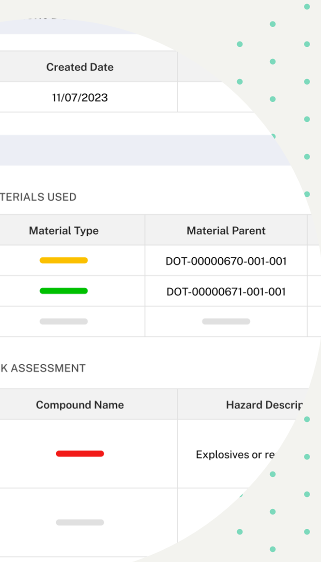
Strategy for Chemicals & Materials Innovation

R&D Digitalization with Data Intelligence in Mind



Dotmatics

Introduction

Digitalization of materials innovation in the chemicals industry has the potential to accelerate innovation and increase productivity but digitalization projects can be challenging to execute and often fail to gain end-user adoption. This whitepaper summarizes our experiences in successful digitalization projects and how to avoid common pitfalls. Beyond the fact that these projects must be viewed as long term business transformation efforts, rather than short term IT deployments, any R&D digitalization project must meet the following key requirements in order to be successful:

1. Adaptability to implement a highly diverse set of research workflows
2. A clear strategy on how to capture, share and leverage highly valuable research data
3. An approach to R&D IT infrastructure that reduces complexity

Challenges in Chemicals & Materials R&D

Overcoming Barriers to Materials Innovation

Developing advanced materials and meeting technological challenges has become increasingly expensive and time-consuming. To address this, specialty chemical companies are tailoring material properties for a wide range of applications, from mobile computing to aerospace and automotive light-weighting. Meanwhile, the complexity of R&D in cosme- and nutra-ceuticals is now rivaling that of pharmaceuticals, and agrochemical companies are using data-driven approaches to optimize crop growth in specific microclimates.

Market requirements are driving regionalization and product differentiation. For example, specialty chemical companies are serving a “market of one” for large customers in the aerospace and automotive sectors, while CPG companies are reformulating products to capture regional preferences and customer sentiments, leading to an explosion of products in their portfolios.

However, as companies consolidate, externalize research, and build strategic partnerships to tackle increasingly challenging research projects, collaboration has become mission-critical. Teams scattered across the globe, socialized in different innovation cultures, face challenges in effectively exchanging information and working together as a cohesive unit. These internal and external collaboration challenges must be addressed for successful innovation in the chemicals and materials industry.

Why Digitalize in the First Place?

The challenges outlined above have a common theme: They all have solutions which involve data being more efficiently generated, captured, leveraged and/or shared. The impact of digitalization can be conveniently divided into three major areas.

Efficiency gains and research capacity:

With researchers spending significant amounts of time on non-value-added tasks such as data pre- and post-processing, repetitive/canned reporting, or transcriptions, digitalization helps with automation both in data handling and workflow execution. Underlying solutions are often point solutions implemented at a department level (LIMS, ELN, LES, …).

The drive to innovate faster and the promise of data-driven R&D:

With the rise of artificial intelligence (AI) and machine learning (ML), larger organizations in particular are looking towards data science to shortcut lengthy lab experiments and use algorithms and simulation tools to identify new product opportunities. Apart from a data analytics platform, this requires the availability of high-quality data that often cannot be achieved with a fragmented point-solution landscape consisting of department-centric applications without common data infrastructure.

Collaboration and externalization:

Specialty chemical companies often sit in the center of an integrated innovation supply chain. They must source raw ingredients (often with highly variable specs) and produce materials that need to match narrow specification ranges provided by their customers. Innovation tasks may be handed off to 3rd parties or may come from mergers or acquisitions. Here, a digitalization infrastructure needs to provide data standardization and openness to facilitate the free exchange of information.

