Innovate UK
KTN

Innovation 4.0 Playbook: Digitalised Research, Development and Innovation in the Chemical Sciences

April 2022

MADE SMARTER
INNOVATION

Innovate UK
KTN

We create diverse connections to drive positive change.

Innovate UK KTN exists to connect innovators with new partners and new opportunities beyond their existing thinking – accelerating ambitious ideas into real-world solutions. Innovate UK KTN is part of the Innovate UK Group – the UK's innovation agency.

MADE SMARTER
INNOVATION

Innovate UK KTN was awarded a 4.5 year contract in October 2020 by UKRI to deliver the networking element of the Manufacturing Made Smarter ISCF challenge.

A £147m investment from the UKRI Industrial Strategy Challenge Fund (ISCF) will support the transformation of UK manufacturing capabilities through the innovation of industrial digital technologies. Key digital technologies in this challenge include:

- Artificial intelligence, machine learning and data analytics
- Additive manufacturing
- Robotics and automation
- Virtual reality and augmented reality
- The Industrial Internet of Things (IIoT) and connectivity (5G, LPWAN)

Contents

1.0	**Introduction**	**04**
1.1	Strategic context	05
1.2	Industry 4.0	07
1.3	Innovation 4.0	08
1.4	Cutting through the hype	09
2.0	**Innovation 4.0 - The Lab of the Future**	**10**
2.1	Lab automation, machine learning and autonomous discovery	11
2.2	Internet of Things (IoT)	14
2.3	Cloud data management	15
2.4	AI and ML	16
2.5	Virtual experiments (computer simulation and modelling)	18
3.0	**Digital Maturity Framework**	**19**
3.1	Notes on the levels	22
3.2	Discussion	23
4.0	**The Opportunities and the Business Case(s) for Change**	**24**
4.1	The urgency of chemical science research	24
4.2	Resilience	25
4.3	Supply chains and material variability	25
4.4	Scale up	25
5.0	**The Challenges**	**26**
5.1	Human factors	26
5.2	Fragmentation	27
5.3	Knowing where to start	27
5.4	Knowing who to engage with	27
6.0	**What Next?**	28
7.0	**Case Study**	**29**

1. Introduction

Digital technologies are now commonplace across many commercial, industrial and civic environments and we rely heavily on them to perform many day-to-day tasks. The advantages of adopting digital technologies for some tasks and processes are easy to recognise, but for others it is not so straightforward.

We have written this Playbook to provide guidance to innovators in companies deploying and creating advanced materials, chemistry and formulations, on how best to understand and describe the use of emerging digital techniques to transform their research, product development, and innovation activities. The Playbook will also be of interest to academic groups looking to increase the productivity of lab-based research in their respective universities.

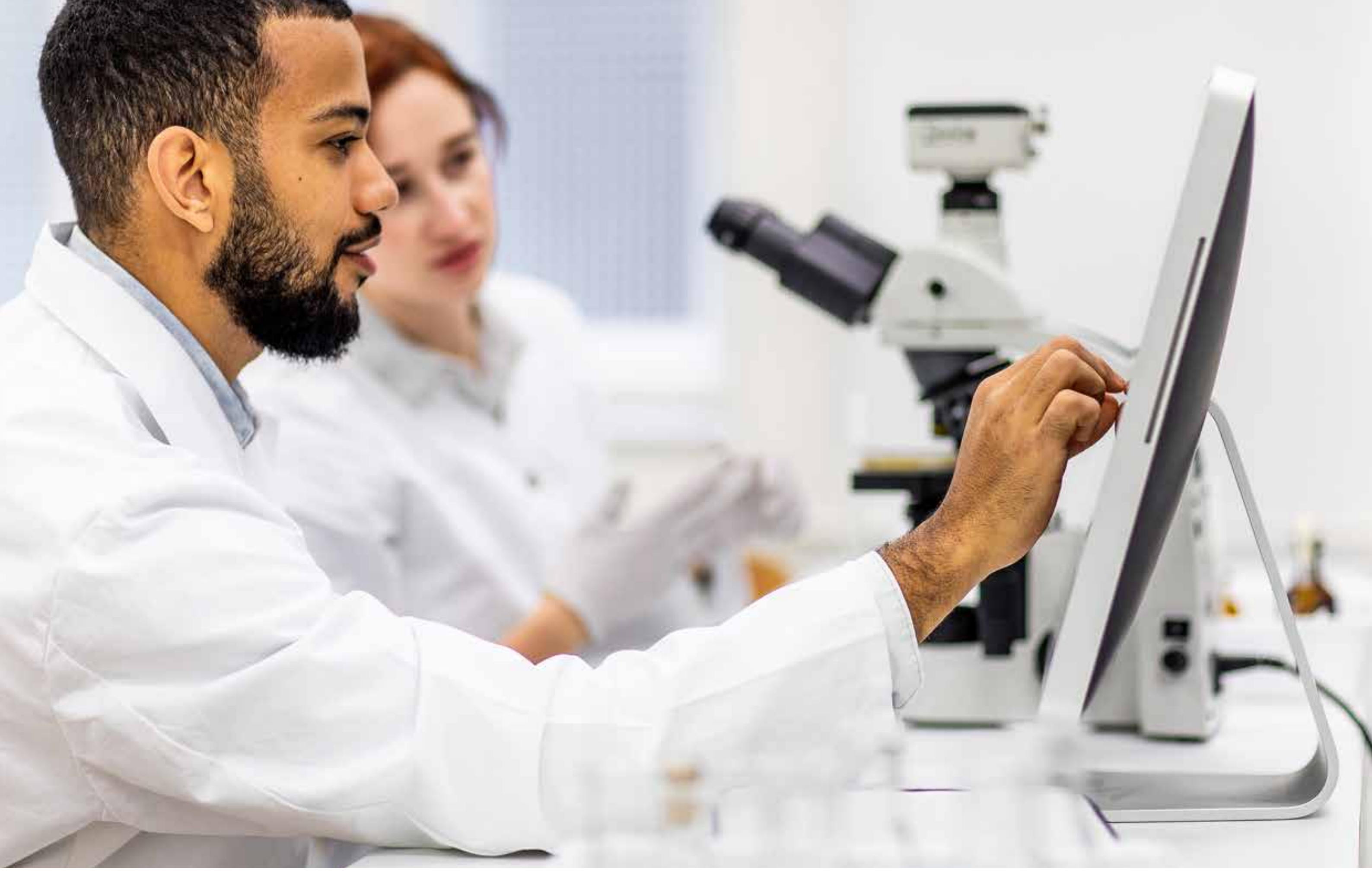
The Playbook has been written for practitioners, and reflects the experience of commercial innovators who have successfully applied digital techniques in their research and development (R&D) processes. It is grounded in an understanding of the strategic imperative of increasing the end-to-end productivity of commercial innovation teams, and also addresses the complex stakeholder management tasks which are needed for a successful implementation of new processes.

