

Cite as: Sorcha Forde, Eagle Genomics – Unleashing the power of genomics, Innovation & Impact, 2017

URL: iai.pubpub.org/pub/eagle-genomics/

DOI: [10.21428/14612](https://doi.org/10.21428/14612)

Eagle Genomics: Unleashing the power of genomics

Sorcha D. Forde

Abstract: Eagle Genomics is a bioinformatics and smart knowledge management company based in Cambridge, U.K. As a pioneer in data-driven discovery and innovation, they provide invaluable question-driven software solutions for the life sciences. NGS-based research has resulted in an explosion of genomic data production. Eagle Genomics had the foresight to realise that this data would require optimal storage, organisation and exploitation. They set out with a vision to provide scientists with data management solutions to exploit and extract the most valuable information from their genomic data. Indeed, Eagle's award winning smart data platform has revolutionised data access and management in the life sciences industry. Their smart data management platform transforms big data into actionable insights by applying a 4-step work flow involving curation, cataloguing, discovery and exploitation. This task is undertaken by the award winning portfolio of Eagle products and enables new and better insights to be found more efficiently.

Keywords: Genomics, Bioinformatics, NGS, eaglecore, eaglediscover, meta-data, Cambridge

1. The success story

Eagle Genomics is a Cambridge based company founded in 2008 that specialises in bioinformatics, data curation and smart knowledge management. They provide rare expertise to life science clients who wish to use genomic data to accelerate the development of new products – be it drugs, agricultural or personal hygiene products. Eagle Genomics founding mission was to enable scientists and researchers to exploit and extract the most valuable information from the huge amount of genomic data produced both in-house and available publicly. Their success can be attributed to the powerful combination of scientific expertise, both in Life Sciences and IT with sharp business acumen. Eagle's executive management team includes top bioinformaticians, whose achievements include leadership roles in several leading open-source bioinformatics projects that have defined the way research in the field proceeds. This wealth of expertise has positioned Eagle as leaders in management and analysis of genomic data and the company's software solutions exploit smart data science and deep statistics to

Sorcha D. Forde: Division of Molecular Histopathology, Department of Pathology, University of Cambridge, Addenbrooke's Hospital, Cambridge CB2 0QQ, United Kingdom , Email: sf533@cam.ac.uk

© Sorcha D. Forde, published at the pubpub.org, under the Creative Commons Attribution 3.0 License.



radically reduce time to complete and cost of research. The overall result is drastic productivity improvements and true data driven discovery. Eagle's established commercial licensing partnerships with many of the world's leading pharmaceutical and biotech companies are aiding in more effective and efficient use of genomic data to bring new products and services to the market, improving customers' business processes, product innovation & commercial performance.

2. How did they start

Eagle Genomics is the brainchild of Abel Ureta-Vidal (now CEO) and William Spooner (now CSO). Both Abel and William had experience working at the European Bioinformatics Institute and combining scientific and business expertise had the foresight to realise that the huge explosion of Next Generation Sequencing (NGS) research in the mid-2000's would result in the need for appropriate storage, organisation and exploitation of this genomic data. The venture began by providing genomic consulting to firms using Ensembl, a powerful tool for visualizing and searching genome data. By 2008 the increased demand resulted in the establishment of the Eagle Genomics company and the evolution of multiple products and services. Initially, the focus was on genomic data integration and analysis around the Ensembl platform with Eagle having a partnership agreement with EMBL-EBI and the Wellcome Trust Sanger Institute providing privileged access to the development team and know-how behind the Ensembl project. Working from home and with low overheads, the company was able to grow and expand its product and services repertoire. In 2009 with the emergence of Amazon Web Services and the closing of Eagle's first £1 million round of investment in November 2013, Eagle scaled up with the eventual development of eaglecore and eaglediscover [1]. Together these products offer invaluable resources to companies to accelerate data usage and optimise data exploitation. 'The New York Times' used the term 'data janitor work' to account for the laborious hours spent 'wrangling' through big data. In fact, it is thought that data scientists spend between 50 and 80% of time collecting and organising complex data sets [2]. Eagle's products aim to cut these figures drastically by allowing the exploration of metadata rather than the raw data itself.

