idw - Informationsdienst Wissenschaft
Nachrichten, Termine, Experten

Pressemitteilung

ZB MED - Informationszentrum Lebenswissenschaften
Ulrike Ostrzinski

05.01.2023
http://idw-online.de/de/news807318

Forschungsprojekte, Kooperationen
Biologie, Ernährung / Gesundheit / Pflege, Informationstechnik, Medizin, Tier / Land / Forst
überregional

Preprints als Informationsquelle besser nutzbar machen - gemeinsames DFG-Projekt von TH Köln und ZB MED

Im Projekt PIXLS – Preprint Information eXtraction for Life Sciences werden die TH Köln und ZB MED – Informationszentrum Lebenswissenschaften in den kommenden drei Jahren eine Anwendung entwickeln, die Preprint Server automatisiert erschließt. Damit kann die Forschungscommunity die aktuellen Informationen, die auf Preprint Servern veröffentlicht wurden – und daher in klassischen Nachweis- und Suchsystemen kaum auftauchen – besser nutzen. Die Deutsche Forschungsgemeinschaft (DFG) fördert das Projekt im Rahmenprogramm e-Research Technologien.

Preprints stellen eine relativ neue Art dar, wissenschaftliche Ergebnisse noch vor dem Peer Review und der Veröffentlichung in einem Fachjournal für die Forschung bereitzustellen. So können aktuelle Erkenntnisse schnellstmöglich verbreitet und nachgenutzt werden. Mittlerweile gibt es verschiedene Server, die Preprints für unterschiedliche Forschungscommunitys bereitstellen. Sie unterscheiden sich technisch oder inhaltlich. Preprints.org ist beispielsweise multidisziplinär ausgerichtet, während bioRxiv ausschließlich Preprints aus den Lebenswissenschaften enthält.

Die Arbeitsgruppen aus der TH Köln – unter Leitung von Prof. Dr. Philipp Schaer – und von ZB MED – unter Leitung von Prof. Dr. Konrad Förstner – werden in PIXLS ein System entwickeln, das die bisher noch vernachlässigten Informationsquellen auf den Preprint Servern systematisch erschließt und durch Mehrwertdienste besser zugänglich macht. So können sowohl die Volltexte als auch die Metadaten einfacher nachgenutzt werden. Das PIXLS-Team setzt dabei auf eine sogenannte Information Extraction Pipeline, eine Anwendung, die aus den unstrukturierten Daten der Preprints – also Fließtexte oder Zahlen – strukturierte Informationen extrahiert. Solche strukturierten Informationen können von Datenbanken besser verarbeitet werden. Sie können dadurch auf vielfältige Weise nachgenutzt werden und sind darüber hinaus einfacher auffindbar.

Die extrahierten Daten werden im ZB MED Knowledge Environment – eine von ZB MED bereits entwickelte Datenbank – zusammengeführt und vereinheitlicht. Diese Datenbasis macht es dann möglich, Mehrwertdienste zu entwickeln und für die Forschung bereit zu stellen. Das können zum Beispiel Linked-Open-Data-Schnittstellen oder innovative Reputations- und Trendindikatoren sein. Die Daten werden auch über LIVIVO, den Discovery Service von ZB MED, verfügbar sein. Im Sinne von Open Science werden sowohl die Daten als auch die Technologie der Bibliotheks- und Wissenschaftscommunity zur Nachnutzung zur Verfügung gestellt.

Preprints stellen eine relativ neue Art dar, wissenschaftliche Ergebnisse noch vor dem Peer Review und der Veröffentlichung in einem Fachjournal für die Forschung bereitzustellen.
Marcus Gloger
ZB MED