

# Digital Transformation in Materials Science: A Paradigm Change in Material's Development

Julian Kimmig, Stefan Zechel, and Ulrich S. Schubert\*

The ongoing digitalization is rapidly changing and will further revolutionize all parts of life. This statement is currently omnipresent in the media as well as in the scientific community; however, the exact consequences of the proceeding digitalization for the field of materials science in general and the way research will be performed in the future are still unclear. There are first promising examples featuring the potential to change discovery and development approaches toward new materials. Nevertheless, a wide range of open questions have to be solved in order to enable the so-called digital-supported material research. The current state-of-the-art, the present and future challenges, as well as the resulting perspectives for materials science are described.

be performed in the future. Thus, here, we will focus on the current developments in terms of automation and digitalization in all fields of material research (covering all material classes), evaluate the current and future challenges, and try to draw a picture of a potential digital material research future, which is beginning to develop. For this purpose, we will focus on the different aspects of material development starting from the experiment itself over the characterization and analysis of the data and finally to the fabrication and application of the obtained materials.

## 1. Introduction

Artificial intelligence (AI), automation, deep learning, big data, machine learning: these buzzwords are dominating the media, in news articles and are already connected to nearly every field and application (see **Table 1** for definition of the most important terms). Consequently, it is not surprising that science in general is not excluded from this trend and there are already first promising results.<sup>[1]</sup> Exemplarily, a deep learning software can be utilized for the identification of specific genetic diseases,<sup>[2]</sup> artificial intelligence can be utilized for the interpretation of images in clinical routines<sup>[3,4]</sup> or the development of new drugs,<sup>[5]</sup> big data handling simplifies the understanding of geographic questions<sup>[6]</sup> or cognitive computing enhances the speed of life science research.<sup>[7]</sup> All these examples show the great potential of the ongoing digitalization for significant advancements in science and one can simply conclude that this process will also influence materials science and the way research will

## 2. Preparation of Materials


The first and one of the most important steps in the development of new materials is the synthesis, formulation or preparation of new compounds. Typically, this step marks the beginning of material development and is followed by characterization (after a potential purification) and interpretation of the obtained data.

In general, new materials are obtained by a chemical reaction, a modification or by formulation/blending of different substances. The reaction conditions must be chosen accordingly to obtain the material of choice with the appropriate properties. Typically in today's research, these syntheses are performed by human, that is, scientists or technicians, who utilize a wide range of different methods for the different classes of materials. While in the field of polymers mostly classical organic synthesis protocols are utilized,<sup>[8]</sup> there are also formulation such as sintering steps (e.g., for ceramics),<sup>[9]</sup> melting processes,<sup>[10]</sup> or sol-gel processes<sup>[11]</sup> for other material classes, for example, metals, respectively.

This approach has the great disadvantage that the results and the materials obtained depend strongly on the researcher performing the experiments and on how much experience one has. For this reason, initial approaches in the digital transformation of material research are focusing mostly on the automation, parallelization, and miniaturization of the synthesis and the development of opportunities for high-throughput processes.<sup>[12]</sup> Two priority methods are robot-based synthesis<sup>[13]</sup> and flow chemistry<sup>[14]</sup> (including microfluidics<sup>[15]</sup>). The former one is based on the utilization of synthesis robots, which can perform a variety of experiments simultaneously and with high precision as well as high reproducibility. In contrast, in flow chemistry microreactors, for instance, are utilized allowing a corresponding high number of experiments due to the parallel experiments as well as a relatively low amount of required

J. Kimmig, Dr. S. Zechel, Prof. U. S. Schubert  
Laboratory of Organic and Macromolecular Chemistry (IOMC)  
Friedrich Schiller University Jena  
Humboldtstr. 10, Jena 07743, Germany  
E-mail: ulrich.schubert@uni-jena.de

J. Kimmig, Dr. S. Zechel, Prof. U. S. Schubert  
Jena Center for Soft Matter (JCSM)  
Friedrich Schiller University Jena  
Philosophenweg 7, Jena 07743, Germany

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adma.202004940>.

© 2021 The Authors. Advanced Materials published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adma.202004940

**Table 1.** Definition of often-utilized terms for the field of automation and digitalization.

Term	Short definition <sup>a)</sup>
Artificial intelligence (AI)	"The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."
Machine learning	"The use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data."
Big data	"Extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions."
Deep learning	"A type of machine learning based on artificial neural networks in which multiple layers of processing are used to extract progressively higher level features from data."

<sup>a)</sup>Definitions were obtained from English Oxford Living Dictionary.

compounds.<sup>[16–18]</sup> Both approaches can also be combined as shown in a very recent study.<sup>[19]</sup> Herein, a robotic arm is utilized to enable flow chemistry approaches. Additionally, the system features an AI-software enabling an improvement in the experimental planning. However, this system can only be utilized for organic synthesis protocols and cannot easily be adapted for other material classes such as metals or ceramics.

These two approaches have been known for a while,<sup>[8]</sup> but they have progressed differently in the different material classes. Consequently, synthetic robots are well-known in the field of polymer synthesis as well as polymer formulation and have been used for a variety of processes.<sup>[20,21]</sup> However, most of the syntheses performed nowadays are still classical ones using flasks and other classical laboratory glassware. In addition, catalysis represents a research field with a strong utilization of combinatorial approaches.<sup>[22]</sup> Also in other material classes there are first examples in this direction,<sup>[23]</sup> for example, in the perovskite development<sup>[24]</sup> or alloys,<sup>[25]</sup> nevertheless, the automated synthesis/formulation has not yet been fully enforced and many syntheses and material manufacturing processes are still performed in a classical manner. This is also due to one of the disadvantages of the robotic approach, which are high costs of the required equipment. However, this investment would be a one-time investment, which can potentially be cheaper on a long-term perspective compared to paying scientist.

