

InBrief

Philip Howard – Research Director, Information Management
 Daniel Howard – Senior Analyst, Information Management and DevOps

“Datactics’ software provides a twenty-fold accurate throughput on product catalogue matching and help us deliver high quality service to our customers.”
 NHS Supply Chain

Datactics

The company

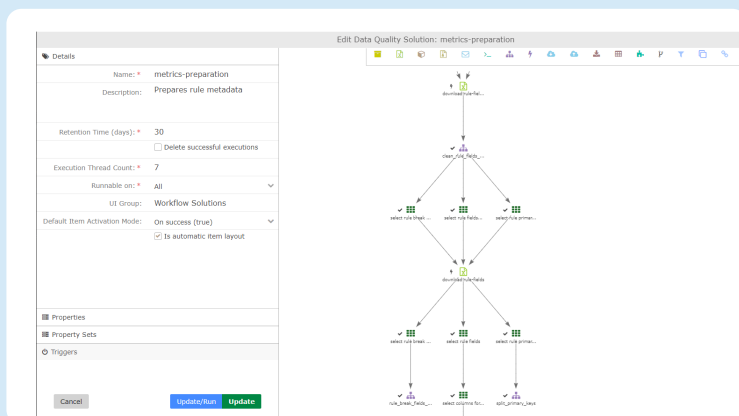
Datactics is a Belfast-based software vendor with nearly 50 employees. It was established in 1999, and it focuses on providing self-service and AI-driven data quality solutions, most notably offering “explainable AI” as a differentiator.

What is it?

Datactics is a data quality platform oriented around self-service. It provides solutions for data quality, data matching, and single customer view, and it puts particular, and notable, emphasis on supporting these capabilities with highly transparent and explainable AI and machine learning. This allows you to benefit from the advantages of AI (smart data matching, for example) while ensuring that it remains accountable, comprehensible, and trustworthy.

The Datactics solution is comprised of a number of separate elements, some of which we will describe in detail in the next section. These include Data Quality Manager (DQM), which allows you to manage and automate your data quality rules and process, and in addition acts as a central hub for administration of the Datactics platform; Data Quality Clinic (DQC), which allows you to manually investigate and remediate low quality records; FlowDesigner, a visual, drag-and-drop rules designer; and the AI Engine that governs the prediction and suggestion AI that runs through it all. The latter is enhanced by ML Monitor, a model monitoring dashboard that leverages MLFlow (an

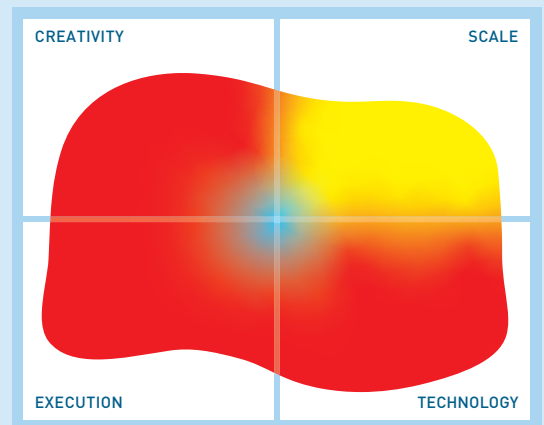
Figure 1 – Datactics FlowDesigner



DATACTICS™

www.datactics.com

Unit C1, Suite B, 2nd Floor
 Lanyon Quay, Belfast BT1 3LG
 Tel: +44 (0)28 9023 3900
 Email: info@datactics.com



The image in this Mutable Quadrant is derived from 13 high level metrics, the more the image covers a section the better. Execution metrics relate to the company, Technology to the product, Creativity to both technical and business innovation and Scale covers the potential business and market impact.

open source platform for managing machine learning lifecycles). In addition, the platform provides more traditional dashboards to expose its data quality results, as well as integration with third-party tools such as Qlik, Tableau, and PowerBI. Note also that the AI engine exists outside the platform itself and is accessed via APIs (which also exposes it for use by other parts of your system, if so desired).

