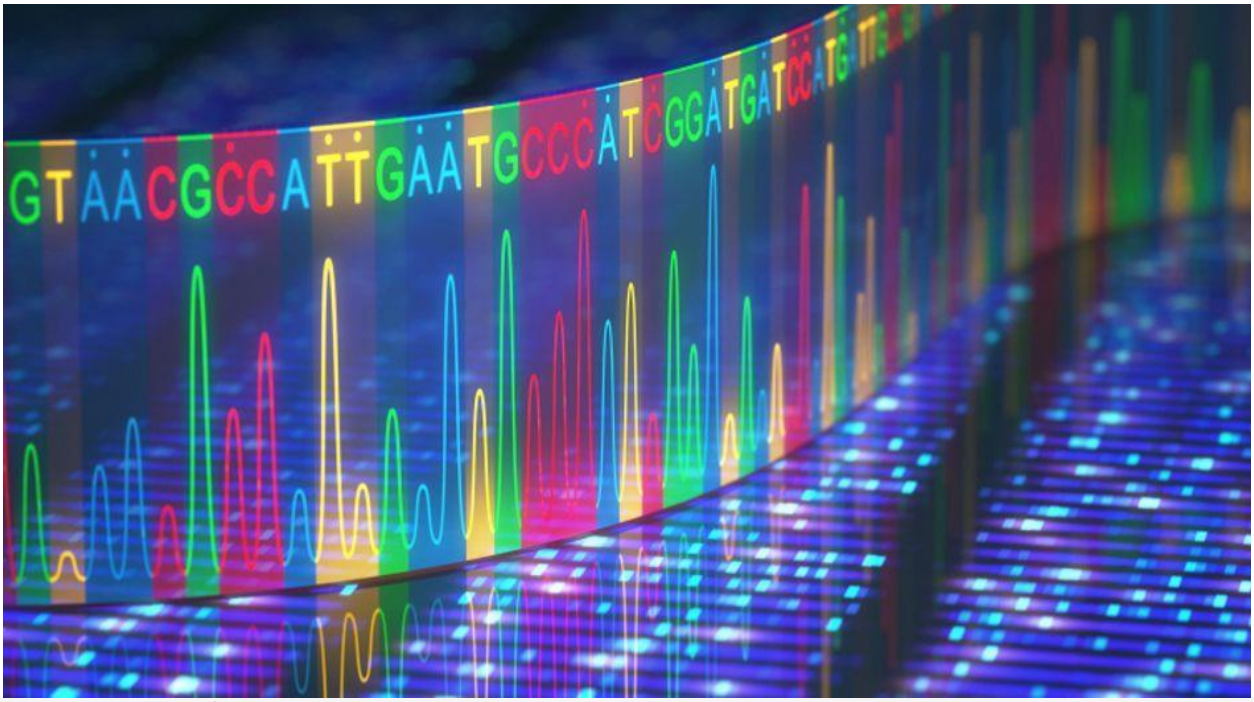
Patent strategies for protecting bioinformatics inventions

Patrick Darno and Peter Fasse

08 February 2023



Shutterstock/ktsdesign

Technology is revolutionising the study of genomes and DNA sequencing, promising advances in personalised medicine that will enable physicians to use a patient's own genetic makeup to pinpoint healthcare interventions with the most promise.

For example, if researchers know the genetic profile of a patient's cancerous tumour, they may be able to target drugs to bind to and kill only cancer cells that display a particular surface marker and not bind to other cells that do not express that surface marker, thereby producing fewer side effects for the patient.

However, as bioinformatics and computational genomics companies push the boundaries of technology in this space, it is important that they take steps to protect the intellectual property behind their hard-earned discoveries. Innovations in bioinformatics and computational genomics are subject to unique subject matter eligibility concerns that any effective patent strategy needs to consider.

With a firm grasp of the requirements of patentability for software and a cogent patent-drafting strategy, bioinformatics and computational genomics companies will be well-positioned to overcome the hurdles inherent in patenting their innovations.

