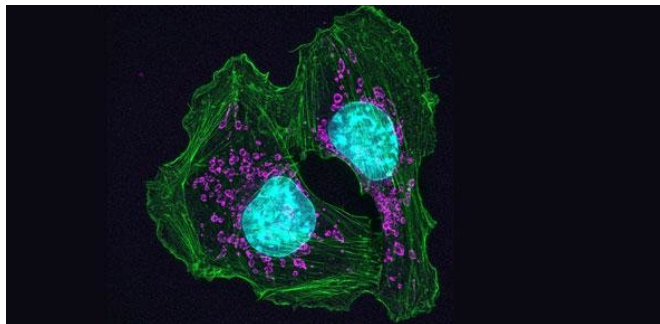


AI system may accelerate search for cancer discoveries

27 November 2018



Skin cancer cells from a mouse show how cells attach at contact points. Credit: [NIH Image Gallery](#)

Searching through the mountains of published cancer research could be made easier for scientists, thanks to a new AI system.

The system, called [LION LBD](#) and developed by computer scientists and cancer [researchers](#) at the University of Cambridge, has been designed to assist scientists in the search for cancer-related discoveries. It is the first literature-based [discovery](#) system aimed at supporting cancer research. The results are reported in the journal *Bioinformatics*.

Global cancer research attracts massive amounts of funding worldwide, and the [scientific literature](#) is now so huge that researchers are struggling to keep up with it: critical hypothesis-generating evidence is now often discovered long after it was published.

Cancer is a complex class of diseases that are not completely understood and are the second-leading cause of death worldwide. Cancer development involves changes in numerous chemical and biochemical molecules, reactions and pathways, and cancer research is being conducted across a wide variety of scientific fields, which have variability in the way that they describe similar

concepts.

"As a cancer researcher, even if you knew what you were looking for, there are literally thousands of papers appearing every day," said Professor Anna Korhonen, Co-Director of Cambridge's Language Technology Lab who led the development of LION LBD in collaboration with Dr Masashi Narita at Cancer Research UK Cambridge Institute and Professor Ulla Stenius at Karolinska Institutet in Sweden. "LION LBD uses AI to help scientists keep up-to-date with published discoveries in their field, but could also help them make new discoveries by combining what is already known in the literature by making connections between sources that may appear to be unrelated."

The 'LBD' in LION LBD stands for Literature-Based Discovery, a [concept](#) developed in the 1980s which seeks to make new discoveries by combing pieces of information from disconnected sources. The key idea behind the original version of LBD is that concepts that are never explicitly linked in the literature may be indirectly linked through intermediate concepts.

The design of the LION LBD system allows real-time search to discover indirect associations between entities in a database of tens of millions of publications while preserving the ability of users to explore each mention in its original context.

"For example, you may know that a cancer drug affects the behaviour of a certain pathway, but with LION LBD, you may find that a drug developed for a totally different disease affects the same pathway," said Korhonen.

LION LBD is the first system developed specifically for the needs of [cancer research](#). It has a particular focus on the molecular biology of cancer and uses state-of-the-art machine learning and natural language processing techniques, in order to detect references to the hallmarks of cancer in the text.

Evaluations of the system have demonstrated its ability to identify undiscovered links and to rank relevant concepts highly among potential connections.

The system is built using [open data](#), open source and open standards, and is available as an interactive web-based interface or a programmable API.

The researchers are currently working on extending the scope of LION-LBD to include further concepts and relations. They are also working closely with [cancer](#) researchers to help and improve the technology for end users.

More information: Sampo Pyysalo et al. LION LBD: a Literature-Based Discovery System for Cancer Biology, *Bioinformatics* (2018). [DOI: 10.1093/bioinformatics/bty845](#)

Provided by University of Cambridge

APA citation: AI system may accelerate search for cancer discoveries (2018, November 27) retrieved 1 October 2022 from <https://medicalxpress.com/news/2018-11-ai-cancer-discoveries.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.