1.1 Strategic context

Many macroeconomic trends and societal issues are driving the 'fourth industrial revolution', not least the acute problems caused by the COVID-19 pandemic, which saw shared working spaces including laboratories either closed entirely or dramatically reconfigured to implement social distancing and hygiene guidance.

This had a clear negative impact on the output of research laboratories as 'in the lab' staff numbers were cut. The pandemic hit as the UK (and other developed nations) struggled with sluggish productivity growth and the need to create more highly-skilled, well-paid employment across the country (the 'levelling-up' agenda). However, as a result of a drastic and sudden change of working practices, companies and organisations have had to quickly innovate and adopt new technologies, as the CEO of Microsoft, Satya Nadella, observed: "We saw two years of digital transformation in two months"[1].

This digital transformation has already changed how many office tasks are carried out, using video conferencing, document sharing tools, and email. However, this is not true for lab-based R&D - for the vast majority of R&D companies (and university research departments), laboratory experiments cannot be run 'from home'. Furthermore, due to reduced numbers of researchers permitted into labs under physical distancing guidelines, many labs will be working at less than 50% capacity for the foreseeable future. For organisations that rely on lab (and related) activity to create commercial value, the impact of long-term restrictions on lab use will destroy substantial value. We can estimate this lost value - the combined turnover of the chemical and pharmaceutical sectors in the UK is about £49 billion, of which £18 billion is value added[2]. If we assume that 20% of this value-added comes from R&D, and 60% of R&D value comes from lab work, then a 75% reduction of lab capacity would lead to the destruction of more than £1.5 billion of value for each year of physical distancing. Moreover, there is an opportunity cost, as companies dedicate this ~25% lab capacity to maintaining current manufacturing output and business continuity whilst postponing or cancelling R&D programmes. This 'innovation gap' is not immediately visible, but in the medium and long term becomes a significant issue for companies and society more generally as the ability of our chemistry-related industries to innovate is curtailed.



We should also highlight here the critical importance and relevance of the chemistry, formulation and materials industries to the UK, and the need to focus on the digitalisation opportunity. The chemicals and pharmaceuticals sectors, for perspective, are approximately the same size as aerospace and automotive, and enable many downstream technologies and innovations. Indeed, the sectors are highly innovative, outperforming the financial services sector in total factor productivity over the last 20 years[3]. They support the employment of 135,000 directly and approximately 300,000 indirectly. However, there is a significant degree of fragmentation, with approximately 80% of the 3,700 registered businesses having a headcount of less than 100. This fragmentation increases the challenge of digital adoption – there are hundreds of laboratories and research teams in the UK with a myriad of experimental processes, data collection and analysis methods, which are unlikely to lend themselves completely to a one-size-fits-all approach.

1.2 Industry 4.0

'Industry 4.0' has become a ubiquitous phrase amongst industry professionals but can mean different things to different people. Nominally, Industry 4.0 is the answer to the manufacturing productivity puzzle.

It also has applicability in an R&D context. A report released in 2017 - The Made Smarter Review[4] - identified the following Industrial Digital Technologies (IDTs) as drivers of Industry 4.0:

- Additive manufacturing.
- Artificial intelligence/machine learning and data analytics*.
- Robotics and automation*.
- Industrial Internet of Things and connectivity (5G)*.
- Virtual reality and augmented reality.

The application of these technologies is beginning to deliver a wholesale transformation in the way that products are manufactured (fuelled in part by the Made Smarter programme funded by Innovate UK KTN). Their adoption is also leading to profound changes in work practices and productivity. However, it is unlikely that the full potential of Industry 4.0 within the chemistry-related sectors in the UK will be achieved without a parallel investment in a range of new innovation platforms and approaches in R&D.

All of the IDTs listed above can be applied in R&D, and in particular, those marked with an asterisk have direct applicability to lab-based R&D activities. Other relevant technologies, deployed to various extents across the industry, include advanced simulation, modelling and computational chemistry. Digital technologies will be used to develop ground-breaking approaches for both academic science and commercial R&D and innovation into the future. We call this Innovation 4.0.

1.3 Innovation 4.0

Many of the key platforms required for Innovation 4.0 exist in prototype form across the UK and Europe in university labs, industrial R&D facilities, within the Catapult centres (e.g. Centre for Process Innovation), and dedicated research centres (e.g. the Materials Innovation Factory, iDMT).