Bioinformatics is a broad, interdisciplinary field that uses computer science to understand biology. In the most general sense, bioinformatics uses computer systems to analyse biological data. More specifically, it refers to computer systems that apply techniques related to data processing, mathematics, statistical modeling, physics, information engineering, or some combination thereof, to analyse and interpret biological data. For example, a bioinformatics system could encompass a system for obtaining genomic reads generated by a nucleic acid sequencer and then performing one or more computational operations on the obtained genomic reads.

Computational genomics is a subset of bioinformatics that uses computer science to draw inferences about a person or animal based on a genomic analysis. While it relies on the same technologies as bioinformatics, it is more focused on analysing and interpreting specific biological data inputs, such as whole genome sequencing data.

Strategically building a bioinformatics patent portfolio

While opportunities in personalised medicine abound and the industry is poised to grow, efficiently managing the massive amounts of data necessary for bioinformatics operations will be a major challenge for bioinformatics companies. As such, data compression is set to play a vital role in this field. Companies in the bioinformatics space must consider the role of data compression and decompression, such as by finding innovative ways to make genomic files smaller, ways to make compression and decompression of genomic files faster, and ways to make compression of genomic files lossless. Success in the bioinformatics industry requires companies to act now to ensure that they are not left behind.

To start, consider your future business needs:

- Does your company currently have a bioinformatics unit?
- If yes, are you taking steps to protect IP rights to your innovations in this space?
- If not, should your company start a bioinformatics unit?

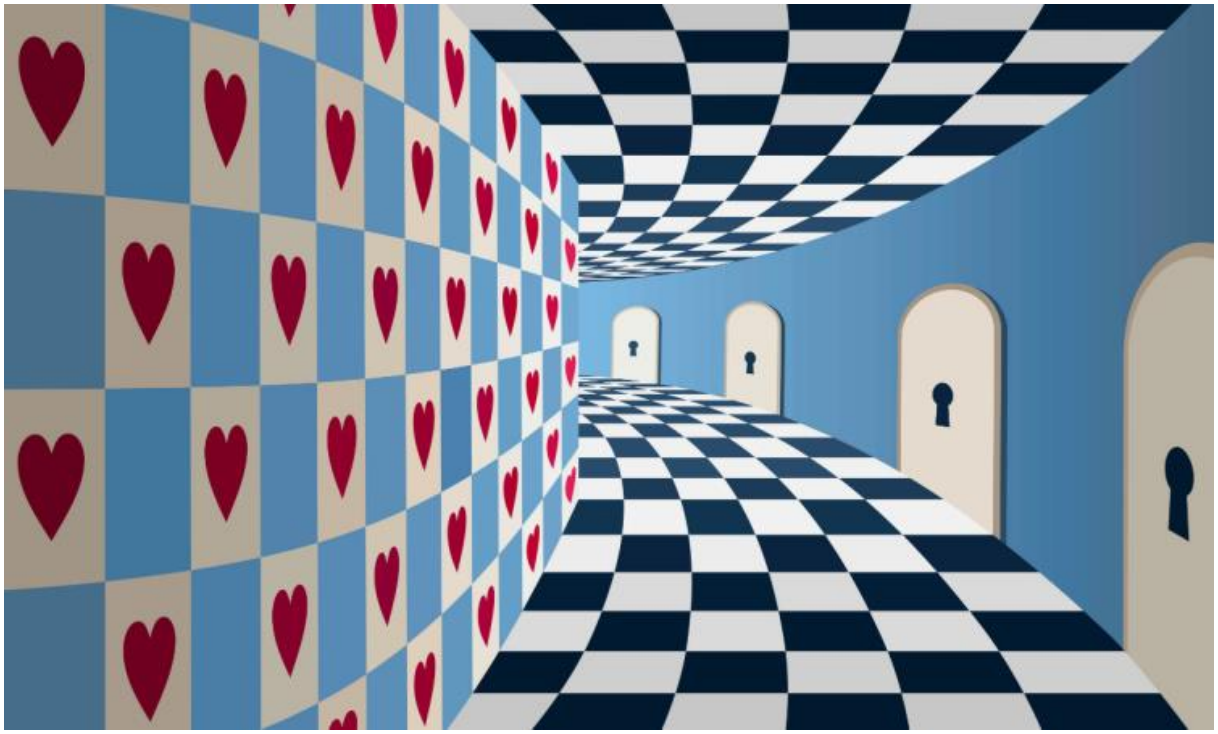
While thinking about today is important, thinking about tomorrow is just as important – where is the industry headed in the next five, 10, 15, or 20 years? It is always beneficial to evaluate and protect technology that is currently in development or that can be deployed today, but do not sacrifice the future by focusing only on today. It is important to perform creative exercises, such as identifying innovations that could be achieved assuming an unlimited team and unlimited budget, and then filing “prophetic” patent applications covering those inventions.

