

Blueprint and the LifeArc Ideas Factory

Leveraging informatics and modeling tools to develop a comprehensive Design Make and Test platform

LifeArc



Dotmatics

LifeArc

LifeArc is a UK medical research charity that has been bridging gaps between the lab and the patient for more than 25 years, providing the scientific know-how, technology transfer expertise and funding to help turn promising early-stage science into medical breakthroughs that transform lives.

With a strategy of ‘making life science life changing’, LifeArc partners with academics, industry, charities and patients to transform the way diseases are identified and treated. The charitable organization is committed to investing up to £1.3 billion (approx. \$1.75B) by 2030 to advance lab-based scientific discoveries into the next generation of diagnostics, treatments, and cures.

Medicines developed from LifeArc’s humanized antibodies are already helping patients with a range of conditions, from cancer and Crohn’s disease to multiple sclerosis and rheumatoid arthritis. A further eight potential medicines are in clinical trials. Recently, as part of the UK’s Therapeutics Antibody Taskforce, LifeArc helped screen over 600 antibodies to create a Covid-19 ‘antibody cocktail’.





A Variety of Therapeutic Areas, Modalities and Partners

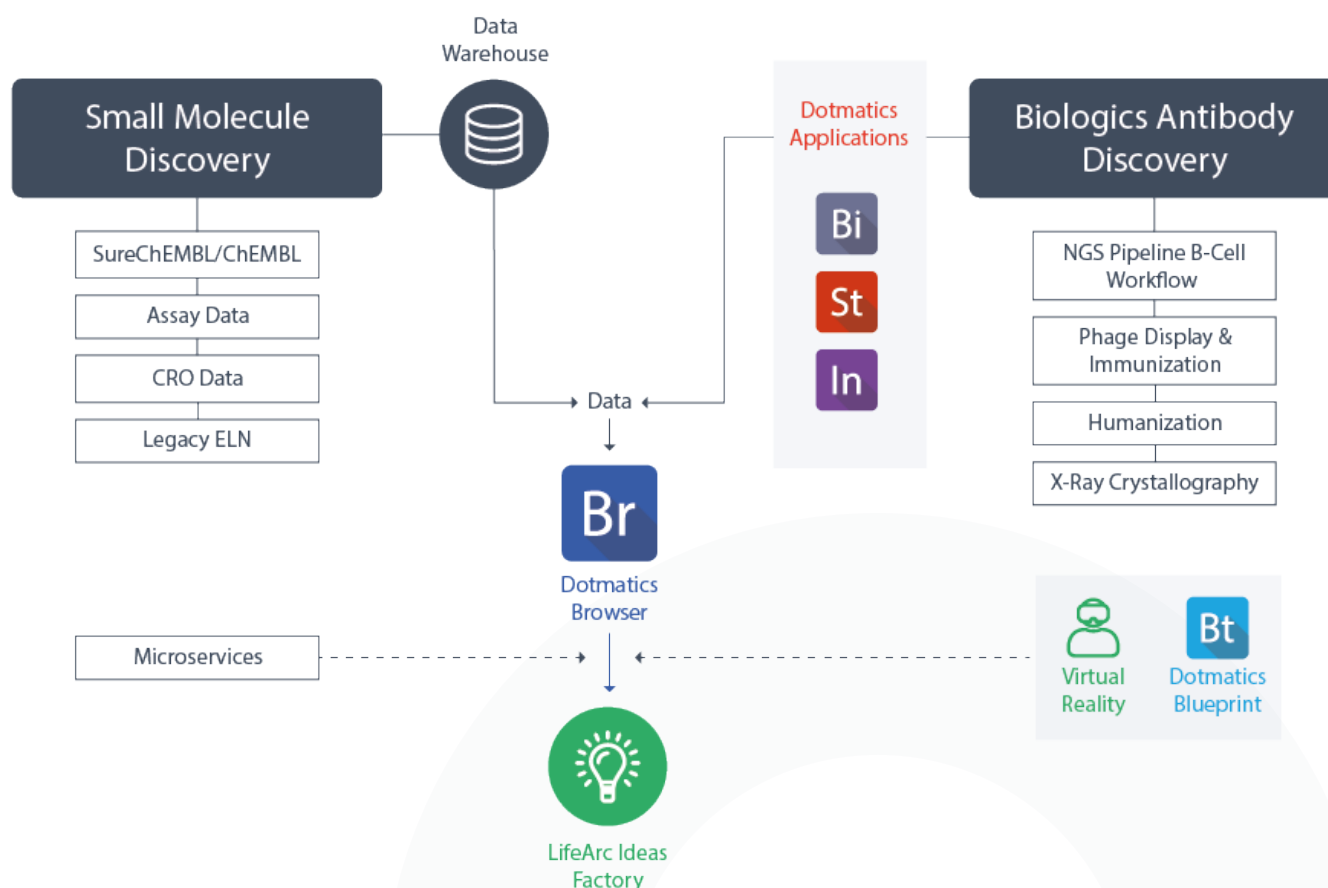
The Informatics Team is based at LifeArc's Stevenage therapeutics discovery center where a dynamic team of over 100 R&D scientists are dedicated to drug discovery across a range of modalities and disease areas. The scientific range across small molecules and engineered antibodies has implications for the type of data generated and how it's captured in an informatics system.