Based on this, a cohesive digitalization strategy builds value by:

a. putting enabling technologies into place (cloud/SaaS infrastructure, a single easily maintainable code base, etc.),

b. establishing personal productivity tools (automation of non value- added tasks),

c. developing operational excellence (data standardization, best practices, data-driven decision making), and finally

d. accelerating new product development (NPD).

In order to support this strategy, R&D digitalization needs to fulfill 3 key requirements: It needs to enable data capture ("data-in"), underpin operational excellence ("data-centric platform"), and ensure data can be provisioned to scientists, management and data analytics experts ("data-out"). See Figure 1.

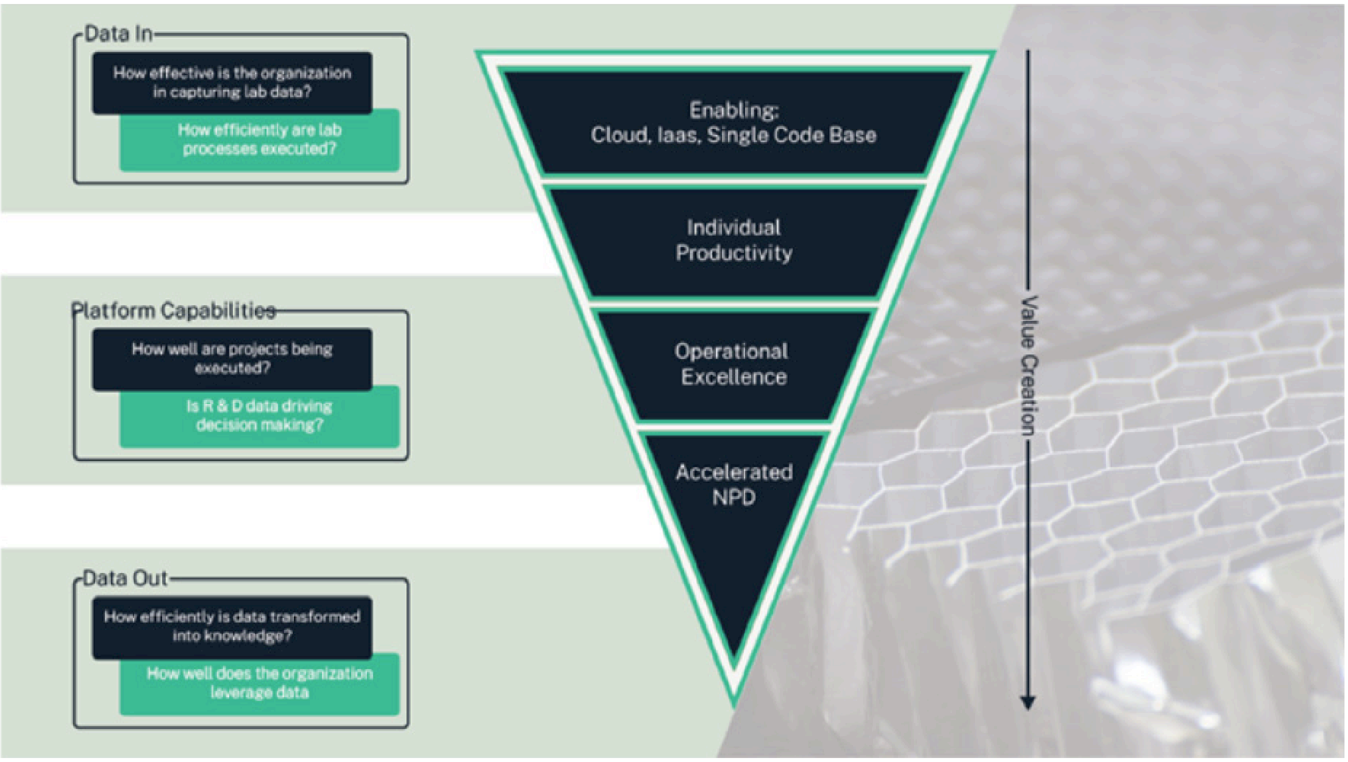


Fig. 1 - Value model for R&D digitalization. "Data-in" objectives generally align to either enabling on personal productivity improvement goals, while platform capabilities enable operational excellence. Finally, "data-out" capabilities are required for accelerated new product development (NPD), Which coincides with highest value creation

Problem Meets Solution

Why is R&D Digitalization so Difficult in Chemicals & Materials?