An increasing trend toward automated synthesis is recognizable, since it not only increases the number of experiments, it further generates more reliable data not influenced by human bias. Each experiment is traceable, the reaction conditions can be read out as a protocol (this data is thus usable, see paragraph data management) and a new experiment under the same conditions provides comparable results (also in different labs as long as the same starting materials are utilized). Therefore, it is not surprising that companies have recognized this trend and have heavily invested in the automation of syntheses/formulation for their research and development departments.<sup>[26,27]</sup>

The challenges in the field of automation and digitalization in the area of material preparation are to be sought above all in the improvements of the robotic systems and processes. In particular, the robotic systems have to be optimized to cover

all different reaction conditions, which are applied in material research (temperature, pressure, type of material etc.). Exemplarily, high-temperature processes have to be mentioned, which is still a challenge using automated synthesis robots. Moreover, the required compounds (liquids, solids, pastes, oils etc.) have to be transferred from bottles, boxes, cylinders etc. into the reaction vial in the right volume, mass using syringes or balances. A novel revolutionary approach using "pins" (or "coins") has been presented in order to overcome the this bottleneck of current robotic based synthesis.<sup>[28]</sup>

Furthermore, multi-step reaction procedures have to be performed automatically.<sup>[29]</sup> In addition, the purification and characterization of the materials must be performed in such an automated fashion. First approaches to this are well-known in the field of organic synthesis and polymer science.<sup>[30]</sup> Thus, the group of Cronin has combined various purification and analysis methods with synthesis modules.<sup>[31,32]</sup> Transferring this approach to the various material classes and characterization methods will be a crucial challenge for the future of the digital transformation of material research.

Recently, a synthetic robot, which can move through a laboratory collecting samples, placing them into analysis equipment, such as gas chromatography or high-performance liquid chromatography, was described by the Cooper group (**Figure 1**).<sup>[33]</sup> Herein, the robot was improved using a learning algorithm in order to search for photocatalysts. For this purpose, the robot performed more than 600 experiments and was able to find a better working catalyst compared to the initial starting material by using a learning algorithm.

### 3. Characterization of Materials

The characterization and analysis of materials have been performed already for a very long time using automated equipment, for example, with auto samplers, well plates, etc. Nonetheless, there is an urgent need for further development in order to facilitate effective digitalization and automation, since the characterization represents a central element for the generation of the (measurement) data.<sup>[34]</sup> For this purpose, also additional on-demand and online approaches are required. A major disadvantage for a uniform procedure across all material classes in the development of the analysis methods is the great variety of the materials, the investigated properties as well as the techniques themselves (provided by different companies and utilizing different software tools). Here, even within one class of material, for example, polymers, one can find a wide range of techniques that are regularly applied and used. This challenge multiplies enormously across the different material classes and makes increasing automation and digitalization even more difficult with increasing networking between the devices.

Generally, the requirements profile of analytical methods changes to a faster characterization at a constant quality standard, enabling, for example, direct real-time (kinetic) investigations. This is usually associated with the term of high-throughput characterization in which a variety of samples can be analyzed within a short timeframe (also online and inline), which are known for different material classes, for example, polymers, alloys or hydrogels.<sup>[35–37]</sup> The increasing need for faster characterization is also due to the increasing amount



**Figure 1.** A robotic mobile chemist at work filling vials and loading the gas chromatography equipment. Reproduced with permission.<sup>[33]</sup> Copyright 2020, Springer Nature.

of data required for machine-learning algorithms, as they can only work effectively if sufficient data is generated and implemented.<sup>[38]</sup> The increasing automation of synthesis will increase the need for new (faster) characterization approaches.

Nevertheless, the general procedure for the characterization of new compounds applied so far consists of the sample collection during synthesis and the subsequent transfer to the characterization device by hand. Afterward, the measurement data are evaluated independently by the scientist. However, this procedure must be optimized so that a more efficient and faster analysis is possible and bottlenecks in the research can be eliminated. Variants for this are, for example, the online characterization during the reaction,<sup>[39]</sup> the combination of synthesis, purification, and characterization (see earlier)<sup>[31,32,40]</sup> and online/inline flow-characterization approaches.<sup>[41]</sup> All of these strategies are already available to a certain extent, however, they are rarely utilized, (combined) expensive and need to further be developed.

Finally, another important strategy in the field of materials analysis is the automatic evaluation of the obtained results. Mostly nowadays the results are evaluated by hand by the scientist. To optimize this process, increasingly automatic evaluation tools are applied. Exemplarily, the machine-based image recognition (e.g., transmission electron microscopy analysis)<sup>[42]</sup> or the automatic analysis of spectra (e.g., nuclear magnetic resonance)<sup>[41]</sup> have to be named.

In particular, the aspect of the accuracy of the analysis should not be lost during all automation digitalization and time-related improvements and the obtained information should not change during improvements. Only trustworthy measurement data can be utilized by machine learning and artificial intelligence. If

this would not be the case, such programs would predict new experiments and properties based on wrong data resulting in incorrect predications.

