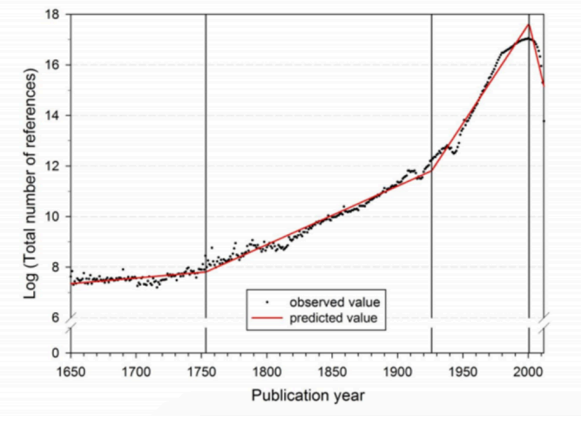
Why Your Keyword-Based Technical Literature Review Is Reducing R&D Productivity by 17X

LEGIT

What's the problem with keywords?

R&D Magazine's 2018 survey found that the three biggest issues facing R&D managers were increasing costs, insufficient project time, and an insufficient R&D budget. With those concerns, anything that can increase R&D efficiency should be adopted as a matter of urgency. Yet the current methods for one of the key tasks in the R&D workflow reduce productivity by as much as 17 times over.

To identify new products, continually improve core technologies and products, and stay competitive, it is critical for R&D teams to keep on top of the constant flow of technical and scientific literature.



A study published in *Nature* (figure on the left) found that global scientific literature doubles every nine years - if you add that to the millions of other documents that R&D teams have to monitor to stay ahead, that adds up to a big haystack!

When trying to understand if an idea is likely to be original and potentially valuable, a current keyword-based review of what's already out has three serious drawbacks

1 Misses conceptually similar ideas with no keywords in common

When using keywords, the R&D team member has to select keywords which best match what they are looking to explore. Then the database being searched will usually only return documents which match the keywords which have been selected. This is a problem where (as is commonly the case) different engineers, researchers, and scientists use different terminology for the same invention or technology.

A case in point is the laser beam:

The first documents outlining this technology and its applications do not feature the word laser at all! The technology was referred to as a maser, or optical maser. Using keyword search here would fail to allow an R&D team to accurately scope out the landscape - indeed they'd miss the key innovation!

2 No personalization

Imagine you are on an R&D team working on medical imaging and you're using free or commercially available tools to scout out the technology landscape. You wonder if someone has worked on portable devices that can perform Computerized Axial Tomography – a CAT scan – as you think there's a market for them.

If you type in "portable CAT device" into Google patents right now, 50% of the first page of results are about cats (the animal) and the rest are totally irrelevant to the area you're exploring.

The current systems that match queries to the database do not know and crucially, cannot learn, that one is interested in medical imaging. Thus they cannot determine that if a user types in CAT, then they are more likely to be using a technical term of art than looking for feline related innovations.

③ Time to review

Current methods will take a keyword-based query from a user and return to them the entire document to review. These documents can be 30-40 pages long if they are a long-form academic article or patent. It's like asking to find a needle in a haystack and being given another smaller haystack - better but not good enough.

R&D teams will need to then review each document to determine whether it's relevant to the project they are working on. Speed-reviewing a 3,000 word piece of technical text to determine its relevance to an R&D project takes on average 4.2 minutes, a more in-depth review could take hours.

How does Legit help?

Legit matches on concepts not keywords

Legit uses AI to process **1.5 million academic articles** and **7 million patents** analyzing their concepts and context so that we can identify when two ideas are similar even if they don't share keywords or terminology in common.

Legit learns with usage

Legit uses machine learning algorithms combined with user profiles to personalize the experience for each user. The system learns what technology you are interested in and adjusts the search space accordingly, meaning that R&D teams can cover more ground and find what's new faster.

Legit returns the exact part of the document

Legit doesn't return an entire document for an R&D team to review - it drills down and returns the exact part the document where the match is strongest. We've found this means that R&D teams can now identify whether the results are relevant in an average of **14.9 seconds - 17 times faster than current methods.**

This **94% time reduction** translates to an average of **176 saved R&D hours per team member** annually.