A digitalization project should not be seen as an IT project, but as a business transformation. This means that the challenges companies face go beyond configuration and deployment issues. Digitalizing R&D fundamentally changes the way people work. Some of the most common challenges are:

Managing the complexity and diversity of chemicals & materials innovation workflows:

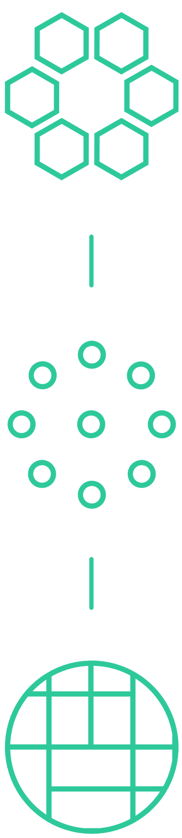
Materials innovation spans many different scientific domains, including organic/inorganic chemistry, formulations, process development, each with highly specialized analytical and physical characterization techniques. These teams often work scattered across the globe and might come together through acquisitions and mergers. Mostly, these teams have established their own workflows and created an inhomogeneous network of IT point solutions.

Addressing the change management bottleneck:

In many cases, digitalization involves taking well-established workflows that are considered tried and true, even if they are not particularly efficient. Analytical labs are generally facing exceedingly high workloads, meaning that digitalization efforts are perceived as disruptive and hard to fit into project schedules. Digitalization efforts should therefore aim at replicating and/or streamlining data entry procedures by either automating repetitive tasks or reducing the number of exposed data entry fields to minimize operator errors and improve data integrity/quality.

Defining a clear strategy for effective knowledge extraction:

Data standardization across sites, divisions and regions can be challenging, especially if data integrity issues exist due to poor end user adoption. Without a unified platform approach, data curation can become a major obstacle to leveraging data as a corporate asset. Planning and oversight may be based on outdated information and experimental planning efforts might either be time consuming or based on incomplete data.



Defining an R&D Digitalization Strategy that Works:

Chemicals and materials R&D organizations achieve digital excellence by mapping their high-level strategic objectives to their IT infrastructure. We see three objectives most commonly pursued: efficiency improvements, accelerated innovation and managing IT infrastructure complexity.

R&D efficiency improvements result from lab digitalization efforts, allowing higher lab utilization, increased testing throughput and shortened turnaround times. It also reduces non-value-added tasks such as data pre- and post-processing or repetitive reporting tasks. Furthermore, lab digitalization solves collaboration challenges and supports externalization efforts.

While increasing R&D efficiency boosts capacity, it does not fundamentally accelerate innovation. To do so, R&D organizations require data intelligence, or the ability to effectively leverage their research data. This allows for knowledge reuse, decision support and establishing best practices.

R&D digital infrastructure incurs complexity and cost. A viable strategy for R&D digitalization needs to take a long-term view on these costs and put measures in place to contain them. Some of the key requirements include maintainability, scalability and system availability.

Implementing a Data-Centric R&D Platform:

Considering the key role of data availability, a data-driven platform needs to be at the center of any strategic digitalization project. A data-centric platform allows for a unified view on all data and allows for the implementation of user roles and workflows. This requirement goes beyond an application-driven platform that has the sole purpose to integrate an application portfolio. Nevertheless, any platform should be open and able to integrate 3rd party data sources & applications.

In order to implement the vast variability of R&D workflows that can be found in materials innovation, configuration-driven interfaces are a way to create experiment templates that reflect domain data models. This means that all data types that belong to a specific experimental workflow can be grouped together and exposed as needed. With this approach, end user adoption is simplified as e.g. lab technicians are only exposed to data fields relevant to their work. It also means that data silos are eliminated, as an experiment template now ties together multiple data sources, such as structural, composition, performance and analytical characterization data.