This Playbook describes how companies who have yet to invest (or wish to invest more to continue their journey) can benefit by digitising their R&D work. But 'Going Digital' is not easy. Many companies are looking with interest at how to transform their research, innovation, and product development by implementing digital techniques, but[5]:

- they don't know where to start;
- they find it hard to make the business and investment case;
- there is no 'one-stop shop' for advice and guidance on adopting digital technologies in chemical lab R&D, and they want to partner with the best leaders and organisations in the field to drive the change;
- they want to implement the changes that will have the biggest impact and lowest risk. Evaluating the impact and risk proves to be a difficult task;
- they have found that many large IT vendors have little real insight into the world of chemical R&D;
- they need to equip their staff with the right digital skills.

These are some of the issues that we address in this Playbook. The concepts and insights we present have been developed in a number of large R&D organisations and leading university labs over the past 15 years. They are presented here to help other organisations, big and small, to understand and describe how digitalisation can transform their innovation practices. This common language will allow us to talk about digitalisation in an R&D context, distinct from the manufacturing context.

1.4 Cutting through the hype

For many innovators the key challenge in adopting digitised solutions in R&D processes is to translate the buzzwords, corporate slogans and jargon into tangible, costed and implementable solutions that make sense for their own business.

To this end, we encourage the reader to pay little attention to the branding of the technologies, and more to their function. Furthermore, the implementation of digitised solutions will, in the majority of cases, require the business to analyse their current R&D processes to be able to make decisions about where changes should be made.

Be assured that the irony of requesting the reader to divert attention from buzzwords, whilst introducing a new one (Innovation 4.0), is not lost on us - we do this only to differentiate the part of the business we are referring to (R&D, as opposed to manufacturing)!

2. Innovation 4.0 – The Lab of the Future

The promise of Innovation 4.0 is the wholesale transformation of industrial research, development and innovation in the chemistry sector, achieved by the widespread adoption of new digital R&D technologies and correspondingly new approaches to innovation work practices.

The outcome of Innovation 4.0 is a portfolio of new digital assets, which can be used to re-engineer the usual innovation process. The benefits will be any or all of the following:

- Speed: Launch bigger innovations faster.
- Innovation: Develop superior, but hard to make, products.
- Scale-Up: Faster and more robust scale-up - no nasty surprises!
- Roll-Out: Consistent delivery and traceability.

2.1 Lab automation, machine learning and autonomous discovery

The way we perform chemistry, materials and formulation research has changed remarkably over the past half-century thanks to the increased range of scientific instrumentation and equipment available to the modern laboratory.

However, one thing that has remained fairly constant is the role of the researcher, in the way that they plan, analyse and conduct experiments to work towards a research objective. Often a large chunk of a researcher's time is spent performing repetitive tasks, from manual execution of laboratory experiments to data analysis. These tasks lend themselves to automation, both in terms of robotics, and AI techniques for decisions making[6]. There is an emerging vision in the chemical sciences that sees a future in fully autonomous labs[7].

The advantages of autonomous labs are numerous. Researchers are liberated from tedious lab work and freed up to apply themselves to the higher-level task of scoping, framing and deciding on good research questions. Autonomous systems can develop new chemical intuition and explore chemical/formulation parameter space in more productive ways. The results from each and every experiment contribute to a database from which new insight can be gained. The safety of lab workers is enhanced as they spend more time in the office and less time in the lab (papercuts notwithstanding).

Fully autonomous labs would be a step up from high-throughput experimentation (HTE). HTE has been in use for at least 30 years, and has significantly improved the process of discovering new functional materials and drugs[8,9]. In HTE, large amounts of experiments are conducted by robots (and in hybrid approaches, simulations - or 'virtual experiments' - are also run). However, traditional HTE approaches are 'exhaustive' or 'brute-force', and their practical performance plateaus due to the combinatorial explosion of design space as the systems of interest become more complex. This is where advances in artificial intelligence (AI) can play a role. By treating the parameter search as an optimisation problem, AI can be used to narrow the search space through optimisation algorithms[7]. Although this approach represents a step-change in autonomous discovery, it is still limited by the application-specific nature of the hardware and software used and is still only practical with about two to five experimental variables[7]. To take the final step towards self-driving labs that can operate independently from humans for long periods of time towards a specified research objective, the 'loop' must be closed.

Realising these closed-loop systems requires the use of machine learning (ML)[10–12], a sub-branch of AI (ML can handle much more complexity than the techniques used previously), along with robust control software for scheduling and running automated experiments.

In this 'closed-loop' system, experiments are automatically executed, analysed, and decisions made about which experiments to do next. A recent project at the Materials Innovation Factory has shown the promise of this approach in the area of catalyst design[13].

Where are the productivity gains in autonomous labs?

A direct R&D productivity improvement is achieved when a wide range of workflows are implemented with lab robotics in chemistry lab work. These systems do not replace human knowledge workers, but instead profoundly augment the human capabilities of an R&D team.

Let's start with the robotic lab systems themselves. Robotic lab systems deliver a simultaneous step-change in three independent aspects of experimental science:

(a) Reproducibility

Robots can, almost by definition, repeat a physical process to a high level of precision. In the case of a chemistry lab robot, this precision applies to all aspects of an experimental protocol: weighing materials, adding liquids, mixing reagents, controlling reaction conditions, measuring reaction conditions, measuring end-points etc. Practical experience has shown that for many manual assays, the adoption of a robotic protocol can deliver a three-fold to five-fold improvement in reproducibility, i.e. experiments run by humans in triplicate can be replaced by a single robotic evaluation.

(b) Traceability

Fully digitally-controlled experiments allow the trivial logging of process conditions throughout the protocol. Instead of relying on human paper-based records of what happened, robotic experiments automatically log (with high fidelity) exactly what happened, and when. Although this data is rarely the desired end-point measurement, the ability to record and store in a retrievable manner the complete 'history' leading up to a measurement is invaluable for troubleshooting and for experiments that seek to understand order-of-addition effects (a particularly important aspect in formulation).