#### 4. Management and Analysis of Data

The collection, analysis and handling as well as the interpretation of data represents the central element in the field of digitalization in material research. However, these aspects have received the least attention so far. There are frequently reports on the utilization of artificial intelligence etc.,<sup>[1,38,43]</sup> but very little on the storage and (automated) interpretation of data leading to new knowledge and an accelerated materials discovery. However, they are the key elements, as shown, for example, by Amazon and others for their purposes, but not yet for materials science. The data are required to train software programs in the field of machine learning and artificial intelligence enabling automated routines to provide personalized solutions and suggestions. A first initiative that has recognized the potential in this field was the material genome project of the US government<sup>[44]</sup> and later the analog initiative in China.<sup>[45]</sup> In doing so, the competencies in the area of material research should be bundled and by a combined effort it should finally be possible to decrypt the “material code”<sup>[1]</sup> opening the field of material informatics, “materialomics” or “polymeromics.”<sup>[46]</sup>

Currently, most of the research data is generated locally in academic and industrial laboratories and only there the primary data are accessible, stored in different formats at different storage location, which are partially offline. Furthermore, even in

the corresponding scientific publications only selected data are presented, typically only the data from successful experiments.<sup>[47]</sup> Thus, only very sparse, the really relevant data are shared within the scientific community. The quality of the provided data is sometimes questionable and the reliability is not given in every case. Another challenge in this context is the “negative” data, which are data obtained from experiments not resulting in the desired material.<sup>[48]</sup> These data are hardly published these days,<sup>[47]</sup> or even stored in a useful manner. Of course, this raises the question of how many experiments have already been carried out meaninglessly several times in the world, even though it has already been observed that the experiments are not promising.

In order to merge all data, positive and negative ones, a comprehensive database structure for all data of the material research is required. Meta-analysis, as known from other disciplines,<sup>[49]</sup> are nowadays hardly possible and higher-level structure-property relationships can only be revealed with great difficulty. In addition, the design and planning of new experiments are not optimal.

One solution to this problem is provided by the FAIR principle of data handling.<sup>[50]</sup> Thus, the data are findable, accessible, interoperable and reusable. Consequently, previously used experiments and the associated data can also be reutilized again. A central element for storing the data is required, which can be offered by a repository. All original data are stored in a meaningful way applying the FAIR principles. Consequently, all these data can be (re)used. Finally, new results can be generated based on existing measurements, and meaningless experiments are not performed multiple times. Moreover, less promising procedures can be later improved in an automated manner. Such an example can be found in the field of metabolomics and is already known and established.<sup>[51]</sup> Herein, the measurement data in a specific format are stored centrally in a public repository and each scientist around the world can access and reuse those data. One approach regarding materials science is the material cloud,<sup>[52]</sup> for example, Aflowlib with over 3 000 000 million substances.<sup>[53]</sup> One current positive trend is that more and more journals already starts to expect the storage of primary data related to a publication. However, this procedure must be improved in the future and the data storage has to become obligatory for each publication.

Nevertheless, there will be further challenges<sup>[54,55]</sup> in the form of generally accessible data formats, the standardization of data and, consequently, the definition of information standards. Moreover, also already obtained knowledge must be converted into data formats in order to make them usable. As a basic requirement, scientists themselves must be convinced of such a system and students have to be trained accordingly. A first national step is currently being observed in Germany, where a national research database infrastructure is to be set up (e.g., the initiative NFDI4Chem<sup>[56]</sup> for chemistry related fields). However, global solutions are required. Moreover, data storage and management work flows must be generated, which are useable for all material classes.<sup>[57]</sup>

## 5. Planning of Experiment/Design of Experiments

The category of experimentation planning is, for most people today, most associated with the concept of digitalization.

Software-assisted synthesis planning and the development of new materials using machine-learning and artificial-intelligence approaches can become a key technology in the future of materials research.<sup>[43]</sup>

Classically, the planning of experiments is performed in such a manner that the scientist considers a new experiment based on literature findings and personal experience. This experiment is then performed and, based on the obtained data, a new experiment may be planned.

The first advances in this field are provided by simulations of reactions and material properties that allow a prediction of particular compositions and, thus, enable more targeted experiments.<sup>[58]</sup> In the first phase of the development of such simulations, the results were still rudimentary and could only partially be utilized for the prediction of properties. Meanwhile, the prediction quality have become much better (also using machine-learning techniques<sup>[59]</sup>) and simulations allow in certain areas, for example, metal-organic frameworks (MOFs),<sup>[60]</sup> the targeted development of new materials.<sup>[61–63]</sup> The simulation of properties and the corresponding structures becomes increasingly more important; however, it is still limited for specific material classes such as alloys<sup>[64]</sup> or MOFs.<sup>[65]</sup> In case of, for example, polymers, the modeling is still in its infancy and the calculation demand is too large to obtain exact predication.<sup>[66]</sup> In the future, simulations will gain more importance in all fields of materials and can also be utilized for the achievement of data, which can be further used for artificial intelligence programs.

