



Setting up Bioinformatics in a Data-Driven Biotech Start-Up

[ebook]

■ Biotech Start-Ups & Big Biological Data


Current technologies applied to biological and medical research are generating huge amount of data that can be used, for example, to develop new medical diagnostics and therapies.

Attracted by this potential, [many biotech entrepreneurs launch start-ups that are deeply rooted in the big biological data.](#)


However, to make the best of the opportunities that the data provides, it is worth to [first consider and make decision regarding the source of data, who will be analysing them and in what computational environment.](#)

In this ebook, we address these three aspects in visuals summarizing our blog posts on the respective topics.




What to consider when looking for a source of biological data for your biotech start-up? 



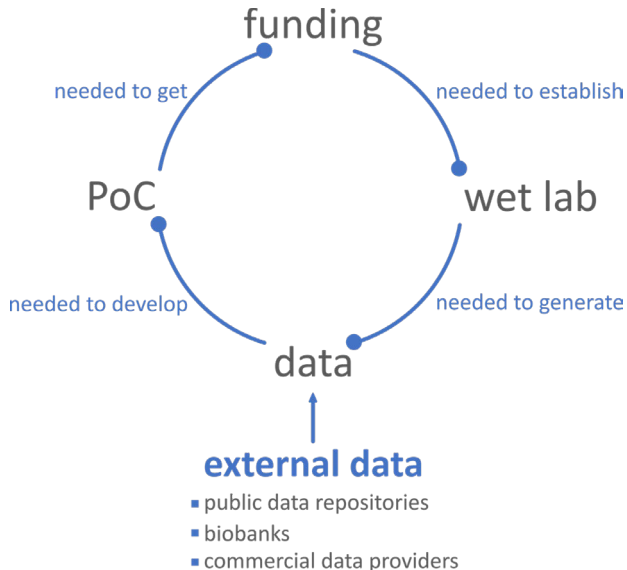
Who to team up with when in need of bioinformatics specialist for your biotech start-up? 



Why should data-driven biotech start-ups opt for a cloud-based computing environment? 

■ Biological Data Sources

Data from external resources enable biotech start-ups to develop PoC and expand their R&D without the need to generate their own input data.



What are the most important **benefits of the external data**?

Moving forward faster

Whether used to develop and test ideas or to explore potential new research direction, external data can greatly speed-up the R&D process.

Lowering costs

Public data repositories and biobanks offer data for free. Even if data needs to be purchased from a specialized company, the cost is still incomparably lower than establishing own infrastructure to generate the data.