(c) Throughput

The capacity of a lab robot is unaffected by its mood, the psychological impact of many external events, and fatigue. Even when a robot appears to be moving sedately, it can continue to operate in a highly reproducible manner, often for 24 hours per day, five or six days per week (with a day dedicated for maintenance work). Overall, this can deliver a markedly increased throughput in terms of the number of experiments per unit of time. In addition, the number of experiments required to answer a research question will be minimised by the analytics performed on the high-fidelity datasets produced by the robotic platform.

All three of these factors are enhanced simultaneously and independently by the adoption of lab robotics. Loosely, we can argue that the total productivity increase would then be a multiplication of the increase in productivity in each factor, i.e. taking a factor of two for each would lead to an overall eight-fold (i.e. 2 X 2 X 2 = 8) improvement in productivity of the experimental protocol.

Now, if we include the 'autonomous' aspect, i.e. the AI/ML system that generates decision-making capability and control of the robotics, there is an additional improvement of productivity. One of the unexpected upsides of implementing automation in lab activities is the fact that a high-quality experimental data creation process, with well-controlled set-up and high-quality data capture, removes a well-known bottleneck in the application of AI – namely the requirement in almost all real ML and AI applications to manage 'dirty-data'[14].

In many practical R&D cases in which a predictive model is the target, it may well be more productive to create a new robotic + AI workflow in parallel to existing manual activity rather than try and collate, curate and cleanse existing legacy data of questionable provenance and quality.

Robotic experimentation also forces R&D staff to address and record key aspects of the data creation process: how the data is gathered, its provenance, what the different parameters mean, etc.

Robotic paradigms

There are two ways to implement robotic experimentation. The first is to put together a fixed or semi-fixed (modular) array of robotic equipment that will execute a sequence of experiments/characterisations (traditional high-throughput experimentation approach). Usually the materials and samples will be loaded in at one end and processed by the array. This approach is more relevant if the R&D processes involve a set of standard tests that need to be executed often, and do not vary too much between sample/product types. A degree of flexibility can be retained if needed, by using interchangeable units, however, if the testing requirements vary too much across a period of time, this paradigm will become unfeasible.

The alternative paradigm is to use a mobile robot[13], which mimics a human researcher in that it will move around the physical lab space, prepare samples and run experiments. This approach requires some modification of the lab space but still allows parallel human participation. The main difference, of course, is that the mobile lab robot can work constantly for days at a time, when coupled with an AI system for decision-making. This paradigm makes less sense if the experiments or tests required do not vary or are unlikely to vary over time.

2.2 Internet of Things (IoT)

The widespread deployment of high computing power, low energy usage, digital sensor and communication devices has become known as the Internet of Things (IoT)[15].

These technologies have recently begun to appear in consumer devices – such as the smart-plugs and voice command devices available from Google, Apple and Amazon.

The same approach can also be deployed throughout an R&D environment[16]. In a lab, small electronic laboratory devices, such as balances, stirrers, thermometers, pH meters, barcode readers etc, are now able to be digitally connected and integrated into Wi-Fi enabled IT systems. When deployed in a coherent way, these devices allow a complete digital lab bench to be built. An R&D scientist can work at a digitised version of their usual lab bench to create digital data at source. This is data which has not been manually inputted, e.g. typed into a spreadsheet or written in a notebook, but captured from the devices they are working with directly to a time-stamped data file – all without having to change their individual work-practices.

In addition, equipping lab staff with mobile technologies, communicating with the equipment in the lab, can alert them to changes in conditions in lab experiments etc, which can help the scientist manage their own productivity and time.

It should be noted that there are a number of outstanding challenges involved with this approach – the digitised chemistry lab bench – which include the lack of widely adopted interoperability standards for lab equipment, poorly understood user experience requirements for lab chemists, data privacy constraints and GDPR implications which are implicit in the creation of personally identifying data by lab infrastructure.

2.3 Cloud data management

Cloud computing is a solution for companies who do not have sufficient in-house IT expertise to host their own datasets and manage their computing infrastructure[17].

Third parties can be contracted to provide these services, and in many cases provide some analytical tools. However, difficulty in accessing high-quality R&D data is one of the core issues faced by data scientists. This needs explicit attention, and is not best served by re-badging enterprise software solutions. Innovation 4.0 needs data management platforms which are designed from the ground up to record domain relevant meta-data as well as process and end-point data.

2.4 Artificial Intelligence and Machine Learning

The use of advanced data science techniques, such as artificial intelligence (AI) and machine learning (ML), which are augmented by the incorporation of domain-specific scientific insights to constrain search strategies and exploration algorithms.

These hybrid science-AI approaches are not off the shelf, nor are they trivial applications of off the shelf AI-ML techniques. They do not rely on so-called Big Data, but rather they rely on very good quality small data and the encoding of pertinent physical insights and laws.

Technological applications of AI require explicit consideration of 'explainability' and trust. Here the provenance provided by world-class academic science becomes very important.

The expression artificial intelligence, or AI, is a widely used and general term that refers to hardware or software platforms that are able to undertake complex tasks which usually require human intelligence or decision-making skills. The field has been researched by computer scientists, neuroscientists and statisticians since the 1950s. Early applications tended to focus on rules-based programmes that could implement some rudimentary decision-making processes in highly constrained contexts. These earliest forms of AI included 'expert systems', which were an attempt to capture what human specialists knew in order to make expert decisions.

Complex creative tasks, such as those involved in scientific discovery and product innovation, are ill-suited to a rules-based AI approach. Real-world scientific and technical challenges are often too complex to be solved by computer programmes that follow sets of rules written by experts. What experts tend to do is ask a series of complex nested questions, often which loop back to the start of the problem during the unfolding process of discovery or innovation.