What does it do?

Datactics uses rules to drive your data quality processes. These rules are created in FlowDesigner, using a visual, drag-and-drop framework, which can be seen in Figure 1. They can be used as you would expect, to create matching rules, construct datasets, identify low-confidence results, and so on.

These rules are underpinned by the Datactics AI Engine. For Datactics, the emphasis is on “explainable AI”, meaning that the models used by the AI engine are designed to be transparent, reproducible, and ultimately trustworthy and accountable. These models also cater to a variety of use cases, and are not strictly restricted to data quality (graph linkage prediction and automated dataset labelling, for example, are both supported, and have significant applications beyond quality). That said, for the purposes of data quality, the most notable applications are AI-augmented data matching and deduplication, and error detection.

Architecture	★★★★★
Automation	★★★★★
Cleansing	★★★★★

Ease of use	★★★★★
Extensions	★★★★★
Matching	★★★★★
Profiling	★★★★★

“The platform’s usability means our SMEs can now develop data quality rules themselves – extending the functionality to multiple regulatory requirements and broader data quality and governance applications – and automatically generate meaningful interactive reports in Tableau, our house reporting tool.”
Retail Bank

In particular, the former is focused on automatically pruning low-confidence match results that would normally need to be manually reviewed (and often arrive in large numbers), thus saving your data quality analysts a significant amount of time and effort.

In addition to providing this level of automation, these AI-enabled data quality processes will generally be more comprehensive, more systematic, and ultimately more accurate than their manual equivalents. Moreover, Datactics’ AI models benefit from unsupervised learning, allowing them to improve and grow more accurate over time without

Why should you care?

There are two big draws to Datactics: the emphasis it places on self-service, evident in its user-friendly design and in particular the visualised, drag-and-drop rules creation provided by FlowDesigner; and the particularly transparent and explainable AI and machine learning functionality that it offers.

For the former, it’s worth caring about because it enables non-technical users to engage with your data quality processes. Datactics describes this as a “logical next step” from self-service analytics, and it provides many of the same benefits, allowing business users that are already equipped with domain knowledge to put that knowledge to use in data quality, enabling those users to get at the high quality data they need more quickly and without going through IT, fostering collaboration between technical and business users, and freeing up IT to work on other things.

Explainable AI, on the other hand, allows you to reap the benefits of AI – such as increased automation, and thereby efficiency – while at the same time making sure your AI models are comprehensible and can be meaningfully understood. In turn, this should allow you to trust the decisions and predictions made by your models. If a model makes what turns out to be a bad call, it’s easy to examine why that call was made. Moreover, being able to explain how your AI models are using your data goes from important to essential when you consider data protection regulations such as GDPR and CCPA. To comply with these regulations, you must be able to explain how your models process any given individual’s personal data. To this end, explainable AI is a very useful tool indeed.

“With Datactics we have a data quality technology that both significantly enhances our customer-marketing function and provides a platform for progressing the information quality agenda for future strategic projects.”
Retail Bank



Figure 2 – Data quality overview in Datactics

manual intervention. Data quality and data matching rules and checks can be executed in DQM, and the results are displayed in a variety of views and dashboards, one of which is shown in **Figure 2**. Notably, these dashboards can include prediction reasoning, and the data matching results view in particular features a prominent ‘explain decision’ button. The latter does what it says, using visualisations to show you which data features led to that particular decision being made and both of these elements provide good, useful examples of Datactics’ “explainable AI” philosophy in action.

The last element we should talk about is ML Monitor, the platform’s model monitoring dashboard. This allows you to track data drift, model performance, and so on. It also uses MLFlow for model management (for example, tracking model versions).

The Bottom Line

Datactics is a self-service data quality platform that puts significant stock in, and devotes substantial functionality to, the idea of explainable AI and machine learning. As fervent proponents of explainability in AI, we certainly approve.

[FOR FURTHER INFORMATION AND RESEARCH CLICK HERE](#)