You can develop a coherent bioinformatics patent strategy using a two-tiered approach:

1. **First tier – look inward at your company’s innovations:** identify streams of products and services from which your company currently derives revenue and those from

which your company plans to derive revenue in the future. Then prepare and file at least one patent family for new and useful “blocking features” (features that protect the product stream that cannot be easily avoided by your competitors) for each identified stream of products or services. Also consider filing patent applications on improvements to current products and services, as well as new products and services you may launch in the future. Finally, monitor public disclosures by employees and file patent applications on blocking features before any public disclosures (to the extent practicable) and employ nondisclosure agreements as appropriate.

2. **Second tier – look outward at your competitors’ innovations**: gather competitor intelligence related to your competitors’ products and services by determining their key features, any improvements that could be made to them, and their potential technological evolution. Then file patent applications on blocking features related to your competitors’ products or services in any of the above categories to build an economic moat. By determining how you can improve on a competitor’s offering and filing a patent application for the improvement, you can potentially cut off a competitor’s revenue stream.



Navigating *Alice* landmines

Subject matter eligibility is often the most difficult hurdle to overcome for bioinformatics and computational genomics innovations, as they are likely to receive "*Alice* rejections" under 35 U.S.C. § 101 by the US Patent and Trademark Office (USPTO). Such rejections allege that a patent claim is directed to patent ineligible subject matter, such as abstract ideas, laws of nature, and natural phenomena as articulated by the US Supreme Court in the 2014 case ***Alice Corp v CLS Bank International***. In its decision, the Supreme Court developed a two-part test to determine whether an invention is eligible for patent protection:

1. Determine whether the claims at issue are directed to a patent ineligible concept, and
2. If so, determine whether the elements of the claims, considered both individually and in combination, are sufficient

to ensure that the patent in practice amounts to significantly more than a patent on the ineligible concept itself.

Are bioinformatics inventions eligible for patent protection?

For software innovations, many examiners are likely to issue *Alice* rejections alleging that bioinformatics claims are directed to abstract ideas, although bioinformatics applicants also frequently receive law-of-nature rejections if an innovation involves genetic sequence information or methods of diagnosis.

The US Court of Appeals for the Federal Circuit in 2021 considered the patent eligibility of a bioinformatics invention in ***In re Board of Trustees of Leland Stanford Junior University***. While the court found the claims in that case to be patent ineligible, a close read of the opinion reveals insights that provide helpful practice tips.