Machine learning (ML) is a subset of AI which is not rules-based a priori. As its name suggests, machine learning seeks to create a computer programme which has learnt something about the structure of the problem through training. Typically, this approach requires a significant set of training data, on which the ML algorithm is built. The reason that ML is attractive is that it uses a computer to create the 'rule set' rather than a team of human experts. This means that pragmatically the cost of starting a project can be lower, and more complex and subtle problems can be tackled (and solved). Machine learning was defined by the American computer scientist Arthur Samuel (1901 – 1990) as the: 'field of study that gives computers the ability to learn without being explicitly programmed'.

Technological applications of AI require explicit consideration of 'explainability' and trust. It is here where the provenance provided by world-class academic science becomes very important. The reader is directed to the references for in-depth reviews and examples of machine learning in chemistry and formulation[7,10,12,18].

2.5 Virtual experiments (computer simulation and modelling)

In addition to physical experimentation, the use of virtual experiments to gain insight into a chemical system of interest are increasingly used as part of the innovation process[19,20].

One of the main advantages of simulation is the ability to probe length and timescales inaccessible to even the most cutting-edge experimental equipment. Thanks to recent advances in computing architectures (e.g. HPC and quantum computing) simulations can be run for longer and/or with more moving parts. New methods of structuring simulations allows for the extrapolation of results to higher and lower length and timescale, boosting their explanatory power. New 'apps' and front-ends are under development to make it easier for more and more researchers to deploy simulation techniques as part of their day-to-day research, without having to become experts in coding or cloud computing. Virtual experiments can either be run on an ad-hoc basis to aid decision making, or as part of the 'closed-loop' autonomous discovery system outlined in section 2.1.

3.0 Digital Maturity Framework

Now we consider the implementation of these technologies and approaches. There are numerous tools and frameworks available under the banner of Industry 4.0 (e.g. 4M) to help companies understand the opportunity and identify ways to implement it.

However, the language and relevance of these are heavily focused on (discrete) manufacturing operations. Innovate UK KTN, working with partners, identified the need to develop a framework for understanding the implementation of Industry 4.0 in an R&D environment i.e. Innovation 4.0. It is important to note that the major changes in both technological and managerial approaches that are required to fully implement Innovation 4.0 do not need to be applied in one go. Innovation 4.0 will not arrive with a Big Bang! Moreover, attaining the highest level of digital 'maturity' need not be the goal.

Here we outline a framework for understanding the stages of complexity involved in adopting Innovation 4.0. The levels build from easy-to-implement digital lab tools (Level 1), through a range of lab automation approaches, specialised machine learning (ML), artificial intelligence (AI) and simulation platforms (Levels 2-3), to fully integrated, digitised, and autonomous R&D (Level 4).

Table 1: High-level descriptions and comments of the different levels of 'digital maturity'

LEVEL	Physical experiments	Virtual experiments	Analytics
Level 0	Experiments are run by hand, with conditions and results recorded by hand or written to local data files which remain distributed on individual machines or notebooks. Only simple office software file storage methods are used.	Minimal, if any, use of simulation and modelling techniques for generating an understanding of the material/chemical/formulation properties.	No deployment of AI or ML methods on data. Only simple offline analysis tools are used. Design of Experiments (DoE) methods underutilised.
Level 1	Direct digital capture of information from manual workflows to a secure repository. Experiments are run by hand, but conditions and results are automatically recorded and stored in a central, secure repository, allowing enhanced access. (IoT, data loggers, sensors, integrated lab notebook systems).	Some use of simulation and modelling techniques for understanding the material/chemical/formulation properties, but manually executed and/or conducted on a project-by-project basis. No direct integration with lab data.	DoE methods used to inform experiment planning.
Level 2	Automated lab robotics with experimental data management. Execution of high frequency, time-consuming tasks with high reproducibility.	Simulation and modelling techniques integrated with high fidelity, robotically generated lab data.	AI and ML methods deployed on the high quality robotically generated datasets for enhanced insight and experiment planning.

LEVEL	Physical experiments	Virtual experiments	Analytics
Level 3	Highly integrated workflow management. A seamless blend of digital data capture, lab robotics, simulations and flexibly-deployed significant computational resources. Physical and virtual experiments are conducted in a single workflow.		AI and ML methods are informing the experimental direction, but the human researcher is still in ultimate control of the workflow, defines the end-point and provides validation input.
Level 4	Algorithmic control of closed-loop workflows. Autonomous control for discovery, optimisation, and routine testing. Application of ‘reasoning AI’ methods in formulating hypotheses and designing, and executing, physical and/or virtual experiments. Research questions are posed by humans, but the experimental workflow is created and executed entirely by AI and lab robotics.		

3.1 Notes on the levels

Level 0

We suspect that Level 0 will be familiar to quite a number of readers, either because it describes their current research practices or because it describes their practices from the not so distant past. The main productivity loss at this level is the lack of centralised and structured data storage, which prevents the implementation of analytics and increases the need for experiments to be repeated in the event that the original experimenter leaves the organisation. Moreover, because the experimental account is significantly subjective, the traceability is sub-optimal.

Level 1

To graduate to Level 1, experimental data generated by manual experimentation should be captured and sent to a centralised and structured repository. This can be done using an electronic lab notebook as long as minimal amount of information is input manually by the user – the important thing here is that the lab instruments produce the data and metadata automatically and it is stored immediately. Simulations and modelling should be a regular part of project workflows (if applicable). For experiment planning, DoE methods should be used.

Level 2

At Level 2, most if not all of the physical experimentation burden should fall on robotics – either in the modular array paradigm or the mobile robot paradigm. This produces higher-quality datasets that allows the productive deployment of AI/ML methods and the subsequent integration with simulation and modelling methods. The role of the research staff becomes more about research direction and workflow planning, and less about actually performing the experiments themselves.