The next step in a modern material research process is the utilization of machine learning and artificial intelligence to facilitate feedback-driven reaction and material planning.<sup>[38,43]</sup> These programs uses the data from previous experiments predicting a new targeted composition of a material, which can subsequently be synthesized. Alternatively, the measurement data may also be utilized for the improvement of the synthesis protocol and, consequently, the yield of the reaction. Finally, a sufficient amount of data can be generated by a large number of experiments resulting in the opportunity to reveal the ideal composition of a material for a certain application. Above all, it is important that sufficient data points are available (as the number of desired properties increases, the number of data required also increases) and that negative data attempts are also recorded,<sup>[67]</sup> since otherwise the desired parameter space cannot be defined in a clear manner.<sup>[38]</sup>

The development of such software-based methods for material development of new materials is still in its infancy. However, there are first examples, such as the use of machine-learning for the predication of material properties,<sup>[59,68]</sup> for example, battery materials,<sup>[69,70]</sup> the use of machine learning and modeling for the targeted selection and subsequent synthesis of conductive MOFs,<sup>[71,72]</sup> the predication of metallic glasses<sup>[73]</sup> or the targeted development of catalysts.<sup>[57]</sup>

The most significant challenge for machine-learning systems is the amount of data and the associated time (and resources) for the required experiments. The better the program should work and the more properties a particular material has, the more experiments are required. To solve this challenge, there are two possible solutions: a) the utilization of simulations<sup>[74]</sup> to predict material properties<sup>[75]</sup> (as well as the utilization system properties and similarities<sup>[76]</sup>) and b) the merging of all generated data in a corresponding repository from which the



software can obtain the required data. This procedure could reduce the number of required experiments and enable a more targeted material research in the future. However, at the present time, the application of simulation is limited to some material classes and a general (worldwide used) repository for materials research containing all classes of materials is not foreseeable. Accordingly, the current descriptions of machine-learning-based material systems are mostly based on the use of own data (partially evaluating the literature) and, thus, demonstrate only the principle feasibility of such systems.

## 6. Fabrication and Processing

The processing and fabrication of materials represent usually the final step in the exploration of future potential applications of the materials. The way of processing largely depends on the materials properties and the potential fields of application, whether the usage is, for example, for batteries, as a coating for anticorrosion or as construction element. Here, a variety of different techniques are already available, which can also run robotically supported, already (partially) applied industrially. The classic processes here are, besides many others, extruding, spraying (e.g., for coatings), roll-to-roll printing, injection molding, deep drawing, etc.

Nevertheless, the trend toward digital fabrication has been apparent for some time in order to further advance digitalization in this area enabling the processing of materials to the desired/personalized shape.<sup>[77]</sup> A benchmark technology here was 3D printing.<sup>[78]</sup> Producing materials in the desired 3D shape has been a technological revolution and is constantly evolving,<sup>[79]</sup> for example, 3D printing of cells resulting in fully functional human hearts components.<sup>[80]</sup> Additionally, there are also industrial applications known for 3D-printing. For example, parts of the Boeing 777 are produced by 3D-printing.<sup>[81]</sup> Meanwhile, almost all classes of materials are printable, although different techniques have to be applied, for example, fused layer molding, laser beam melting, screen printing, and two-photon-polymerization. The latest trend in digital fabrication is so called 4D printing, which involves printing materials that change their shape over time (fourth dimension).<sup>[82]</sup> A classic example of this behavior can be found for shape-memory materials, which are then also usable for various purposes, for example, medical applications.<sup>[83]</sup>

In the field of digital transformation of materials research, fabrication is probably the most advanced area, but there is still enormous potential for future developments. In particular, the printing of all kinds of materials (e.g., glass<sup>[84]</sup>) or the printing of high-performance materials are still at the beginning of research.<sup>[85]</sup> Furthermore, new developments regarding automated approaches combined with self-optimizing programs enable new possibilities, such as recently reported for film-preparation of solar cell materials.<sup>[86]</sup>

## 7. Conclusion—The Next Generation of Materials Research (Challenge and Possibilities)

The future of material research is becoming increasingly digital. However, this process will require significant efforts and

changes. Nevertheless, the first trends are already recognizable and there is progress in many areas. The different classes of materials are currently at different levels, which are also due to the fundamentally different methods of synthesis, producing, analyzing and simulating the materials.

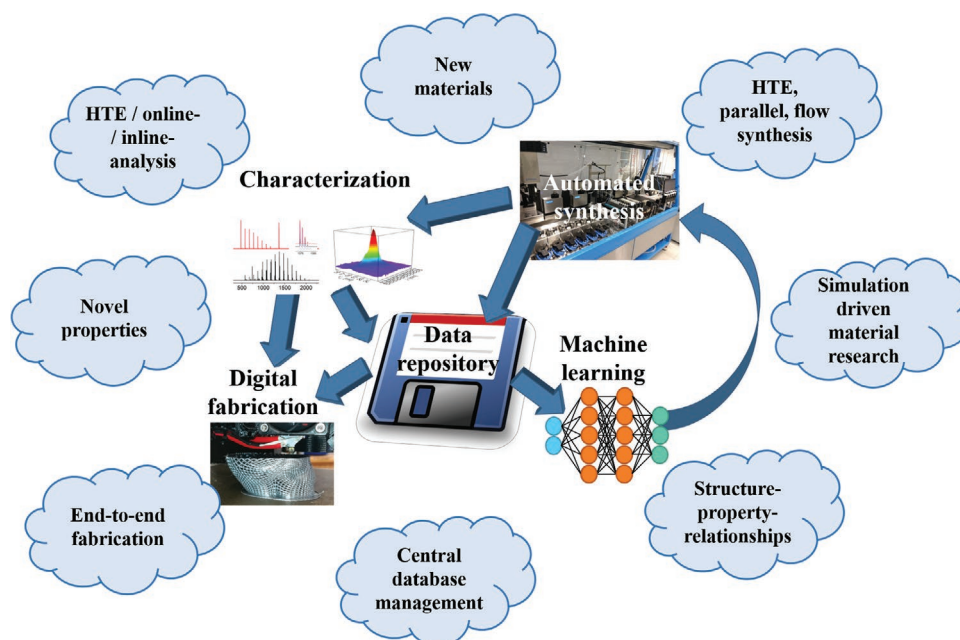
In general, substantial progress beyond the individual sub-areas is required in order to reach the next level. A combination of all areas—from synthesis to characterization, data management and machine learning to digital fabrication—is possible and required (**Figure 2**). Engineering a full chain of knowledge will enable a fully automated, digital, and robot-assisted material research and will change the way we discover, develop, and process new materials. Currently, only individual and isolated approaches are described.

