TaxonConcept - TaxonRank a synonymy ranking algorithm for earth science data networks

Eine Vielzahl von Arten sind ausreichend beschrieben und untersucht worden und sie sind heute ein extrem wichtiges Werkzeug für z.B. die Stratigraphie oder Paläoklimatologie. Diese taxonomischen Informationen sind allerdings nur dann nutzbar, wenn sie effektiv organisiert und verfügbar sind um die gerade benötigte Information zu finden. Die klassische Methode hierzu in der Paläontologie ist die Veröffentlichung von taxonomischen Primärdaten in Zeitschriften mit Hilfe von sorgfältig vorbereiteten Synonymielisten Die meisten Fachzeitschriften fordern die Verwendung der sogenannten ‚Offenen Nomenklatur', die es erlaubt auch mit unklaren Klassifikationen zu arbeiten und es dem Autor erlaubt die Identifizierung von Individuen durch andere Autoren zu kommentieren.

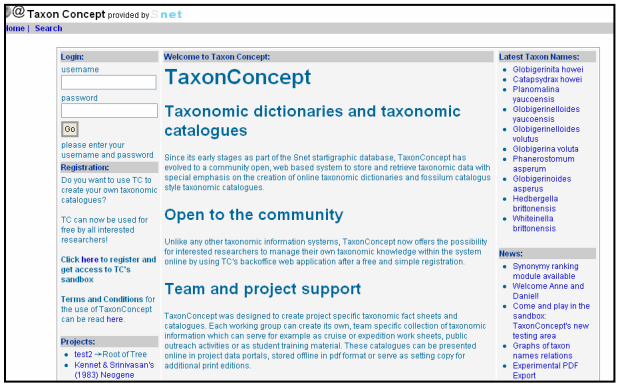
Diese Synonymielisten spiegeln eher die taxonomische Meinung dieses Autoren als dass sie objektive Synonymien beinhalt Die Nutzung dieser Listen ist deshalb zwar hilfreich für die Anwendung etwa in der Biostratigraphie. Für den, in dieser Organismengruppe nicht spezialisierten, Geowissenschaftler ist es allerdings schwierig der verfügbaren taxonomischen Literatur zu folgen um auch die Arten zu finden für die ebenfalls Daten existieren und die womöglich auch als Synonyme verwendet wurden.

Eine Synonymieliste besteht im Grunde aus einer Liste aus Zitaten die sich auf das verwendete Taxon beziehen und die in einer bestimmten Notation die Meinung des Autoren zu den jeweiligen Synonymien ausdrücken. Diese bibliographischen Eigenschaften einer Synonymieliste laden dazu ein, bibliometrische Techniken anzuwenden um etwa ein Ranking von Synonymen vorzunehmen. Wir stellen dazu TaxonRank vor, ein neuer Rankingalgorithmus, der auf bibliometrischen Methoden und Internettechniken basiert

TaxonRank nutzt publizierte Synonymielisten die digitalisiert und in TaxonConcept, einem zu Stratigraphy.Net gehörigen taxonomischen Informationsystem, vorliegen.

Der bibliometrische Algorithmus ist weiter dahingehend modifiziert als auch die Zuverlässigkeit der Artidentifikation, ausgedrückt durch die Offene Nomenklatur, berücksichtigt wird.

Synonymien sind ein wohl bekanntes Problem für geowissenschafliche Datenbanken, die meist die ursprüngliche Klassifikation durch die Datensatzautoren belassen und Änderungen in der Nomenklatur nicht mitführen können. Manche dieser Datenbanken sind deshalb schwierig abzufragen und es mag schwierig werden so Datennetzwerke aufzubauen. Technologien wie TaxonRank werden hilfreich für solche Netzwerke sein um ausgeklügelte Ontologien zu erzeugen und zu analysieren, die es dem Endnutzer zumindest ermöglichen adäquate Anfragen zu formulieren.



Internet: http://taxonconcept.stratigraphy.net/