Chief Information Officer
It’s an exciting time to join us as we continue to build an international centre for scientific, business, cultural and educational activities emanating from Genomes and BioData. With a significant expansion of our Wellcome Genome Campus on the horizon, we are now able to shift our horizons to ask and answer even bolder questions.

Our mission is “to maximise the societal benefit of knowledge obtained from genome sequences”. The ambition of Genome Research Limited, with our Campus partners, is to progressively strengthen its well-established foundations in scientific research and discovery, and to build on them, developing the Wellcome Genome Campus over the forthcoming 25 years.

Genomic research is still in the foothills of extracting and using the knowledge buried in the 6 billion letters of code in the human genome. The ever increasing numbers of human genomes sequenced for research or clinical diagnosis will reveal patterns and motifs that will shape health and disease research for decades to come. When we also consider the rest of the genomes on Earth the potential is vast and the Sanger Institute will be in the vanguard of this revolution in science and society.
Wellcome Sanger Institute

The mission of the Wellcome Sanger Institute is “To use information from genome sequences to advance understanding of biology and improve health.”

The Wellcome Sanger Institute is a world leading genomics research centre. We undertake large-scale research that forms the foundations of knowledge in biology and medicine. We are open and collaborative; we share our data, results, tools and technologies across the world to advance science. Our findings are used to improve health and to understand life on Earth.

Our science is large-scale and organised into Programmes, led by our Faculty who conceive and deliver our science, and supported by our Scientific Operations and IT teams responsible for all data production pipelines, compute and storage facilities at the Institute.

With secured funding from Wellcome Trust, we are able to strategically focus our work in five key research fields:

Cancer, Ageing and Somatic Mutation
Provides leadership in data aggregation and informatics innovation, develops high-throughput cellular models of cancer for genome-wide functional screens and drug testing, and explores somatic mutation’s role in clonal evolution, ageing and development.

Cellular Genetics
Explores human gene function by studying the impact of genome variation on cell biology. Large-scale systematic screens are used to discover the impact of naturally-occurring and engineered genome mutations in human iPS cells, their differentiated derivatives, and other cell types.

Human Genetics
Applies genomics to population-scale studies to identify the causal variants and pathways involved in human disease and their effects on cell biology. It also models developmental disorders to explore which physical aspects might be reversible.

Parasites and Microbes
Investigates the common underpinning mechanisms of evolution, infection and resistance to therapy in bacteria and parasites. It also explores the genetics of host response to infection and the role of the microbiota in health and disease.

Tree of Life
Investigates the diversity of complex organisms found in the UK through sequencing and cellular technologies. It also compares and contrasts species’ genome sequences to unlock evolutionary insights.

Connecting Science

Wellcome Genome Campus Connecting Science’s mission is to “enable everyone to explore the impact of genome science on health, research and society.”

They connect researchers, health professionals and the wider public, creating opportunities and spaces to explore genomic science and its impact on people. Connecting Science inspires new thinking, sparks conversation, supports learning and measures attitudes, drawing on the ground-breaking research taking place on the Wellcome Genome Campus.
Information Technology at Sanger Institute

The last three decades have witnessed an extraordinary increase in the amount of data generated from biological systems. This has primarily been initiated by the advent of systematic genome sequencing, and fuelled by iterative reductions in its cost, but other sources, including biological images, are increasingly contributing. In consequence, there has been a transformation in the influence of computational approaches in biology; to store, organise, analyse, share and present these data. Together, these changes are having a substantial and pervasive impact on biological science and its applications in medicine, agriculture, management of the environment and in other contexts.

These changes continue to accelerate and we therefore have a pressing need to more closely integrate data, informatics and IT through an integrated strategy. To help drive this strategic change we are creating a senior CIO role to lead in the coordination and delivery across the whole organisation.