In *Stanford*, the patent application (US application No 13/486,982) was directed to computerised statistical methods for determining haplotype phase. Claim 1 of the '982 application states:
A computerised method for inferring haplotype phase in a collection of unrelated individuals, comprising:

...

building a data structure describing a Hidden Markov Model, where the data structure contains

a set of imputed haplotype phases...

a set of parameters comprising local recombination rates and mutation rates

wherein any change to the set of imputed haplotype phases contained within

the data structure automatically results in re-computation of the set of parameters

...

storing the at least one final predicted haplotype phase for the individual on a memory

of a computer system

The patent examiner rejected the claim as being directed to an abstract idea, which the Patent Trial and Appeal Board (PTAB) affirmed. The board explained that Stanford failed to identify any specific disclosures in the specification that indicate the method of claim 1 improves computer functionality. The Board distinguished ***McRo, Inc v Bandai Namco Games America Inc***, because the process used in that case combined specific rules that rendered information into a specific format, which was then used and applied to create desired results (i.e., a sequence of synchronised, animated characters). The PTAB acknowledged that claim 1 of the '982 application may be useful in medical or population genetics studies, but was devoid of a specific step that the claimed calculations were "integrated" into a practical application.

On appeal, the Federal Circuit affirmed the PTAB's decision, holding that the claims were directed to an abstract idea. At oral argument, Stanford asserted that the claim steps resulted in a more accurate haplotype prediction. This argument was based on the principle that an abstract claim can still be patent eligible if it provides an improvement to the functioning of a computer or to a particular method. The Federal Circuit disagreed, pointing out that the '982 application clearly stated

that the mathematical steps were “conventional” and “well-understood” and that claim 1 recited no application, concrete or otherwise, beyond storing data. The court explained that the **alleged** increase in computational accuracy did not qualify as an improvement to a technological process. The court distinguished an alleged improvement from an actual improvement supported by evidence in the specification, as filed, which implies that an **alleged** improvement based solely on after-the-fact attorney argument is not enough to overcome a § 101 rejection.

Practice tips for drafting bioinformatics patent applications

Stanford holds a few key lessons for practitioners.

1. Drafting a patent application to survive patent eligibility challenges is a multi-faceted approach; a claim must be drafted to capture a practical application of a process resulting in a technological improvement *and* the specification must be drafted to explain this technological improvement.
2. Technological improvements cannot merely be supported by attorney argument. Both the PTAB and the Federal Circuit rejected attorney arguments of alleged increased accuracy while also highlighting the lack of support in the application for such increased accuracy (or, more bluntly, pointing out that the specification said the mathematics was conventional).
3. Execution of a computer process on a data structure should be claimed and described at an appropriate level of detail – i.e., a level of detail that recites how the process executed on the data structure creates a formatted data result that improves computer functionality.
4. Claims must bring a process executed to achieve a technological improvement to a conclusion that achieves a practical application of the claimed process; they should not merely claim “generating and storing data”.

Giles Rich, a former Chief Judge of the Federal Circuit, once famously stated that “the name of the game is the claim”. However, in contrast to this bedrock principle of patent law, the Federal Circuit in *Stanford* was more critical of the specification of the ‘982 application than the claims themselves. Accordingly, when drafting patent applications, it is important that a patent practitioner follow the lessons learned from *Stanford* and draft an appropriately detailed description that explains a process having a practical application and a technological improvement. To this end, a patent practitioner should not leave the practical application or technological improvement undefined, as doing so may invite future accusations that the technological improvement is merely alleged. When drafting, a patent practitioner can consider including a paragraph or two in the summary to create a concise description of the invention that captures the practical applications and technological improvements. This will ensure that you will always know where to find this helpful disclosure and may even provide support for arguments against future *Alice* rejections.

It can also be helpful to include experimental results like the results of comparative testing, when practicable, to provide evidentiary support for a technological improvement. For example, if the invention is related to the compression of genomic quality scores, provide data comparing compressed file sizes and compression speed of the claimed methods to test data of conventional compression algorithms performed on the same test data. This evidence can help to show how your new method is better than existing methods.

