

## Case Study

**MIROCULUS**  
SCIENCE SIMPLIFIED**INDUSTRY**

Life Sciences

**USE CASE**

Knowledge Graph

**GOAL**

Find the right microRNA biomarker to support a blood test for gastric cancer.

**CHALLENGE**

Assimilating findings from voluminous medical research into microRNAs.

**SOLUTION**

Create a knowledge graph of the latest microRNA research.

**RESULTS**

- Accelerated insights into medical research
- Faster identification of microRNA indicative of stomach cancer

**Miroculus****Machine Learning and Graph Technology Accelerate Medical Research**

*Miroculus created a platform that detects microRNAs in blood. To scour the ever-expanding corpus of relevant medical literature, Miroculus designed a machine learning pipeline that captures research on microRNAs and genes to diseases in a knowledge graph stored in Neo4j.*

**The Company**

[Miroculus](#), a San Francisco-based molecular diagnostics company, developed a digital microfluidics platform that enables scientists and healthcare workers to detect diseases using microRNAs. The diagnostic platform handles various kinds of testing and is flexible and simple to use.

**The Challenge**

For most types of cancer, early diagnosis improves patient survival rates. The problem is that many tests used to diagnose cancer are invasive and require special equipment.

For example, stomach cancer is still diagnosed with an endoscopy, technology from 1965. Further, endoscopies are not an effective screening tool: only 2% of patients screened have stomach cancer. A simpler diagnostic method could eliminate unnecessary procedures and increase detection.

Miroculus saw the promise of microRNAs for cancer detection, but microRNAs were thought to be locked in cells. In 2008, one of the company's advisors discovered circulating microRNA. Whenever there's a problem in the cells at the tissue level, they break apart and release microRNAs into the bloodstream.

Today, detecting microRNAs still requires highly skilled scientists, expensive reagents and machinery and very complex protocols. Miroculus saw the potential to change the face of microRNA detection.

Conducting a study to find a microRNA biomarker for stomach cancer meant keeping up with an explosion in related medical research. Typical research methods involved searching for articles, selecting a relevant article, attempting to retrieve it and assimilating it.

With the increase in microRNA research, absorbing all the relevant information would take several lifetimes. Miroculus needed a way to accelerate that process and connect scientists directly to pertinent research.

**The Solution**

Miroculus wanted to find a microRNA biomarker for stomach cancer. It is a compelling research area – of 1 million people diagnosed with stomach cancer, 80% survive less than 18 months.

## Case Study



“In order to make sense of all the newly available microRNA information, we stored this high volume of data in a searchable graph database.”

–Antonio Molins  
VP of Data Science, Miroculus

Their work required keeping up with the latest publications connecting genes, diseases and microRNAs. “In order to make sense of all the newly available microRNA information, we stored this high volume of data in a searchable graph database,” said Antonio Molins, VP of Data Science at Miroculus.

The Miroculus team gathered more than a billion articles in Hadoop. Next they used natural language processing (NLP) to extract specific sentences with keywords for gene, disease and microRNA. Inferring the relationship between keywords required yet another step. The team developed an [unsupervised machine learning model](#) to classify relationships, which are then stored in Neo4j.

“We think it’s good to use the right tool for the right problem,” said Molins. “Graph databases are the right tool if you are focusing on relationships.”

The team created an interactive visualization searchable by microRNA, gene and disease. The user’s search criteria becomes the central node of the visualization, with surrounding nodes connecting specific microRNAs. Selecting a particular node pulls up papers that relate them, with the specific sentence cited and a link to the publication.

With the latest medical research in hand, Miroculus designed a study following FDA guidelines. The study – conducted in collaboration with the NIH, the National Cancer Institute and experts in Chile – included 650 people eligible for an endoscopy to diagnose stomach cancer.

## The Results

The Miroculus team discovered relevant links between microRNAs and gastric cancer faster with medical research connected in Neo4j. “It helped us a lot trying to find specific microRNAs that were related to gastric cancer,” said Molins.

The Miroculus team identified microRNA that successfully screens for gastric cancer using a blood sample. Their platform offers a minimally invasive approach, saving endoscopies for those who need them most. “The test predicts well in early as well as later stages. This is important; we want to find people with early stages of cancer, where the prognosis is much better,” said Molins.

Miroculus made its visualization of the relationships between genes, diseases and microRNAs [publicly available](#). The color-coded visualization interactively queries Neo4j. “People find a lot of value in first going to the findings that you want to look at and then pulling the paper, instead of the other way around. It seems to be more efficient,” said Molins.

Selecting an element retrieves the specific sentence in the article that includes the search elements, along with a link to the full article, connecting researchers with relevant studies.

“Machine learning and graph technology like Neo4j have provided opportunities in the medical field to accelerate scientific research,” said Jorge Soto, co-founder and CTO, Miroculus.

Neo4j is the leader in graph database technology. As the world’s most widely deployed graph database, we help global brands – including [Comcast](#), [NASA](#), [UBS](#), and [Volvo Cars](#) – to reveal and predict how people, processes and systems are interrelated.

Using this relationships-first approach, applications built with Neo4j tackle connected data challenges such as [analytics and artificial intelligence](#), [fraud detection](#), [real-time recommendations](#), and [knowledge graphs](#). Find out more at [neo4j.com](#).

Questions about Neo4j?

Contact us across the globe:  
[info@neo4j.com](mailto:info@neo4j.com)  
[neo4j.com/contact-us](https://neo4j.com/contact-us)