As a consequence, a faster and higher reproducible development of materials will be possible. Moreover, the number of duplicates in experimentation will be reduced significantly and the human resource will have to do less synthesis work and can focus more on planning new materials and their properties in order to enable a targeted experimental design. Moreover, a higher number of experiments and a central storage of the required data will enable much more innovations resulting also in materials with properties, which are hardly imaginable up to now.

There are many challenges for achieving automated, digital, and robot-assisted materials research (summarized in **Table 2**). First of all, the further development of existing technologies has to be pushed forward, for example, robots and machine-learning. Herein, the programs of AI have to be modified to more suitable for materials science and the different requirements of the different classes of materials.

The establishment of a data repository is also crucial, as this is the only way to ensure utilization of the produced data by every scientist. Moreover, the data formats have to be defined in a standardized manner in order to have the same format across the world. Thus, all data can be reused sufficiently to evaluate general structure-property relationships. One of the greatest challenges, however, is the training of skilled students and scientist. Classically, scientists today receive training in materials research, chemistry, engineering or physics. In the future, knowledge in the field of computer science, database management, and robotics will be required as well. Integrating this content into teaching and training represents a key challenge for the future of digital material research and will result in a new generation of scientists.

Finally, the question of the perspective and the chance for scientists remains in such a construct. What tasks will you have, and which skills do you have to bring and what can be the benefit? First, it should be noted that the realization of a digital material research is still far away and many scientists are required in order to optimize and expand the individual areas until this vision can become reality. Herein, many new discoveries are required, which must be performed by experts. In addition, properties, which should be obtained at the very end, must also be specified by a scientist and a rough direction for the starting parameters of experiments must be guaranteed. Otherwise, the number of iterations is far too large. Finally, programs need to be developed further, robots have to be optimized and made available in a cost-efficient manner for the broad community of researchers, and the design of experiment (at least in the early



**Figure 2.** Schematic representation of a completely automated and digital-transformed material research combining automated synthesis, online/inline characterization, machine learning, and digital fabrication. All obtained data will be stored centrally in a data repository enabling new materials as well as general structure-property relationships.

stages) will remain a long-term task for scientists. The chance of a more digital material research lies in the higher productivity, the better reliability of the data and in a better understanding

of general structure-property relationships. Scientists can focus on the major part of the discovery of new materials, which is the designing of desired materials and their properties, since

**Table 2.** Summary of advantages, disadvantages as well as challenges of a more automated and digital material research.

Category	Advantages	Current disadvantages	Challenges
Synthesis	<ul style="list-style-type: none"> <li>– Faster</li> <li>– Automated</li> <li>– More reliable</li> </ul>	<ul style="list-style-type: none"> <li>– Expensive</li> <li>– Not applicable for all materials</li> <li>– Not accessible for all scientist</li> <li>– Required space for synthesis robots</li> </ul>	<ul style="list-style-type: none"> <li>– Access for all</li> <li>– More reaction parameters, miniaturization</li> <li>– Application for all materials</li> <li>– Purification of materials</li> </ul>
Characterization	<ul style="list-style-type: none"> <li>– Faster</li> <li>– Online and inline</li> </ul>	<ul style="list-style-type: none"> <li>– Reliability?</li> <li>– Accuracy of results</li> </ul>	<ul style="list-style-type: none"> <li>– Faster, but still reliable</li> <li>– Automated evaluation</li> <li>– Combining methods (also with synthesis)</li> </ul>
Machine learning	<ul style="list-style-type: none"> <li>– Automated optimization</li> <li>– Material design, which is not so easy imaginable</li> </ul>	<ul style="list-style-type: none"> <li>– Black box—hardly controllable</li> <li>– Loss of unexpected results</li> <li>– Number of iterations—time consuming experimentation</li> </ul>	<ul style="list-style-type: none"> <li>– Better programs</li> <li>– Data input must be ensured</li> <li>– Output parameters have to be defined</li> <li>– Establishment of simulation as data source</li> </ul>
Digital fabrication	<ul style="list-style-type: none"> <li>– Device of choice can be produced</li> <li>– Personalization</li> </ul>	<ul style="list-style-type: none"> <li>– Partially difficult for some materials</li> <li>– Not efficient for high number production</li> </ul>	<ul style="list-style-type: none"> <li>– More development, for example, 4D-printing also for other materials</li> <li>– More cost-efficient</li> <li>– Access to the devices</li> </ul>
Data handling	<ul style="list-style-type: none"> <li>– Central access to data</li> <li>– Negative results can be stored</li> <li>– Less duplicates of experiments</li> <li>– Reliable data</li> <li>– Meta-analysis</li> <li>– Structure-property-relationship</li> </ul>	<ul style="list-style-type: none"> <li>– Invalid data has to be filtered out</li> </ul>	<ul style="list-style-type: none"> <li>– Development of such a system</li> <li>– Data format must be uniform</li> <li>– Information standards have to be defined</li> <li>– Rethinking of scientist</li> </ul>
Teaching	<ul style="list-style-type: none"> <li>– Broader teaching</li> </ul>	<ul style="list-style-type: none"> <li>– Changed profile of teacher</li> <li>– Partially loss of details in academia teaching</li> </ul>	<ul style="list-style-type: none"> <li>– Combination of materials science and other areas (informatics, robotics)</li> <li>– Implementation into curriculum</li> </ul>