3. Their technology

Eagle Genomics has designed eaglecore as a data federated, secure platform from which to deliver an integrated web-based solution for information management of large-scale experimental datasets. eaglecore easily manages and monitors experimental information within and between organisations allowing for acceleration of the key components of research - discovery, integration and collaboration. It helps researchers spend less time finding information and more time developing new products for pharmaceuticals, biotech and fast moving consumer goods. Using metadata as the integration layer to bridge between datasets, eaglecore provides consistent access to underlying assets, irrespective of file format or data type. eaglecore encourages data reuse, allowing for further analysis or reuse in new experiments and the secure platform promotes collaboration by enabling sharing of datasets. Eaglediscover then built upon this by offering a unique web-based platform for the

measurement and profiling of data value, Figure 1. By providing users with prioritized data based on scientific value, eaglediscover improves data selection, encourages data reuse and informs experimental programs. A contextually-relevant metadata catalog is created, with the data being statistically scored for relevance based on the scientific and business questions being posed. eaglecore can then be employed to manage the data.

Using robust quantitative and probabilistic techniques based on decision theory eaglediscover objectively and statistically measures the value (as opposed to only quality; quality is only a component of value but by no means the only one) of data assets as defined by their usefulness and contextuality. The catalogs created allow researchers to conveniently find what data is available, where it is, who produced it and what format it is in. This has the key benefits of increasing productivity and reducing time spent on data wrangling. The models are hierarchical; an expert-driven process assigns scores across various dimensions (value components) according to multiple stakeholder perspectives.

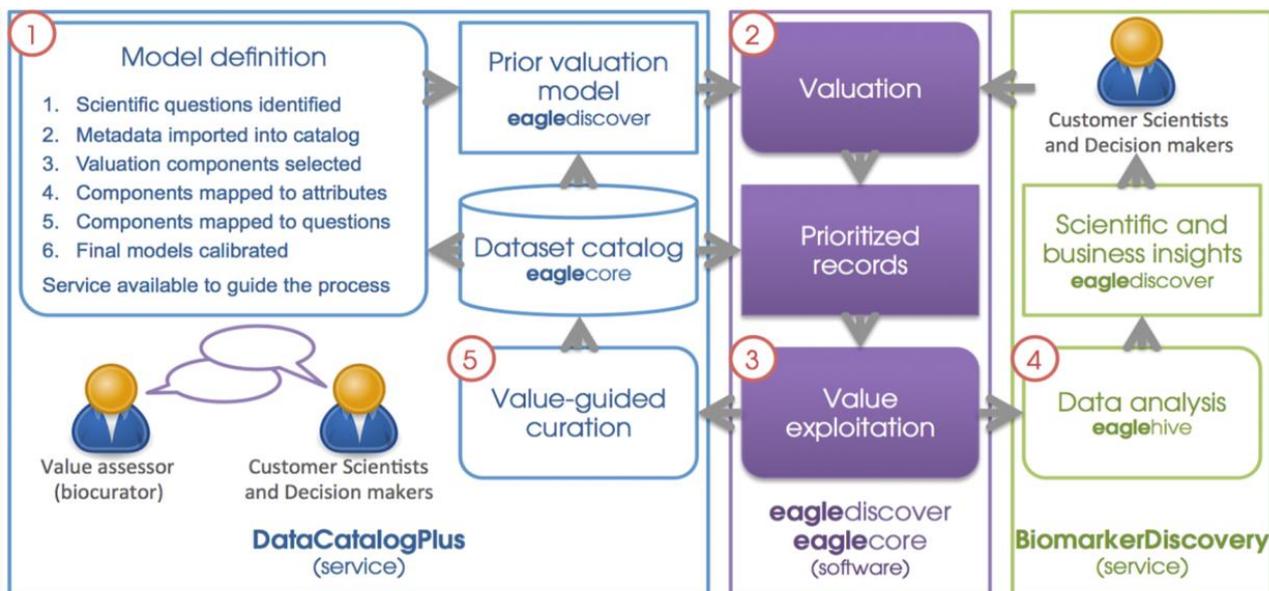


Figure 1. Eaglediscover: Overall process from data valuation to usage and exploitation.