To support data analytics initiatives and to allow scientists to innovate faster, digitalization needs “data-out” capabilities, i.e. the ability to perform queries against all aspects of a domain model. The query capabilities are best combined with a data visualization system that can help with data exploration, knowledge extraction and decision support. Interactivity with the data platform is key, e.g. to allow for design of experiment (DoE) or reporting capabilities.

“Considering the key role of data availability, a data-driven platform needs to be at the center of any strategic digitalization project.”

To illustrate the data-centric platform concept, we show a Dotmatics interface example on a query hit in a formulations experiment. Figure 3 shows how a researcher sees all relevant data in one view. Compositional data, with links to the inventory system, includes both planned and actual amounts, and where applicable, also links to formulations used as ingredients. Next, we see sample information with barcode, sample location and available amounts followed by analytical and application testing results. In traditional R&D IT systems, these data would be stored in dedicated but disconnected systems e.g. spec management, inventory, sample management, and potentially multiple LIMS systems used by different departments.

Once a query hit list is generated, users of a data-centric platform can pass the results on to a data visualization environment that allows for exploration of the entire materials design space. In Figure 4, a high-throughput experiment involving an inorganic catalyst generated performance data (conversion rate, mechanical stability) as a function of composition and processing.

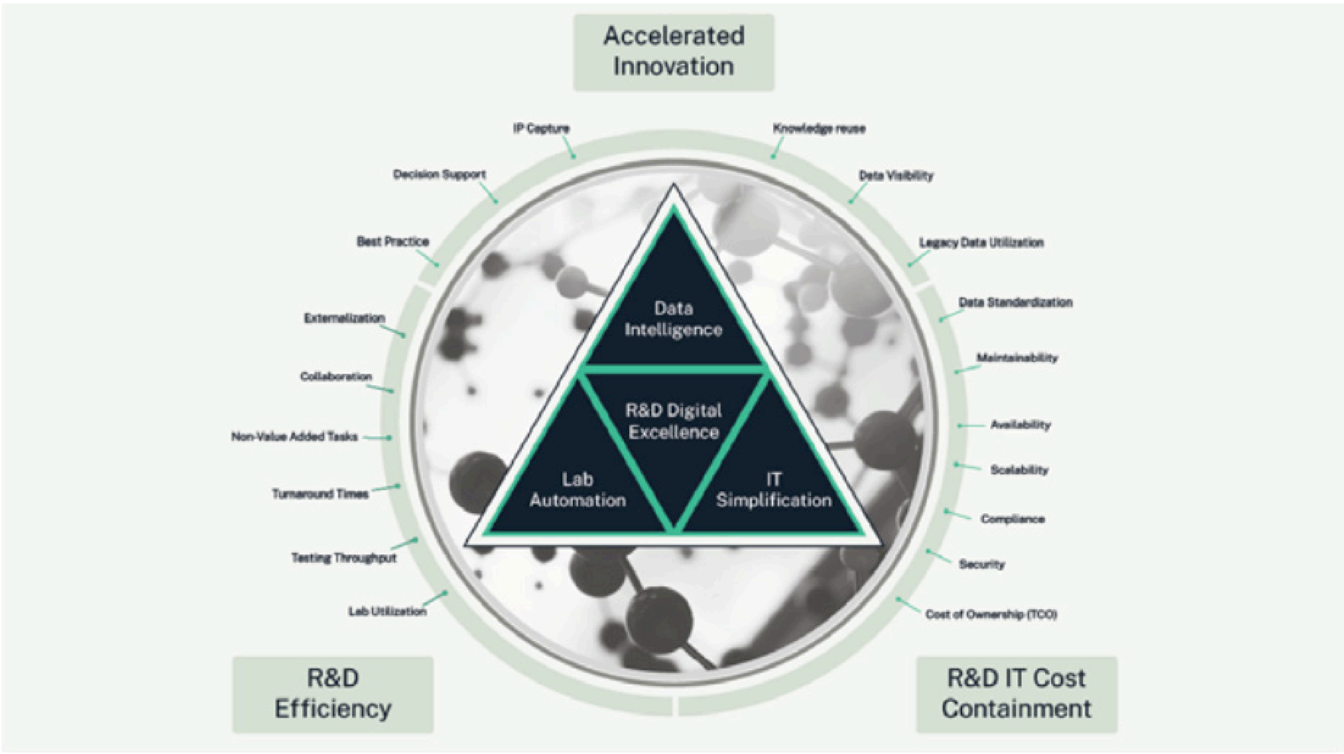


Fig. 2 - Mapping high-level objectives of R&D digitalization to infrastructure requires a comprehensive view on lab digitalization, data intelligence and containment of R&D IT infrastructure complexity.

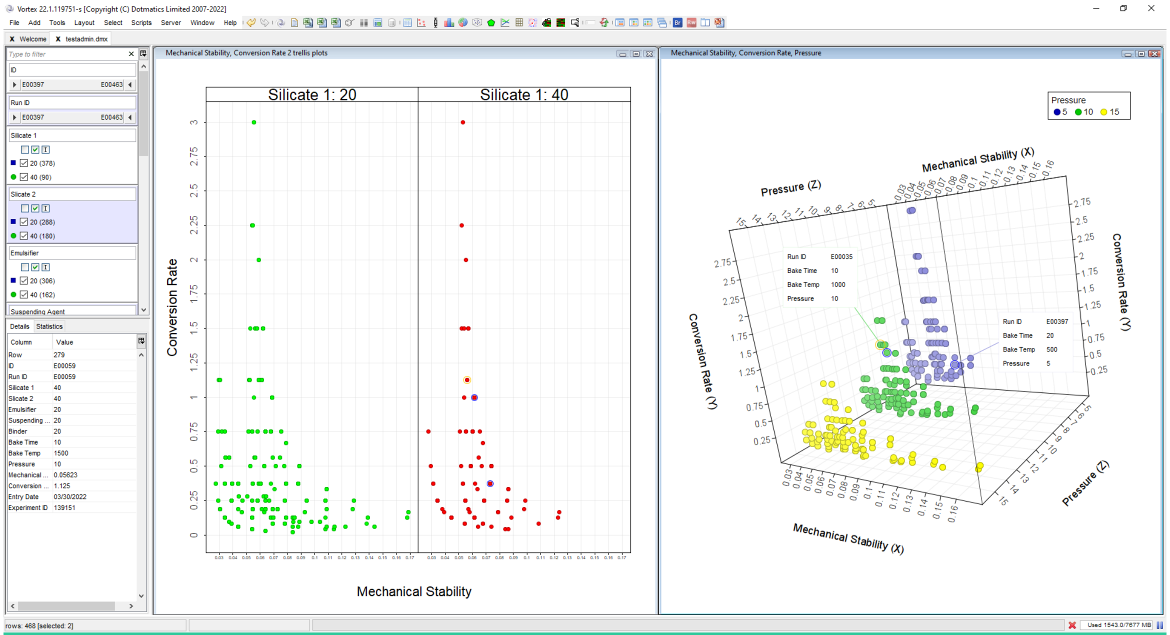


Fig. 3 - Dotmatics interface example. This query combines composition, sample, analytical and application testing data. All information is hyperlinked, allowing scientists to access related data contained in different data siloes.