Stronger case towards VCs, other funding bodies and stakeholders

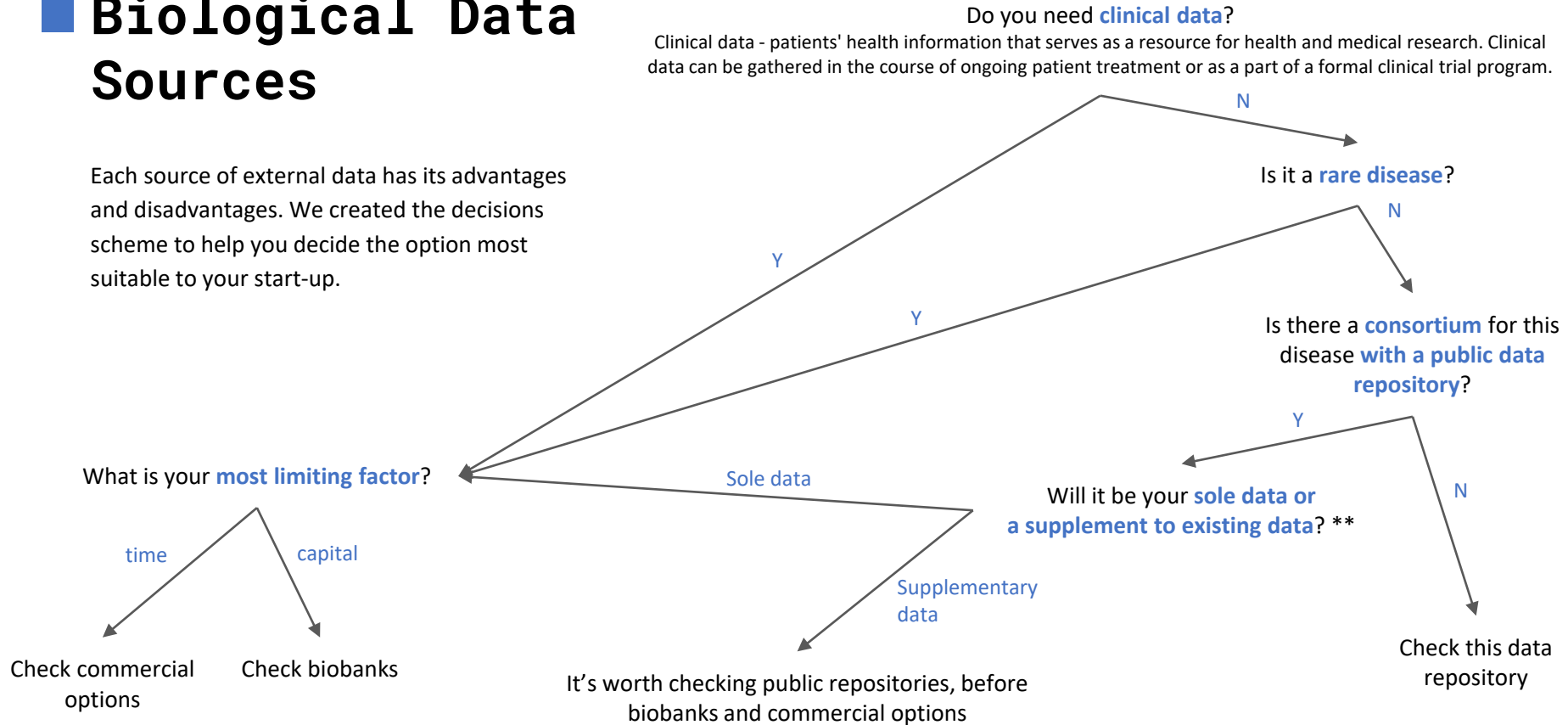
Factual and quantitative validation makes an R&D idea more appealing and therefore increases the probability of financial back up.

Pro tip: Scoring on trust

Ideas validated with data from renowned biobanks can increase trust and confidence in a start-up and attract more funds and the finest talents.

■ Biological Data Sources

Each source of external data has its advantages and disadvantages. We created the decisions scheme to help you decide the option most suitable to your start-up.



Do you need **clinical data**?
Clinical data - patients' health information that serves as a resource for health and medical research. Clinical data can be gathered in the course of ongoing patient treatment or as a part of a formal clinical trial program.

Is it a **rare disease**?

Is there a **consortium** for this disease **with a public data repository**?

Will it be your **sole data** or **a supplement to existing data**? **

What is your **most limiting factor**?

time

capital

Check commercial options

Check biobanks

It's worth checking public repositories, before biobanks and commercial options

Check this data repository

** For statistically relevant results you will have to have sufficient amount of data. If the data you need will be the sole input for analysis, you should focus on sources that usually have larger datasets (biobanks, commercial providers). If you need some extra data to supply to the existing data you have, also public repositories with their smaller datasets might already be enough.

■ Expert Support Sources

Bioinformatics expertise source that aligns well with a biotech start-up business and R&D strategy poses an important asset for the new venture development and growth. Yet, finding out which type of collaboration will work best is not easy.

What are the important aspects to consider?

Project

Size | Stage | Complexity | Development Potential | Importance

Is it a small-scale, new test project or a side project? Or is it an extensive, full-scale, multi-stage project that is at the core of the start-up's R&D? | Does the project have a narrow focus requiring highly specialized tools and techniques? Or is it a wide, explorative, complex project that will benefit from the extensive know-how of various biological aspects, and computational tools and techniques? | How likely is the project to develop into a large, multiplex undertaking? | How much does the start-up's competitive advantage depend on the project's results?

Time frame

What are the deadlines for the results & interpretation (e.g. next funding round, meeting with investors)? What should be the turnaround time?