Example of successful claim-drafting

While *Stanford* demonstrated the importance of the specification, these elements of a practical application and technological improvement should also be captured in the claims. Effective claim drafting to overcome *Alice* rejections requires claiming a practical

application and a technological improvement, as well as reciting a final claim step that includes the practical application and the technological improvement. For example, a useful hypothetical claim for an invention directed to compression of nucleic acid sequence data using multi-stage encoding might read as follows:

Hypothetical Claim 1

1. A method for compressing nucleic acid sequence data, the method

comprising:

obtaining, by one or more computers, nucleic acid sequence data;

determining, by one or more computers, whether the read sequence

data satisfies a first criteria;

based on a determination that the read sequence does not satisfy the

first criteria, generating, by one or more computers, a first encoded data set to

reduce size of the nucleic acid sequence data; and

using, by one or more computers, a second encoding process to encode

the first encoded data set, thereby compressing the obtained nucleic acid

sequence data

Claim 1 above uses all three strategies for effective claim drafting to overcome *Alice* rejections: A practical application (i.e. a method of

compressing nucleic acid sequence data), a first technical improvement (i.e. reducing the size of input data to a second encoding process), and a final step (i.e. completing the process by compressing the input data of reduced size).

Drawings can also be used to enhance the detailed description. Often, patent applicants in the computer science space use block diagrams of system elements, but those system elements alone may not disclose much about the invention. Some patent applications include many pages of drawings, but each drawing shows only one component of the overall invention, depriving the patent examiner of a bird's eye view of the entire invention. A more effective strategy is to describe an entire invention in a single figure, as this may make it more difficult to argue that the invention is abstract. When paired with specification support that describes the technological improvement and claims that describe a practical application, technological improvement, and a final step, this can create a powerful package to argue against deficiencies commonly cited in § 101 rejections.

While the comments above apply generally to inventions in the broad fields of bioinformatics and computational genomics, we next review considerations to protect specific types of technologies in these fields.

Strategies along the bioinformatics pipeline

A bioinformatics pipeline can include a device, software engine, or a combination thereof that:

1. Performs one or more operations to obtain or generate biological data (e.g., perform sequencing operations on plant or animal DNA to generate genomic reads);
2. Performs one or more secondary analysis operations on biological data to interpret the biological data (e.g., map genomic reads to a reference sequence, align genomic reads to a reference sequence, determine variants between the genomic reads and the reference sequence); or
3. Performs one or more tertiary analysis operations on output from the secondary analysis engine (e.g., generates targeted diagnosis, treatment, or a combination thereof based on obtained variants)

Patent strategies to protect innovations relating to bioinformatics pipelines can include a two-part analysis. Two key questions to ask when framing a bioinformatics patent strategy are where in the pipeline the invention lies and who is likely to perform that invention. Such initial strategic planning can help avoid common pitfalls that can arise in patent enforcement such as joint infringement defences raised by infringers by ensuring that the claims focus on the relevant stage of the bioinformatics pipeline where the innovation resides and the target entity practices.

First, determine the stage of the bioinformatics pipeline that is the main part of the invention (i.e., sequencing, secondary analysis, or tertiary analysis). Also consider whether the invention includes any pre- or post-operation processing. Pre-operation processing can encompass, for example, formatting reads output by a sequencer or decompressing reads received by a network prior to performing subsequent processing operations on the formatted or decompressed reads. Post-operations processing can include compressing reads generated by a secondary analysis engine, evaluating a diagnosis generated by a tertiary analysis engine, or administering a treatment generated by a tertiary analysis engine.

Second, identify the entities that are likely to perform the invention wherever it falls within the pipeline. These could include:

- A lab operating a sequencer that performs sequencing operations (sequencing stage)
- A data analytics company that performs secondary analysis operations in a cloud server (secondary analysis stage)
- A pharmaceutical company that runs tertiary analysis software to generate a diagnosis or treatment (tertiary analysis stage)
- A doctor who administers a treatment (post-operations processing)

A good practice is also to consider how the “where” and the “who” may change over time as technology evolves. Such strategies can function to future-proof a patent portfolio by, for example, drafting patent claims that are unlikely to result in divided infringement. In addition to avoiding divided infringement, having patent claims that read on a single entity performing all operations of the claim(s) makes the infringement read easier.

We will now take a closer look at each stage of the bioinformatics pipeline and provide some practice tips for drafting claims on innovations at each respective stage.