The therapeutics discovery team collaborate closely with diagnostic group colleagues in Edinburgh where molecular diagnostics capabilities are used to reassess drugs and targets.

LifeArc also partners with a number of external organizations including large pharma, small biotech and academia, and multiple CROs. As a result of these collaborations, data sources and types are varied, adding complexity to the informatics system.

“With such an eclectic mix of therapeutic areas and modalities, LifeArc needed an informatics platform that was fit for purpose. We believe that the Dotmatics platform gives us the flexibility to plug in and play any other tools that we have now or might need in the future.”

LifeArc Scientist



Informatics at LifeArc supports data from platforms for both antibody generation and chemistry, which is stored in operational databases. Antibody generation encompasses humanization activity, an NGS pipeline that generates human antibodies from a transgenic mouse, immunization and ADCs. The chemistry platform involves a HT Screening set of approximately 250k compounds. Additional operational databases store assays from CROs, and external public databases accessed and mined include TAB, CovAB and ChEMBL.

Scientists at LifeArc are constantly making decisions using visualization and analytics and it's important to have access to all data flowing between operational and public databases. The informatics team has a data warehouse into which clean, annotated data is pulled, using automated python ETLs, from all data sources. Microservices inject calculated properties into the data warehouse as well as further downstream into analytics.

LifeArc data warehouses for small molecule and antibody discovery include assay data, predicted data, annotations e.g. Unichem, ChEMBL, ontologies, and patent information via SureChEMBL.

Dotmatics Browser gives project-based views of the data that can be analyzed further in Vortex and Blueprint.

Entity registration and IP protection

Guarding intellectual property (IP) is intrinsic to ensuring that scientific discoveries are attractive to investors and progress along the innovation pipeline. At LifeArc, all biologics - antibodies, antigens, drug targets, or any entity generated from protein production units in the antibody drug discovery platform are registered using Dotmatics Bioregister. The informatics team have used an API layer to automate registration of clones from NGS sequencing of antibodies. Entities are screened using Studies Screening and tracked with Inventory.

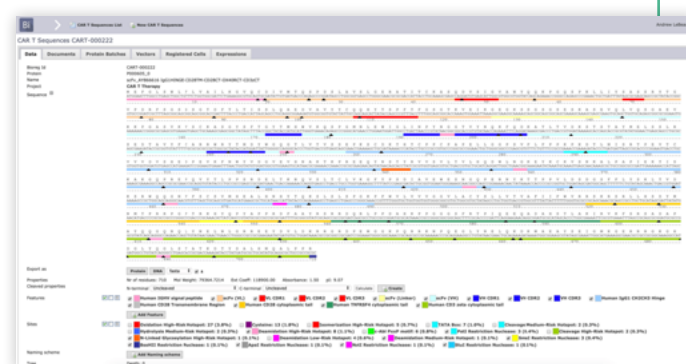
"The flexibility that Dotmatics brings means we are able to push a mix of data sources and operations into seamless workflows. A day-to-day example is our Antibody Drug Conjugate (ADC) workflow: a small molecule generated by an external CRO, with assay data and annotations stored in the small molecule warehouse, can be used in combination with an internally registered antibody to register a new ADC entity using Bioregister. This new ADC is subsequently screened with Studies."

"All data are available for scientists to use and easy to locate, which saves a lot of time. For example, it is possible to see if a compound has a SureChEMBL patent in Browser, without the need to look up elsewhere. Our growing team of data scientists can also re-use the data for predictive models, machine learning and AI."

LifeArc Scientist

Bioregister

Dotmatics Bioregister is a web-based application for registering sequence-based, chemically-modified and structureless biological entities, allowing biologics discovery organizations to ensure entity uniqueness and protect their intellectual property. Bioregister supports management of a broad set of biologics entities, including DNA, RNA, peptides and proteins, antibodies, conjugates, non-natural peptides and nucleotides, plasmids, cell lines and user-defined entities. It also enables users to record batches and samples for these entities, purification and expression information, and other protein production data.