the synthesis, characterization, and optimization of the process (improvement of the synthetic procedure) will be performed automatically. Consequently, faster material design and more target-oriented research will be accessible.

## Acknowledgements

The authors would like to thank the German Research Foundation (DFG) for funding (SFB 1278 “PolyTarget,” project number 316213987, project Z01).

Open access funding enabled and organized by Projekt DEAL.

## Conflict of Interest

The authors declare no conflict of interest.

## Keywords

artificial intelligence, automation, combinatorial science, digital transformation, machine learning

Received: July 20, 2020  
Revised: September 1, 2020  
Published online:

- [1] N. Nosengo, *Nature* **2016**, 533, 22.
- [2] Y. Gurovich, Y. Hanani, O. Bar, G. Nadav, N. Fleischer, D. Gelbman, L. Basel-Salmon, P. M. Krawitz, S. B. Kamphausen, M. Zenker, L. M. Bird, K. W. Gripp, *Nat. Med.* **2019**, 25, 60.
- [3] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, H. J. W. L. Aerts, *Nat. Rev. Cancer* **2018**, 18, 500.
- [4] E. J. Topol, *Nat. Med.* **2019**, 25, 44.
- [5] S. Ekins, A. C. Puhl, K. M. Zorn, T. R. Lane, D. P. Russo, J. J. Klein, A. J. Hickey, A. M. Clark, *Nat. Mater.* **2019**, 18, 435.
- [6] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvahais, Prabhat, *Nature* **2019**, 566, 195.
- [7] Y. Chen, J. D. Elenee Argentinis, G. Weber, *Clin. Ther.* **2016**, 38, 688.
- [8] R. Hoogenboom, M. A. R. Meier, U. S. Schubert, *Macromol. Rapid Commun.* **2003**, 24, 15.
- [9] M. N. Rahaman, *Ceramic Processing and Sintering*, Vol. 2nd, Marcel Dekker, Inc., New York **2003**.
- [10] X. Liu, N. Fechler, M. Antonietti, *Chem. Soc. Rev.* **2013**, 42, 8237.
- [11] A. E. Danks, S. R. Hall, Z. Schnepf, *Mater. Horiz.* **2016**, 3, 91.
- [12] C. A. Tweedie, D. G. Anderson, R. Langer, K. J. Van Vliet, *Adv. Mater.* **2005**, 17, 2599.
- [13] W. F. Maier, G. Kirsten, M. Orschel, P.-A. Weiss, A. Holzwarth, J. Klein, *Macromol. Symp.* **2002**, 165, 1.
- [14] D. Webb, T. F. Jamison, *Chem. Sci.* **2010**, 1, 675.
- [15] K. S. Elvira, X. C. Solvas, R. C. R. Wootton, A. J. deMello, *Nat. Chem.* **2013**, 5, 905.
- [16] C. Mateos, M. J. Nieves-Remacha, J. A. Rincón, *React. Chem. Eng.* **2019**, 4, 1536.
- [17] N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo, X. Ye, A. J. Mijalis, M. D. Simon, A. Loas, S. Liu, C. Jessen, T. E. Nielsen, B. L. Pentelute, *Science* **2020**, 368, 980.
- [18] S. Chatterjee, M. Guidi, P. H. Seeberger, K. Gilmore, *Nature* **2020**, 579, 379.
- [19] C. W. Coley, D. A. Thomas, J. A. M. Lummiss, J. N. Jaworski, C. P. Breen, V. Schultz, T. Hart, J. S. Fishman, L. Rogers, H. Gao, R. W. Hicklin, P. P. Plehiers, J. Byington, J. S. Piotti, W. H. Green, A. J. Hart, T. F. Jamison, K. F. Jensen, *Science* **2019**, 365, eaax1566.
- [20] S. Oliver, L. Zhao, A. J. Gormley, R. Chapman, C. Boyer, *Macromolecules* **2019**, 52, 3.
- [21] C. Guerrero-Sanchez, J. Zhang, J. Vitz, U. S. Schubert, in *Encyclopedia of Polymer Science and Technology*, Wiley, Hoboken, NJ, USA **2019**, pp. 1–21.
- [22] R. C. Pullar, in *Information Science for Materials Discovery and Design* (Eds: T. Lookman, F. J. Alexander, K. Rajan), Springer, Cham, Switzerland **2016**, pp. 241–270.
- [23] T. A. Stegk, R. Janssen, G. A. Schneider, *J. Comb. Chem.* **2008**, 10, 274.
- [24] S. Chen, Y. Hou, H. Chen, X. Tang, S. Langner, N. Li, T. Stubhan, I. Levchuk, E. Gu, A. Osvet, C. J. Brabec, *Adv. Energy Mater.* **2018**, 8, 1701543.
- [25] B. R. Ortiz, J. M. Adamczyk, K. Gordiz, T. Braden, E. S. Toberer, *Mol. Syst. Des. Eng.* **2019**, 4, 407.
- [26] *Digital Transformation Initiative—Chemistry and Advanced Materials Industry*, World Economic Forum, Cologne/Geneva **2017**.
- [27] M. N. O. Sadiku, S. M. Musa, O. M. Musa, *J. Res. Technol. Eng. Manage.* **2018**, 2, 21.
- [28] Chemspeed Technologies SMOLE, <https://www.chemspeed.com/smole/>, (accessed: July 2020).
- [29] C. A. Shukla, A. A. Kulkarni, *Beilstein J. Org. Chem.* **2017**, 13, 960.
- [30] M. Rubens, J. H. Vrijssen, J. Laun, T. Junkers, *Angew. Chem., Int. Ed.* **2019**, 58, 3183.
- [31] S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone, L. Cronin, *Science* **2019**, 363, eaav2211.
- [32] J. M. Granda, L. Donina, V. Dragone, D.-L. Long, L. Cronin, *Nature* **2018**, 559, 377.
- [33] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes, N. Rankin, B. Harris, R. S. Sprick, A. I. Cooper, *Nature* **2020**, 583, 237.
- [34] T. L. Burnett, P. J. Withers, *Nat. Mater.* **2019**, 18, 1041.
- [35] J. N. Kumar, Q. Li, Y. Jun, *MRS Commun.* **2019**, 9, 537.
- [36] F. G. Coury, P. Wilson, K. D. Clarke, M. J. Kaufman, A. J. Clarke, *Acta Mater.* **2019**, 167, 1.
- [37] F. Xu, B. Corbett, S. Bell, C. Zhang, M. Budi Hartono, Z. J. Farsangi, J. F. MacGregor, T. Hoare, *Biomacromolecules* **2020**, 21, 214.
- [38] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **2018**, 559, 547.
- [39] T. C. Majdanski, J. Vitz, A. Meier, M. Brunzel, S. Schubert, I. Nischang, U. S. Schubert, *Polymer* **2018**, 159, 86.
- [40] T. Schuett, J. Kimmig, S. Zechel, U. S. Schubert, *Polymers* **2020**, 12, 2095.
- [41] V. Sans, L. Porwol, V. Dragone, L. Cronin, *Chem. Sci.* **2015**, 6, 1258.
- [42] M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse, S. V. Kalinin, *ACS Nano* **2017**, 11, 12742.
- [43] B. Sanchez-Lengeling, A. Aspuru-Guzik, *Science* **2018**, 361, 360.
- [44] The Materials Genome Initiative, <https://www.mgi.gov/> (accessed: July 2020).
- [45] X.-G. Lu, *Sci. Bull.* **2015**, 60, 1966.
- [46] E. Altuntaş, U. S. Schubert, *Anal. Chim. Acta* **2014**, 808, 56.
- [47] U. Dirnagl, M. Lauritzen, *J. Cereb. Blood Flow Metab.* **2010**, 30, 1263.
- [48] D. Mehta, *Nature* **2019**, <https://doi.org/10.1038/d41586-019-02960-3>.
- [49] M. Borenstein, L. V. Hedges, J. P. T. Higgins, H. R. Rothstein, *Introduction to Meta-Analysis*, Vol. 1, Wiley, Chichester, UK **2009**.
- [50] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, et al., *Sci. Data* **2016**, 3, 160018.