Our IT infrastructure supports the entire workflow of data, from data generation to storage in databases and archival systems, research analysis by Faculty teams and publication on our website. To service the demand from DNA sequencing and other data analysis, the Institute has:

- An IT team of around 60 people with additional expertise across out research programmes and scientific operations
- Centre-based high performance compute cores 25,000 approx.
- Centre-based cloud-based flexible compute cores 20,000 approx.
- Usable storage in the Data Centre 65-70Pb approx.
- 20Gbit resilient connection to JISC
- Network backbone speed 400GB/sec

Working with our partner organisation, EMBL-EBI, there are significant benefits to be realised in harmonising our approach to data and informatics, including compression software, and developing the right tools for our researchers so they can achieve their scientific and translational goals.

There are a number of strategic goals in our 2021-2026 timeframe:

- Reliable and robust IT services supporting the campus’ scientific and business activities – encouraging resilient distributed Informatics software approaches.
- World-class customer services – solutions driven by customer need with exceptional customer service.
- Certified/Accredited IT systems – provide secure IT Infrastructure & Services.
- Research capabilities to support the needs of our Science Programmes, e.g. long-read sequence assembly and managing imaging data at volume – evaluate technologies and create innovative software and hardware solutions to scientific challenges.
- Flexible, dynamic environment(s) for large-scale genomes and biodata analysis – creation of software services which make public/private cloud transparent for our researchers.
# Role Profile

<table>
<thead>
<tr>
<th>Current job title:</th>
<th>Chief Information Officer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade</td>
<td>PSG1</td>
</tr>
<tr>
<td>Reports to:</td>
<td>COO</td>
</tr>
<tr>
<td>Management responsibility for:</td>
<td>IT (scientific, high performance, cloud computing, data storage, core software services, database management, service delivery, business systems support, cyber security) and the core informatics team</td>
</tr>
</tbody>
</table>

### Interacts with:

- **Strategic Informatics Committee (SICom)**: Present for discussion and sign-off the Institute’s Technology Strategy. Present progress against the strategy and bring any changes for decision to the committee.

- **Science programmes**: To establish researchers’ requirements for scientific data storage and computing, and core informatics and deliver them.

- **Management Operations**: To establish business enterprise needs for the organisation and deliver them.

- **EMBL-European Bioinformatics Institute**: To ensure a partnership approach to IT between the organisations

- **External partner Institutes, suppliers and providers, other entities on the Genome Campus, large funding agencies and the wider research community**: To contribute to and promote the Institute as a thought leader in scientific IT, informatics and setting data standards in Genomes and Biodata.
**Role purpose and primary objective:**

- Translate Institutional scientific objectives to an organisational data and informatics strategy and through this specify the requirements for underpinning technologies (e.g. hybrid of on premise and public cloud) and business and scientific services (e.g. in-house and external providers).

- Leadership of IT and core informatics staff, with oversight and delivery of a wide range of programmes and projects to deliver clear business improvements and benefits across the organisation and drive efficiencies in both data storage and computing.

- Strategic and operational responsibility for all information and communication technology services and infrastructure across the organisation.

**Core accountabilities [in approximate order of importance to role purpose]:**

1. Develop and maintain a technology strategy to ensure the enduring integration of IT and data management across the organisation and establish a core informatics team to support this.

2. Working with key stakeholders to create a data strategy among the Science Programmes and Scientific Operations.

3. Lead the organisation in identifying, aligning and resolving institutional data storage and computing challenges.

4. Undertake robust management of the data life cycle – related to both the cost of data storage, by reducing long-term data storage where possible, and maintain regulatory (e.g. GDPR) compliance with our data assets.

5. Improve reproducibility and portability of informatics workflows, develop and maintain software as shared tools and platforms to operate on our on premise and offsite infrastructure and maintain accessible libraries of these tools.

6. Identify opportunities and develop further our technology partnerships with both hardware vendors and informatics organisations, especially with regard to data management, computation and collaborative environments.