Eaglediscover was awarded Best of Show at the Bio-IT World Conference and Expo 2016, demonstrating novel insight on the world’s largest public cancer patient records, the International Cancer Genome Consortium [3].

4. The journey so far

Eagle Genomics began with two people and is continuously growing, now with over 20 employees providing invaluable question-driven software solutions for the life sciences. Eagle Genomics has helped one of the world’s leading research-based pharmaceutical companies unravel the secrets within terabytes of cancer genetic data by introducing big data workflow technology using expression profiling of RNA sequence reads (RNA-Seq). This allowed development of a process to compare the RNA sequences of 100 tumors in a clinical study

accurately and efficiently with Eagle providing access to sufficient computational resources to enable storage and processing of the terabytes of data. Scalability is paramount to big data projects and Eagle's workflow was tested when the company wished to move to the next level - processing and interpretation of expression levels in the RNA-Seq data – which was a smooth and problem-free process. Eagle has also had a huge role in supporting another company, Unilever's, digital data program over the years. They needed a scalable solution that would provide parallel computing capabilities to keep up with the demands of R&D, as well as be flexible so that it could just use the extra computing capacity when it was needed. Eagle's cloud solution provided an opportunity for limitless expansion and enabled parallel analysis. The system was created to be secure, stable, robust and user-friendly, with low running costs. The result was 20 times faster genomic data analysis with the architecture supporting ten times as many scientists working simultaneously. Eagle's strength in exploiting data to yield its maximum value was evident in its work with Discuva. As well as requiring assistance in scaling up their antimicrobial resistance NGS profiling 10-fold they also needed a more efficient tool for comparing compound effects on bacteria. Eagle Genomics created a tool that could pull data out of the database and build a matrix to examine the relationships between compounds. A dendrogram is produced which gives a visual representation of the closeness of relationships between the mechanisms of action, helping researchers to determine the risk and rate of development of cross-resistance. Eagle Genomics believes that researchers are on the tip of major breakthroughs, such as new methods for antenatal and cancer diagnostic testing. These efforts are increasingly inhibited by lack of data transparency, inability to share and compare relevant information and access datasets in a comprehensive and logical way. By addressing these challenges Eagle believes they will expedite development of products that promise to improve quality of life worldwide.

5. Looking to the future

After eight years working very closely with many customers in the Life Sciences, dealing with the ever-growing amount of data they must deal with, Eagle Genomics has accumulated an unrivaled insight into the pain points and challenges the industry is facing. Over the years, they have built a rich network of collaborations both in industry and academia. As of today, in late 2016, Eagle Genomics is taking the high road and is pursuing even bigger opportunities with the aim of delivering the reference information architecture necessary to address the challenges, risks and opportunities of the genomics era.

Acknowledgments

We would like to thank Dr. Abel Ureta-Vidal from Eagle Genomics Ltd. for invaluable discussions regarding Eagle Genomics technologies, the company's scientific and business development milestones and success stories, and reading of this manuscript.

The company



Eagle Genomics
The Biodata Innovation Centre
Wellcome Genome Campus
Hinxton
Cambridge, CB10 1DR
United Kingdom
<https://www.eaglegenomics.com/>

References

- [1] <https://www.eaglegenomics.com/eaglediscover/>
- [2] For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights, The New York Times, Steve Lohr, Aug. 17, 2014.
- [3] eaglediscover Wins Bio-IT World Best of Show Award for Eagle Genomics, Business Wire, April 12, 2016.

Sorcha D. Forde has a Bachelor's degree in Molecular Medicine from Trinity College Dublin and is currently a final year PhD student at the University of Cambridge. Her research area is in paediatric lymphoma, with a focus on relapse in B Cell Non-Hodgkin Lymphoma.

Sorcha has a strong interest in innovations that help researchers extract the most possible data from their work and she is involved in several such ventures. Her interest in innovative technology development was strengthened during her time as a member of Innovation Forum and her organisation of the IF16 global conference in Cambridge. Sorcha also has an interest in consulting, particularly helping start-ups enter the market and is a member of the Cambridge Consulting Network.