- [51] R. M. Salek, P. Conesa, K. Cochrane, K. Haug, M. Williams, N. Kale, P. Moreno, K. V. Jayaseelan, J. R. Macias, V. c. Nainala, R. D. Hall, L. K. Reed, M. R. Viant, C. O'Donovan, C. Steinbeck, *GigaScience* **2017**, 6.
- [52] Materials Cloud, <https://www.materialscloud.org/> (accessed: July 2020).
- [53] Automatic-Flow for Materials Discovery, <http://aflowlib.org/> (accessed: July 2020).
- [54] O. J. Reichman, M. B. Jones, M. P. Schildhauer, *Science* **2011**, 331, 703.
- [55] V. Marx, *Nature* **2013**, 498, 255.
- [56] S. Herres-Pawlis, O. Koepfer, C. Steinbeck, *Angew. Chem., Int. Ed.* **2019**, 58, 10766.
- [57] L. Banko, A. Ludwig, *ACS Comb. Sci.* **2020**, 22, 401.
- [58] *Nat. Mater.* **2016**, 15, 365.
- [59] L. Ward, A. Agrawal, A. Choudhary, C. Wolverton, *NPJ Comput. Mater.* **2016**, 2, 16028.
- [60] S. Arni, H. Melanie T., K. Alec, L. Caleb, Y. Sunghyun, H. Feier, F. Zhenxing, W. Christopher E., C. Yamil J., C. Yongchul, S. Daniel, S. Cory, *ChemRxiv* **2019**, <https://doi.org/10.26434/chemrxiv.7854980.v2>.
- [61] T. Le, V. C. Epa, F. R. Burden, D. A. Winkler, *Chem. Rev.* **2012**, 112, 2889.
- [62] K. Tran, A. Palizhati, S. Back, Z. W. Ulissi, *J. Chem. Inf. Model.* **2018**, 58, 2392.
- [63] X. Liu, D. Shi, S. Zhou, H. Liu, H. Liu, X. Yao, *Expert Opin. Drug Discovery* **2018**, 13, 23.
- [64] L. J. Santodonato, P. K. Liaw, R. R. Unocic, H. Bei, J. R. Morris, *Nat. Commun.* **2018**, 9, 4520.
- [65] A. Tarzia, M. Takahashi, P. Falcaro, A. W. Thornton, C. J. Doonan, D. M. Huang, *ACS Appl. Mater. Interfaces* **2018**, 10, 40938.
- [66] A. Gooneie, S. Schuschnigg, C. Holzer, *Polymers* **2017**, 9, 16.
- [67] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, A. J. Norquist, *Nature* **2016**, 533, 73.
- [68] S. K. Kauwe, J. Graser, R. Murdock, T. D. Sparks, *Comput. Mater. Sci.* **2020**, 174, 109498.
- [69] M. Aykol, P. Herring, A. Anapolsky, *Nat. Rev. Mater.* **2020**, 5, 725.
- [70] C. Chen, Y. Zuo, W. Ye, X. Li, Z. Deng, S. P. Ong, *Adv. Energy Mater.* **2020**, 10, 1903242.
- [71] Y. He, E. D. Cubuk, M. D. Allendorf, E. J. Reed, *J. Phys. Chem. Lett.* **2018**, 9, 4562.
- [72] K. M. Jablonka, D. Ongari, S. M. Moosavi, B. Smit, *Chem. Rev.* **2020**, 120, 8066.
- [73] F. Ren, L. Ward, T. Williams, K. J. Laws, C. Wolverton, J. Hatrick-Simpers, A. Mehta, *Sci. Adv.* **2018**, 4, eaaq1566.
- [74] O. A. von Lilienfeld, *Angew. Chem., Int. Ed.* **2018**, 57, 4164.
- [75] G. X. Gu, C.-T. Chen, D. J. Richmond, M. J. Buehler, *Mater. Horiz.* **2018**, 5, 939.
- [76] C. M. Childs, N. R. Washburn, *MRS Commun.* **2019**, 9, 806.
- [77] S. Keating, N. Oxman, *Rob. Comput.-Integr. Manuf.* **2013**, 29, 439.
- [78] T. D. Ngo, A. Kashani, G. Imbalzano, K. T. Q. Nguyen, D. Hui, *Composites, Part. B* **2018**, 143, 172.
- [79] N. Kleger, M. Cihova, K. Masania, A. R. Studart, J. F. Löffler, *Adv. Mater.* **2019**, 31, 1903783.
- [80] A. Lee, A. R. Hudson, D. J. Shiwarski, J. W. Tashman, T. J. Hinton, S. Yerneni, J. M. Bliley, P. G. Campbell, A. W. Feinberg, *Science* **2019**, 365, 482.
- [81] B. Jackson, *GE Prepares Boeing 777X for Takeoff with 3D Printed Turbine Blades*, <https://3dprintingindustry.com/news/ge-prepares-boeing-777x-for-takeoff-with-3d-printed-fan-blades-146566/> (accessed: July 2020).
- [82] F. Momeni, X. Liu, J. Ni, *Mater. Des.* **2017**, 122, 42.
- [83] Q. Ge, A. H. Sakhaei, H. Lee, C. K. Dunn, N. X. Fang, M. L. Dunn, *Sci. Rep.* **2016**, 6, 31110.
- [84] D. G. Moore, L. Barbera, K. Masania, A. R. Studart, *Nat. Mater.* **2020**, 19, 212.
- [85] F. Kotz, P. Risch, D. Helmer, B. E. Rapp, *Adv. Mater.* **2019**, 31, 1805982.
- [86] S. Langner, F. Häse, J. D. Perea, T. Stubhan, J. Hauch, L. M. Roch, T. Heumueller, A. Aspuru-Guzik, C. J. Brabec, *Adv. Mater.* **2020**, 32, 1907801.