7. Build on our hybrid model of on premise and public cloud computing: increasing our computing flexibility, responsiveness to scale, collaborative opportunities and data interoperability, while not compromising data security or integrity.

8. Drive technological advances to improve research and business capabilities and the cost effectiveness of IT, informatics and data solutions.

9. Facilitate the development of scalable collaborative research environments with diverse partners.
Describe the most complex/challenging aspects of the role

There is a need to establish a strategy that allows the Institute to better manage data, integrate informatics and optimise the underlying IT infrastructure. You will be responsible for reviewing the current and required technologies, processes and teams to deliver this new strategy. The strategy will take into account Good Research Practice as this can be challenging in innovative bioinformatics research due to rapidly changing software and complex global data sharing agreements and standards.

You will work collaboratively to more broadly harmonise our approach to data and informatics, especially with our strategic partner EMBL-EBI. This will include the research and development of compression software and establishing the right tools for our researchers so they can achieve scientific and translational impact.

Given the scale of our science, it is essential to achieve efficiencies in both storage and compute to ensure that IT costs do not become a rapidly increasing proportion of the cost of any large scale project, especially given the increasing rates of data generation.

Our projects are often undertaken in collaboration with academic, commercial or governmental partners. The rigour of the IT control environment, risk management, and customer service is becoming more important given the scale of these endeavours and the partners involved.

The increasing scale of our science and wider adoption of cloud technologies means there will be greater focus on IT Security and Data Governance, in particular having the ability to demonstrate the Institute’s compliance with regulations. Moreover, the advent of increasing research opportunities with health record data, working with NHS Digital and others, will need demonstrable security and data protection above that previously required in the Data Centre.

You will need to build, coach and manage a team that provides leadership to the Institute across technology, data management, informatics, analytics, platform architecture and product research and development.

Knowledge, skills and experience required:  E = Essential  D = Desirable

• Experience of leading technology enabled transformational organisation-wide change programmes with key focus on delivering value and innovation within the IT, informatics and data sectors. (E)

• A strategic thinker, with experience of developing integrated IT, data and informatics strategies linked to an organisational mission, in a complex and uncertain environment. (E)

• A proven record of turning strategy into reality through proactive stakeholder management and effective leadership of their teams. (E)

• Ensuring the customer needs are taken in to account when establishing and delivering quality products or services. (E)

• A credible leader with experience of developing staff and establishing high-performing teams. (E)

• Knowledge of IT and cloud vendor partner relationships management, and of reviewing the feasibility of outsourced operations. (E)
A strong technical background, with research and development experience, in aspects of IT, informatics and data. (E)

Broad knowledge of planning and operations, design and deployment and system life cycle management. (E)

Be a champion for equality, diversity and inclusion, and work collaboratively with a range of different stakeholders. (E)

Experienced influencer at the most senior level within a research-focused organisation. (D)

Knowledge of the regulatory environment for commercial and life science data. (D)

Ability to improve operational efficiency and service delivery effectiveness across all lines of business and technology platforms. (D)

Strong financial management and resource management skills. (D)

Knowledge of the challenges of working with large data flows (petabytes) in Genomes and BioData (D)

Experience with working in an organisation delivering scientific impact primarily delivered through the practice of open data and open source software(D)

Other information about the role not covered elsewhere

An outstanding communicator who has strong skills in building effective relations, negotiation and stakeholder management. The CIO is the figurehead for IT, data and informatics across the organisation; they will establish a common purpose across the teams and ensure they are working to the behavioural values of the organisation.
We have a dedicated Equality, Diversity and Inclusion (EDI) structure, along with:

- Equality in Science
- Equality, Diversity & Inclusion Forum
- Athena Swan project board

Our strategy is to foster an inclusive culture where everyone can thrive and diversity is celebrated. For more information about EDI at GRL see our Equality in Science Programme: https://www.sanger.ac.uk/about/equality-science