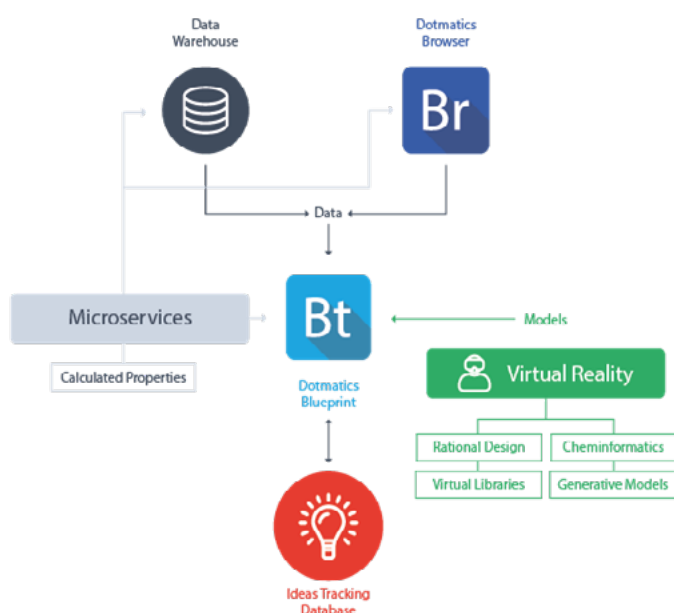


Blueprint and the Ideas Factory

With a flexible informatics system giving R&D scientists access to all of their data to streamline workflows, the LifeArc IT team considered the broader question of ideation, a process which historically separated informatics from ideas and ideas generation. They decided on the term 'LifeArc Ideas Factory' to describe leveraging available informatics and modelling tools to develop a comprehensive Design Make and Test platform.

Rational drug design ideation in the Ideas Factory involves bringing all data to the modelling suite and allowing scientists to use Dotmatics Blueprint in a concerted manner.

The system is built so that the data warehouse provides data via Dotmatics Browser. Blueprint and Browser are already well connected so micro services are used to generate calculated properties, machine learning or models that can be injected into the Blueprint environment. For example, the team have worked with Dotmatics to schedule a calculation for each molecule drawn by LifeArc scientists.



“There’s been a lot of talk over the past 5 years about how to capture ideation in the context of the data and the model and Blueprint has given us a pathway to be able to develop something that works for us.”

“One of the exciting things that we’ve been able to do is bring Blueprint into a VR environment so that you can pull your data into Blueprint and, while still in VR, you can do a search, (R-group analysis for example) come up with some ideas, draw them into Blueprint and then track them in the ideas database.”

LifeArc Scientist



Alongside the data and predictions, LifeArc scientists want to capture all ideas which come from multiple locations: a modelling suite, cheminformatics platform where chemists draw ideas, and more recently in a virtual reality (VR) environment to view molecular and target site context. Researchers have been inspired to collaboratively design compounds in VR where they can combine data from Blueprint and molecular models and track the results in an ideas database.

While still in VR, LifeArc scientists also use Dotmatics Chemselector. This allows chemists to look at their ideas, search publicly available databases for analogues or matched-pairs, pull them into VR, do some modelling and use Blueprint for further analysis.



Blueprint

Dotmatics Blueprint is a web-based visualization and scientific analytics application. It is designed to help project scientists working on small molecule discovery projects to visualize and analyze their data, and to help them to design their next set of compounds to make.



Enhancements to the Ideas Factory

R&D scientists have enthusiastically embraced the Ideas Factory and are inspired to suggest additional capabilities. The informatics team are collaborating with Dotmatics to use Dotmatics Reaction Workflows with Blueprint to consider which compounds can be based on in-house reagents and reactions.

Another interesting project will involve navigating all reactions in the LifeArc databases to design molecules within the context of the data available.

“What’s really useful is seeing chemists use Blueprint within the VR environment and looking at the data while they’re looking at the crystal structure. The idea of activity cliffs makes more sense in 3D - you can say, “this methyl is magic because we predict it is binding in this way,” and when you can see Blueprint in VR it generates a lot of discussion. Suddenly scientists are excited about making specific modifications from looking at both the data and the model”

LifeArc Scientist



Contact

Info: info@dotmatics.com
Sales: sales@dotmatics.com
Support: support@dotmatics.com
Careers: careers@dotmatics.com

Dotmatics