Level 3

Level 3 requires an integration between the robotic experimentation platform (either mobile or static) and simulation and modelling methods, as well as a more prominent role for AI/ML in defining the experimental direction. Research staff retain ultimate control at this level; however, the workflows are largely defined by AI.

Level 4

Level 4 is the ‘closed-loop’ autonomous system that, in real-time, conducts and plans experiments in a continuous loop until an acceptable answer to the research question is reached. This represents the pinnacle in computer and robot-aided discovery and product development for chemistry, materials and formulation R&D.

3.2 Discussion

In general, the Digital Maturity Framework aims to be an antidote to the manufacturing-focused reference material already published, and to help companies accelerate their digitalisation journeys – by providing a common language for defining their starting point, and their ambition.

It should be viewed as illustrative, and not prescriptive. Level 4 maturity does not have to be the ambition for every organisation, and the progress through the maturity levels can be more flexible than the rubric implies.

The specific instrumentation, robotic equipment and software required for any given lab will vary from organisation to organisation, but the high-level concepts described in this table should apply to all. Companies should be able to look at these descriptions and begin to figure out what level of digital maturity they are at - and what the next level might look like.

It is important to note that innovators should not think that they need to digitally mature every single one of their R&D workflows simultaneously - it is possible to choose the workflows that most readily lend themselves to automation and start from there. For example, if one workflow involves experimentation on only one or two pieces of equipment - e.g. a liquid and powder dispenser, a sonicator (for dispersing powders in a liquid), and a rheometer (for measuring flow properties) - this may be digitally matured in isolation from other workflows, perhaps as a test case to prove the concept of Innovation 4.0 in the company or organisation. Other workflows that involve sample transport between rooms, e.g. for microscopy or electrochemical testing, or manipulation that requires a high degree of dexterity (not provided by the current state-of-the-art in robotics), can be matured later. Hence, digital maturity can be applied either to individual workflows or the organisation as a whole, taking into account all R&D activities.

It is also possible that maturity is enhanced along one of the domains, e.g. physical experimentation, whilst the others are left alone. However, at some point, the maturity in the other domains may need to catch up. For example, suppose you have a robotic experimentation platform producing mountains of high-quality data. In that case, it does not make much sense if these datasets are not being analysed by state-of-the-art AI/ML methods and/or being used to inform relevant simulation and modelling activity. Indeed, to progress to Level 3, the physical and virtual domains should be integrated, and to progress to Level 4, further integration with AI/ML methods for decision making is required.

4.0 The Opportunities and the Business Case(s) for Change

4.1 The urgency of chemical science research

The chemical sciences, of course, have a great deal to contribute towards solutions for many outstanding societal challenges (climate change, living standards, health outcomes, food supply etc). A chemical science company's capacity to contribute is dependent on many factors, including employee availability, skills, access to finance, access to facilities and manufacturing technologies. When it comes to a typical product launch, the bottleneck is likely to be in the initial discovery/invention/design and scale-up phase, which requires weeks and months of experimentation, computer modelling, testing, which often proceeds in a circular manner until the product is ready for manufacture. The need for new materials, pharmaceuticals, consumer products could be met sooner if this process was streamlined, automated, and scaled.

4.2 Resilience

The COVID-19 pandemic of 2020 significantly disrupted the performance of chemical science research in the UK, through a combination of enforced physical distancing, staff absenteeism (self-isolation) and the same disruption to supply chains. Industries less affected by the pandemic have been those who rely much less heavily on human presenteeism at specific locations, e.g. labs and offices, allowing work to continue from home. Working from home is, of course, not feasible for researchers involved in lab-based work – however, under Innovation 4.0, laboratories would contain automated experimental platforms requiring little human presence.

Beyond the pandemic, resilience will remain an important issue in the face of regulatory changes, reformulation requirements, the growth of personalised medicines and other consumer products as well as the impact of overseas competition. The UK has the right set of expertise that can make the UK the go-to place for digitalised innovation.

4.3 Supply chains and material variability

Products that contain a large number of components or ingredients which are processed in often convoluted ways are particularly susceptible to slight changes in the properties of raw materials. These can often lead to quality control issues. A digitised laboratory will be better equipped to document, predict and troubleshoot these issues, especially if properly integrated with a digitalised supply chain.

4.4 Scale up

Scale-up exercises tend to be the most troublesome aspect of product development in chemistry and formulation. A digitally mature R&D function will be better equipped to determine the scale-up rules for new products and optimise the transfer of knowledge from lab-based R&D to full-scale manufacturing, and encode these rules in a way that allows AI methods to derive useful insights.

5.0 The Challenges

5.1 Human factors

The introduction of Innovation 4.0 approaches into an existing R&D workforce is non-trivial.

Some will view the encroachment of robotics and AI into the realm of R&D as a threat to their employment or career prospects; however, it is the prevailing view that these technologies will augment rather than replace the work of scientists, as human intervention will be required to provide validation of experimental results, research direction and knowledge generation. Scientists will be able to dedicate more energy to the intellectual exercises involved in knowledge generation due to AI-driven robotic experimentation platforms taking over much of the physical tasks.

There are a number of critical micro-level and macro-level changes which need to be made for successful implementation of Innovation 4.0. For researchers used to driving their programmes 'one experiment at a time' or with 'paper lab-books', the impact can be marked. Robotic platforms, DoE, virtual experimentation and AI/ML imply a major disruption to the working style of many scientists. Digital skills will become much more important in undergraduate and postgraduate education, and continual professional development will be required to upskill the workforce.

Another interesting pitfall to be aware of is that as R&D becomes more automated and streamlined, the research culture may prioritise simple, automatable workflows over more complex, manual ones, which may have the net effect of discouraging innovation rather than promoting it[21]. It may turn out to be a naive assumption that because machines will be doing all of the simpler, repeatable drudge work, scientists will spend more time doing the complicated stuff - perhaps not much of the complicated stuff will get done at all. To counter this, it will be important that digitalised solutions provide for new modes of working, rather than exclusively speeding up older ones.