**Julian Kimmig** studied chemistry at the Friedrich Schiller University Jena. He joined the group of Prof. U. S. Schubert for his master's degree in 2015. Since 2016, he has been working as a Ph.D. student, investigating new high-performance materials. Starting in 2019, he began implementing state-of-the-art machine-learning techniques in his work, which have become the main focus since then.



**Stefan Zechel** was born in Blankenburg (Germany) in 1988. He studied chemistry in Jena and performed his Ph.D. under the supervision of Prof. U. S. Schubert. He obtained a stipend from the Verband der Chemischen Industrie (VCI). After completing his Ph.D., he started a post-doctoral research in the group of Prof. Schubert. From 2017 to 2019, he was funded by the Carl-Zeiss foundation.





**Ulrich S. Schubert** performed his Ph.D. studies at the Universities of Bayreuth and South Florida. After a postdoctoral training position at the University of Strasbourg, he moved to the TU Munich and obtained his Habilitation in 1999. In 1999–2000 he was professor at the University of Munich, and during 2000–2007 full-professor at the TU Eindhoven. Since 2007, he has been a full-professor at the Friedrich Schiller University Jena, Germany. He is the founding director of the Center for Energy and Environmental Chemistry Jena (CEEC Jena) and coordinator of the DFG priority program “Polymer-Based Batteries” (SPP 2248).