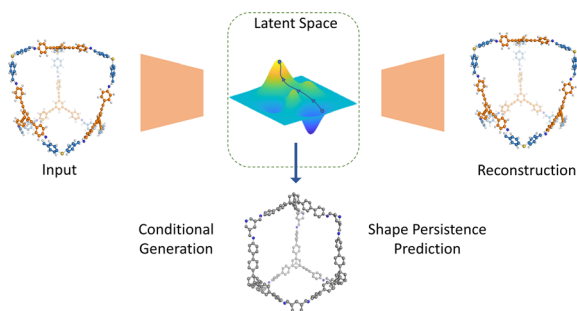
1911



Inorganic synthesis-structure maps in zeolites with machine learning and crystallographic distances

Daniel Schwalbe-Koda,^{*} Daniel E. Widdowson, Tuan Anh Pham and Vitaliy A. Kurlin^{*}

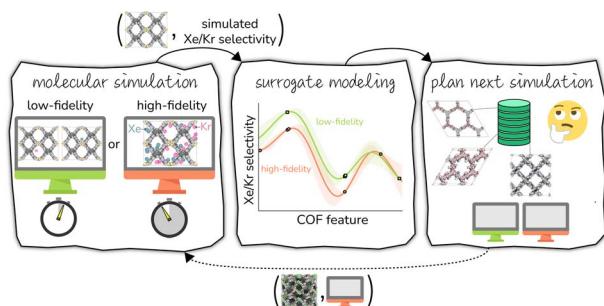
1925



Deep generative design of porous organic cages via a variational autoencoder

Jiajun Zhou, Austin Mroz and Kim E. Jelfs^{*}

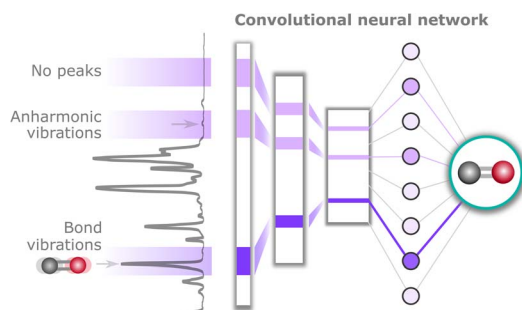
1937



Multi-fidelity Bayesian optimization of covalent organic frameworks for xenon/krypton separations

Nickolas Gantzler, Aryan Deshwal, Janardhan Rao Doppa^{*} and Cory M. Simon^{*}

1957



Understanding the patterns that neural networks learn from chemical spectra

Laura Hannemose Rieger, Max Wilson, Tejs Vegge and Eibar Flores^{*}

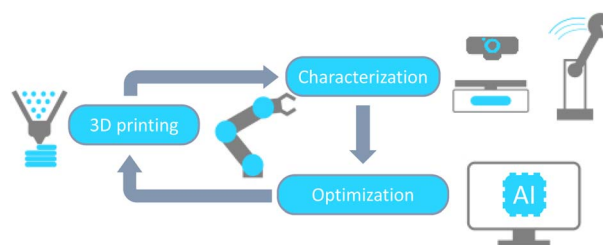


PAPERS

1969

Robotically automated 3D printing and testing of thermoplastic material specimens

Miguel Hernández-del-Valle, Christina Schenk, Lucía Echevarría-Pastrana, Burcu Ozdemir, Enrique Dios-Lázaro, Jorge Ilarraza-Zuazo, De-Yi Wang and Maciej Haranczyk*



1980

Towards a modular architecture for science factories

Rafael Vescovi, Tobias Ginsburg, Kyle Hippe, Doga Ozgulbas, Casey Stone, Abraham Stroka, Rory Butler, Ben Blaiszik, Tom Brettin, Kyle Chard, Mark Hereld, Arvind Ramanathan, Rick Stevens, Aikaterini Vriza, Jie Xu, Qingteng Zhang and Ian Foster*