Other emergent issues are the lack of user-friendly interfaces for lab work and data analysis (AR/VR has a role to play here), the lack of interoperability standards which allow machines (often of differing manufacturer) to communicate with one another. The impact of systems that have some degree of agency or rely on autonomous algorithms for part of the decision making process. Larger scale organisation of work will also be impacted.

5.2 Fragmentation

The chemicals sector is significantly fragmented, with approximately 80% of the 3,700 registered businesses having a headcount of less than 100. This, coupled with the variation in experimental methods, data collection and analysis requirements, creates a great challenge for digitalisation. Platform technologies will overcome some of this challenge. Still, at some stage of the sector's digital evolution, a certain degree of democratisation will be required to allow companies to modulate, adapt and tailor the digital technologies for their individual needs. This will be greatly assisted by the development of, e.g. data/communication standards and the upskilling of chemists and formulators in digital techniques.

5.3 Knowing where to start

Companies invariably take their own paths when it comes to investing in infrastructure, skills and workflows, and will be at different stages of a digitalisation 'journey'. We have written this Playbook to set out a framework for understanding what digitalisation can do for a company's R&D and how mature a particular company might be at the time. Hopefully this will serve as a way to identify what a company might do to take another step on that journey and become more mature. Innovate UK KTN can help with this by identifying and introducing you to some key stakeholders and/or solution providers relevant to the elements of the R&D you want to digitalise, whether that's increasing your use of virtual experimentation, the adoption of a robotic experimentation platform, or simply connecting your lab devices to automatically record and store experimental information. Generally there isn't a one-size-fits-all solution to a company's digital aspiration, but by working with the right people tailored solutions can be developed.

5.4 Knowing who to engage with

The starting point for many companies will be to engage their IT teams in the implementation of digital technologies. This might lead to engagement with the big IT vendors who have little understanding of the nature of chemical science innovation. A better path would be to engage directly with the solutions providers that do have a working knowledge of chemical science innovation where digital technologies can fit in. This will require a large scale socialisation and democratisation exercise whereby companies and solution providers can work together to co-create the technology 'products' or 'assets' that can be widely adopted at an affordable price, especially in the area of lab robotics[22]. This may require further investment in testbeds and open access facilities.

6.0 What Next?

We have written this Playbook to be the beginning or part of a conversation about what R&D digitalisation means for your business or organisation.

In this regard, Innovate UK KTN is here to continue the conversation, connect you with a community of innovators actively putting Innovation 4.0 into practice and provide introductions to digital technology providers and the wider innovation ecosystem in the UK. If you would like to have this conversation, please reach out to us – we're here to help!

7.0 Case Study

Domino Printing Sciences has kindly provided the following case study, to highlight the opportunity, benefits and challenges of digitalising their R&D activities.

Company background

Founded in 1978, Domino Printing Sciences is a world-class provider of coding, marking, and digital printing technologies. With more than 40 years' experience working at the heart of the industrial and commercial printing markets, Domino has established a global reputation for excellence and innovation within its core technologies, worldwide aftermarket products, and best-in-class customer service. Domino's continued growth is underpinned by an unrivalled commitment to product development.

The digitalisation opportunity

We are seeing an increasing demand for new ink products caused by the release of new printer platforms, new customer applications being required and an ever-changing regulatory and supply chain landscape. Reducing the time to market for new inks, with no compromise on quality, is therefore critical to ensure continued business success. The digitisation of our ink development process has been identified as a way of innovating the way we develop our products with successful implementation allowing us to increase productivity within R&D. To achieve this, we have set ourselves a goal of introducing and implementing automated equipment and formulation workflows to our ink design process.

Adopting digital technologies

As a first step, we identified the key properties which are routinely measured when designs are tested and sought to introduce commercially available equipment which had autosampling capabilities. As an example, for every formulation that is produced a Chemist will measure viscosity and density. To save time and effort, we introduced a viscometer which has the capability to run multiple samples via an autosampler. We have estimated that we have saved five minutes operator time per sample and, as it can be run unattended overnight, we can run many more samples than would have been possible using the conventional equipment we had. This has led to equipment payback within six months of introduction.

The next step has been the introduction of a robotic formulation platform to allow us to perform high throughput workflows. We installed the equipment into our laboratories in Spring 2020 and have since trained a group of super-users how to use the equipment and design workflows. We took this approach to ensure that, as well as effectively using the robotic platform, we were diffusing the knowledge and skills across the wider team. Since its introduction we have produced approximately 1,750 samples, including formulations for all our ink technologies. The equipment is now being used to augment the work performed by our Development Chemists in their new product development activities. We are targeting an efficiency saving of 10% within our product development cycle which, for a typical development project, will result in us getting a product to market around two months faster. We are also expecting a decrease in the experimental 'touch time' for our Chemists and are aiming to reduce this by a minimum of 20%. This should free around 1,500 work hours across the team over a year which can then be used for other project activities. Along with these tangible benefits, we are hoping for an increase in the amount of innovation capacity within the team as our Chemists find more time and space to explore novel chemistries and ways of working.

We also had to consider how to best store and use our data. After searching for commercially available electronic lab notebooks, we decided to create and build our own databases using our internal software development capability. We are currently in the process of creating different applications to enable users to view the data that we collect, along with standard reporting and modelling tools to help us get the most from the information. The decision to do this internally was driven by the fact that none of the available tools exactly suited our requirements. Although this has been a large piece of work, we have noticed a rapid uptake of the software in the department as our Chemists have been heavily involved in the design and creation process.

Benefits

Utilising automated pieces of equipment to produce and test trial formulations has been incredibly useful to maintain experimental throughput during the COVID-19 pandemic where we have had to significantly limit occupancy numbers in our laboratories.