FORMULATIONS

FORMULATION EXPERIMENT

Name	Experiment ID	Description	Scientist	Created Date	Completed Date	Protocol	Project
139206-1	139206	Zeolite	polly_mer	04/01/2022 01:56:58	04/01/2022 18:45:03	Formulation ELN	

FORMULATION SUMMARY

Name	Description	Amount Required	Amount Actual	Units	Mass % Planned	Mass % Actual
139206-1		100	101.5	g	100	100

FORMULATION INGREDIENTS

Tier	Addition Order	Name	Ingredient Name	Role	Active %	Amount Planned	Amount Actual	Units	Mass % Planned	Mass % Actual
1	1	139206-1	Aluminum Silicate	Default	100	20	21	g	20	20.69
1	2	139206-1	Zirconium Silicate	Default	100	40	40	g	40	39.409
1	4	139206-1	Pyrophyllite	Default	100	5	5.1	g	5	5.0246
1	5	139206-1	Sodium Magnesium Silicate	Default	100	20	20.4	g	20	20.099
1	6	139206-1	Fuller's Earth	Default Diluent	100	15	15	g	15	14.778

PROCESS

Stage	Operation	Duration (mins)	Component	Equipment	Rate	Temp	Validated?	Comments
Preparation	Add at Once		Zirconium Silicate	BAL-M332		20	Y	
Preparation	Add at Once		Aluminum Silicate	BAL-M332		20	Y	
Preparation	Add at Once		Fuller's Earth	BAL-M332		20	Y	
Preparation	Add at Once		Pyrophyllite	BAL-M332		20	Y	
Reaction	Heat	20		RCT05L-0001		500	Y	
Transfer	Dose at Rate	20	Sodium Magnesium Silicate	BAL-M332		500	Y	

SAMPLES

Sample ID	Barcode	Location Barcode	Container Location	Quantity	Units
139206-1-S1	PM03312022Z1	BOX_6	> UK> Lab 1> Shelf 2> Box 6	50	g

TEST RESULTS

Sample ID	Parent Exp ID	Service Exp ID	Test Type	Test Specification	Low Spec	High Spec	Measured Value	Spec Unit	Pass/Fail	Override	Sample Notes	Filename	Attachment
139206-1-S1	139206	139207	Tensile, Transverse	FTM-000008_01	2000	2300	2140	PSI	PASS			Tensile Strength Graph 1.png	

ANALYTICAL ATTACHMENTS

Sample ID	Experiment ID	Experiment Type	Filename	Extension	Attachment	Creation Date
139206-1-S1	139206	IR	FTIR_B.gif	gif		04/01/2022 02:05:28

Fig. 4 - Data visualization of a high-throughput experiment involving development of an inorganic catalyst. Data include composition, processing, and performance data.

The mapping of strategic objectives to R&D infrastructure discussed in Figure 2 also provides a convenient way to categorize functional requirements for a data-driven platform. Within lab automation, we find technologies required for data capture and workflow execution ("data-in") see Figure 5. These include electronic lab notebooks (ELN) that provide user front ends and implement the various domain capabilities (chemical synthesis, formulation, process development, analytical/ application testing). Together with request management and inventory/ sample management solutions, they can provide high-level LIMS workflows that broker the interactions between scientists and testing labs. Registration systems also fall into this category as they are a means to establish uniqueness of newly synthesized/explored materials.

On the "data-out" or data intelligence side we find data query and data visualization capabilities. Note that these play a completely different role from query interfaces in operational roles e.g. in sample management systems or ELNs (finding specific compounds or experiments). Data intelligence is the infrastructure part that enables knowledge management, data analytics and decision support applications. By providing access to the entire body of scientific data, researchers can understand how chemistry, composition and processing affects desirable material properties and guide them in next steps. Furthermore, the entire product life cycle benefits from data intelligence. During planning stages for product extensions, availability of existing data is key to avoiding unnecessary exploration of configuration spaces that have proven fruitless during previous studies. During pilot phases technicians may reach back to process control data to help scale-up. In regulatory or customer complaint scenarios, easy accessibility of research data is vital for quick turnarounds.

Fig. 5 - Mapping of R&D digital excellence to functional application areas.

In order to manage IT infrastructure complexity, the platform needs to provide various functions, including a data federation service to simplify tying in 3rd party data sources that can then be accessed by "data-out" technologies. Dotmatics has invested in building a cloud-first IT infrastructure, which means that the same system can be run on- premise, on the cloud, or hosted, providing distinct advantages in scalability, availability, accessibility and security. A key distinction between a data- centric platform approach and an application portfolio is how workflows are implemented. Within an application- centric approach, the end users carry out their work by accessing functionalities within a set of applications with rigid user interfaces. This often leads to end user adoption and change management issueswhen ways of working undergo a significant shift.