Budget

How much is the start-up able to invest in the bioinformatics project(s)?

Current Bioinformatics Expertise

Is there a bioinformatics expert in the start-up team? Do they need extra support with specific topics, tools, or techniques? Or do they simply need extra hands to do the work?

Needed Support Timeline & Capacity

- Full-time, permanent
- Full-time, temporary; e.g. large complex project that has to be accomplished as soon as possible
- Part-time, permanent; e.g. continuous support of an in-house bioinformatician in side-projects
- Part-time, temporary; e.g. small, test side-project

Other

Legal support | Project management

■ Cloud vs on-premises computing environment

Though traditionally companies have been setting up their computational centers on-premises, more and more forward-thinking biotech start-ups chose to set up their data storage and analysis in the cloud.

What are the important benefits of the cloud computing environment?



Frog-leap the time- and resource-intensive wet lab phase when working on PoC



Set up and scale the computation environment faster, cheaper and more agile than with on-premises infrastructure



Quickly scale up or down, depending on the needs and circumstances - pay for what you use only



Off-the-shelf cloud formation templates dedicated to analyses of biological data assuring reproducibility



Some data security measures in place



Ability to utilize state-of-the-art AI and ML algorithms, sharpening the competitive edge



Progressive appearance more attractive to potential partners, investors and talents



Automation of storage and analyses of data coming from connected lab devices



Platform to provide tools for customers

■ Cloud vs on-premises

(part 1)

On-premises computing infrastructure

Cloud-based computing environment

Set up requirements

Hardware, servers
Software
Safe physical space & power
IT specialist

Cloud specialist

Usage costs

Hardware costs
Software costs
Maintenance costs
Power consumption (runs even if not used)

Hardware, software and maintenance costs distributed among all cloud users
Pay for what you use only

Scaling

Changes in hardware take time
Delays in scaling up lead to missed opportunities
Delays in scaling down incur unnecessary costs

Instant scale up and down enhances agility

■ Cloud vs on-premises

(part 2)

On-premises computing infrastructure

Cloud-based computing environment

Competitiveness - technological

Advanced algorithms executable only upon
ext(p)ensive expenditure on powerful enough infrastructure

Ability to run state-of-the-art algorithms

Competitiveness - talent

Appeal to potentially more forward-thinking talents

Security & Safety

Entire security has to be set up and continuously monitored
and updated by IT specialist
Requires specialized software
Requires external backup system to mitigate the risk of loss
due to physical accidents
Lower probability of data recovery in case of an accident

Some security assured by cloud providers (e.g encryption)
Other security measures, (e.g. data access rights) still need to
be configured by cloud specialist
Higher probability of data recovery in case of an accident

■ Useful links & resources

Biotech Market Size Report

[Grand View Research](#)

Public data repositories

[Nature's list of exemplary public repositories](#)

[PLOS-recommended repositories](#)

Biobanks

[UK Biobank](#)

Cloud solutions

[AWS for Health](#)

[Azure HPC for health and life sciences](#)

[Google Cloud Life Sciences](#)

■ Authors



Maarten Braspenning
CEO



Fleur Leenen
Team Leader US



Matthias Döring
Development Team Lead



Lucyna Włodarczyk-Gruber
MarCom Specialist

■ About BISC Global



Founded in 2017 BISC Global has grown to be a Top 10 Bioinformatics Consulting and Services Firm supporting the world's leading pharmaceutical and biotech companies. Through our offices in Belgium, Switzerland, The Netherlands, Germany, Boston (MA), and San Francisco (CA); BISC Global provides data analytics, custom tool/pipeline development, and cloud solutions delivered on-site or remotely.

By leveraging our global workforce of expert consultants, we support analysis across multiple domains, including machine learning (data mining, pattern recognition, image analysis, (un)supervised learning, neural networks, deep learning, etc.), bioinformatics (single-cell transcriptomics, metabolomics, metagenomics, epigenomics, etc.), and statistics (experimental design, mathematic modelling, signal-to-noise enhancement, and independent data cross-validation).

Visit us at biscglobal.com

Follow us   