Phase 1: Sequencing inventions

Sequencing inventions typically relate to operations performed by a nucleic acid sequencer, such as sample preparation, nucleic acid sequencing, quality scoring of base cells, and read data formatting. Data generated by a sequencing device is often used in bioinformatics inventions. When drafting claims for sequencing devices, ask whether sequencing is a part of the invention. Avoid independent claims that include an active sequencing operation if the novel or non-obvious subject matter of the invention is not found at the sequencing stage. Similarly, avoid drafting an independent claim to include an active sequencing operation if the target entity you wish to cover with your claim is not likely to perform a sequencing operation itself.

Phase 2: Secondary analysis inventions

Secondary analysis inventions can relate to operations performed on previously-generated biological data, such as mapping of genomic reads to a reference sequence, alignment of reads to a reference sequence, and determination of variants between aligned reads and a reference sequence. When drafting patent claims at the secondary analysis stage, consider the relationship between the secondary analysis engine to other stages of the bioinformatics pipeline.

Answer these questions:

- Would an entity that performs the secondary analysis of generated biological data also perform sequencing operations or certain pre- or post-processing operations?
- Where is the secondary analysis engine located with respect to the nucleic acid sequencer, databases of genomic data, or other input sources?
- Where will the secondary analysis engine send the outputs it generates?
- How will the secondary analysis engine send those outputs?

Answers to these questions may help to focus the scope of patent claims on the point of innovation of the bioinformatics pipeline, while keeping an eye on your competitors.

Phase 3: Tertiary analysis inventions

Tertiary analysis inventions can relate to operations performed on the output of a secondary analysis engine, such as a determination of a diagnosis, a determination of a treatment, or both. When drafting claims to cover the tertiary analysis stage, consider the relationship between the tertiary analysis engine to other stages of the bioinformatics pipeline by, for example, considering the answers to strategic questions regarding this relationship

Such questions include:

- Would an entity that performs tertiary analysis of the genomic reads also perform sequencing or secondary analysis?
- Where is the tertiary analysis engine located with respect to other pipeline devices and engines?
- Will the tertiary analysis engine perform pre-processing operations?

- Where will the tertiary analysis engine transmit generated outputs?
- Will a tertiary analysis engine perform post-processing operations on output before transmitting?
- Would the entity that performs tertiary analysis to determine a treatment administer the treatment?

In addition, try to avoid drafting a claim that requires a treatment step (unless the target entity actually administers the treatment), but note that adding a treatment step can be helpful to overcome 35 U.S.C. § 101 rejections.

Consideration of these factors enables a patent practitioner to balance infringement and validity issues strategically when drafting and prosecuting patent claims.

Computational genomics inventions may fall within the tertiary analysis stage of the bioinformatics pipeline. As such, consider the same relationships to the other phases of the bioinformatics pipeline as those for tertiary analysis inventions. Also consider the unique challenges related to processing the large data set of an entire genome, such as storing of a genome, access of a genome, processing of a genome, and parallel processing to increase throughput.

Patenting bioinformatics outside the US

When building a global portfolio, it is important to consider how patent offices outside the US handle this type of subject matter, as claims drafted to comply only with US law may fail patent eligibility tests abroad.

In general, innovations that **are not** patent eligible in most countries outside the US include:

- Computer programs “as such” or “per se”
- Methods that can be performed in the human mind
- Known methods executed within a general-purpose computer
- Known analytic technique (e.g., machine learning) applied to data organised in a known manner
- Methods of surgery, treatment, or diagnosis of humans

In general, innovations that **are** patent eligible in most countries outside the US include:

- Inventions in an improvement in a technical field outside of a computer
- Computer-readable recording mediums containing computer programs that cause computers to execute certain new steps
- Diagnostic methods using isolated samples and that do not involve “clinical determination”
- Devices, systems, and compositions used for diagnostic methods

The good news for practitioners is that claims drafted for US or European Patent Office (EPO) standards are also generally useful in other countries.