On the other hand, in a data-centric platform, workflows are configured with the applications acting behind the scenes to fulfill specific tasks. This is illustrated in Figure 6 (overleaf). For example, knowledge management solutions may be used at various stages throughout the innovation life cycle by many different roles (marketing, planning, task coordination, pilot, registration, regulatory). Configurability means that each of these roles can have specialized interfaces tailored to their needs. The same is true for lab technicians that access inventory, sample management and request handling functionality from dedicated interface configurations that closely match pre-digitization Workflows.

> "Dotmatics has also invested in building a cloud-first IT infrastructure, which means that the same system can be run on- premise, on the cloud or hosted, providing distinct advantages in scalability, availability, accessibility and security."

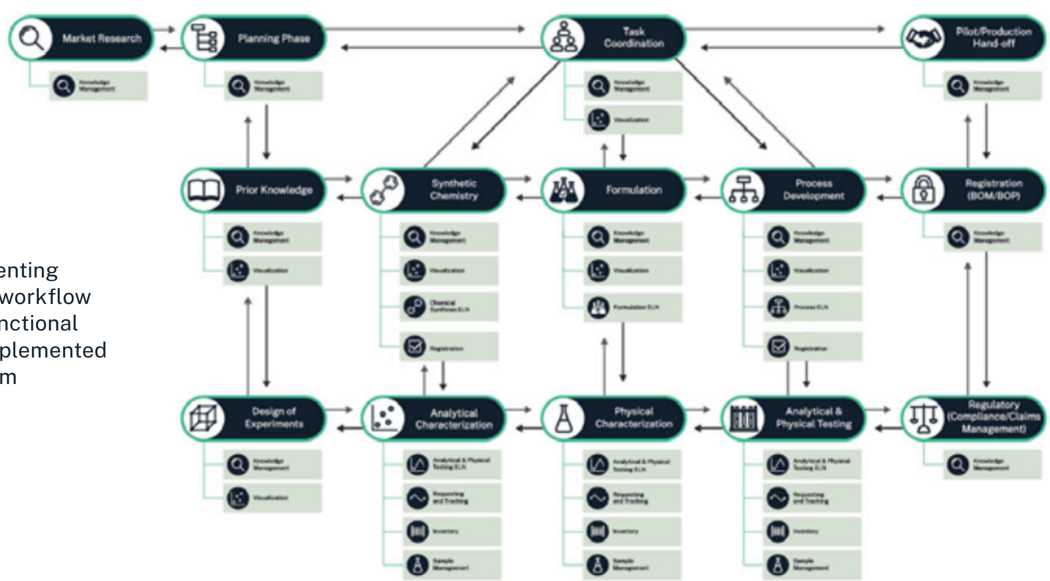


Fig. 6 - Implementing roles based on workflow analysis and functional applications implemented within a platform approach.

Dotmatics Chemicals and Materials Discovery Solution

For many innovation-driven companies, digitalization of R&D is already an important part of their overall corporate strategies. We expect its impact and presence to continue to grow. Value drivers such as individual productivity and operational excellence will remain a focus of lab digitalization. However, leveraging the data that digitalization can make accessible will only be available to those organizations putting a data-driven platform strategy into place.

Dotmatics is a fast-growing, science-focused company that develops its entire software suite by a co-located team of programmers and scientists. With a global network of support offices, Dotmatics prides itself on a partnership approach to engaging with science-driven organizations worldwide.

We invite you to contact us in order to discuss your specific R&D goals and needs.

With 15 years of experience in R&D digitalization and a highly skilled team of scientists, project management experts and product development professionals, we can provide you with the expertise and technology needed to successfully implement your digitalization projects.

Request Demo