Our aim is for automated workflows to become a routine part of our product development process during 2021. We will be measuring the number of samples produced each month and overlaying this with effort saved and knowledge gained. We are also exploring how to assess the cultural change that we expect this equipment to bring and will be looking to measure staff motivation, sharing of best practices and the increased innovation capacity and creativity within the team.

What next?

Uptake of equipment has been really positive to date, and we are confident that this will transform the way we perform our ink development activities. We will be using the equipment as part of our upcoming research and development activities in new ink design projects as well as some reformulation activity. We will also use the equipment to help us explore the design space of our products and evaluate novel materials. This will enable us to extend and expand our formulation capabilities and deliver innovative products for our customers.

We will continue to explore digitisation opportunities over the next year and will be looking for solutions to introduce automated ways of performing some of the more complex measurements we make including colour, adhesion and resistance properties. Unlike the existing measurements, these all require a film of ink to be deposited and cured before a property measurement can be made.

A final focus will be to explore how we can get more value from the data that we generate. We will be investigating how data analytics and machine learning approaches could be used to augment our design intelligence.

Domino UK's Ink Automation Platform lead said:
"Our new automated formulation platform removes any hesitation to make lots of samples, allowing us to change how we approach our formulation experiments. The precision, accuracy and speed of robotics, coupled with the creativity of our scientists, is already yielding some really exciting results."

Acknowledgements

The first draft of this Playbook was written by Dr Matthew Reeves (Innovate UK KTN) and Prof Matt Reed (Materials Innovation Factory, University of Liverpool). It has benefitted from detailed comments by Michael Burnett, Steve Fletcher, and Peter Clark. It has benefitted from detailed comments from the wider KTN team: Michael Burnett (Knowledge Transfer Manager for Process Manufacturing), Peter Clark (Head of Chemistry and Industrial Biotechnology), and Steve Fletcher (former Head of Chemistry and Industrial Biotechnology).

The idea for writing the material came from an Innovate UK KTN organised event in February 2020: Going Digital - The Impact of Computational Tools on Chemical R&D, held at the Hartree Centre, Daresbury.

Funding

The Innovate UK KTN team received funding from Innovate UK KTN Materials & Manufacturing Team as well as from the Made Smarter Innovation Network. The University of Liverpool facilitated Matt Reed working on this Playbook.

References

1. Spataro, J. 2 years of digital transformation in 2 months. https://www.microsoft.com/en-us/microsoft-365/blog/2020/04/30/2-years-digital-transformation-2-months/ (2020).
2. Chemistry Council. The Chemistry Council Sector Deal. (2019).
3. Jones, R. On the UK's chemicals industry. http://www.softmachines.org/wordpress/?p=2513 (2020).
4. UK Government. Made Smarter Review. https://www.gov.uk/government/publications/made-smarter-review (2017).
5. Reeves, M. Innovate UK KTN survey. (2020).
6. King, R. D. et al. The automation of science. Science 324, 85–89 (2009).
7. Häse, F., Roch, L. M. & Aspuru-Guzik, A. Next-Generation Experimentation with Self-Driving Laboratories. Trends Chem. 1, 282–291 (2019).
8. Dar, Y. L. High-Throughput Experimentation: A Powerful Enabling Technology for the Chemicals and Materials Industry. Macromol. Rapid Commun. 25, 34–47 (2004).
9. Selekman, J. A. et al. High-Throughput Automation in Chemical Process Development. Annu. Rev. Chem. Biomol. Eng. 8, 525–547 (2017).
10. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. Nature 559, 547–555 (2018).
11. Goh, G. B., Hodas, N. O. & Vishnu, A. Deep learning for computational chemistry. J. Comput. Chem. 38, 1291–1307 (2017).
12. Mater, A. C. & Coote, M. L. Deep Learning in Chemistry. J. Chem. Inf. Model. 59, 2545–2559 (2019).
13. Burger, B. et al. A mobile robotic chemist. Nature 583, 237–241 (2020).
14. Compton, J. Data Quality: The Risks of Dirty Data and AI. Forbes https://www.forbes.com/sites/intelai/2019/03/27/the-risks-of-dirty-data-and-ai/?sh=4ccc5fec2dc7 (2019).
15. Morgan, J. A Simple Explanation of the Internet of Things. Forbes https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/?sh=5e80145e1d09.
16. Sinclair, M. The internet of things comes to the lab. Nature 542, 125–126 (2017).
17. Xie, C., Li, C., Ding, X., Jiang, R. & Sung, S. Chemistry on the Cloud: From Wet Labs to Web Labs. J. Chem. Educ. 98, 2840–2847 (2021).
18. Bereau, T., Andrienko, D. & Kremer, K. Research Update: Computational materials discovery in soft matter. APL Mater. 4, (2016).
19. Liu, A. J. et al. Opportunities in theoretical and computational polymeric materials and soft matter. Soft Matter 11, 2326–2332 (2015).
20. Rovigatti, L., Romano, F. & Russo, J. Topical Issue on Advances in Computational Methods for Soft Matter Systems. Eur. Phys. J. E 41, 41–43 (2018).
21. Lowe, D. The Downside of Chemistry Automation. Science https://www.science.org/content/blog-post/downside-chemistry-automation (2019).
22. Reed, M. Lab Automation for Materials Chemistry. Royce Inst. Pap. 0–54 (2021).

Connecting for
Positive Change.

Innovate UK
KTN

Head Office

Innovate UK KTN
Suite 218 Business Design Centre
52 Upper Street
Islington
London N1 0QH

Telephone: 03333 403251
Email: enquiries@ktn-uk.org
ktn-uk.org
@KTNUK

Contact Person
Dr Matthew Reeves, Innovate UK KTN
Prof Matt Reed, Materials Innovation Factory, University of Liverpool