

Our company has developed several innovative software components to overcome technical barriers in text mining, information extraction, document classification and annotation.

We have developed powerful grammars that are able to precisely identify complex terms such as chemical entities, scientific names and many classes of unique identifiers.

A KWIC index (see reverse) provides a view the context of any phrase across all documents in a corpus.

N4L::Scribe is a powerful document annotation service that can locate important phrases in documents (e.g., Microsoft Word, XML) and embed semantic links to third party resources. This service can also enrich documents with glossaries and summaries of resources (e.g., gene sequences), giving additional insight for readers and examiners. Custom terminologies can be added to extend the capabilities of this software.

Our unique visualization tools can help to quickly locate the diagnostic terms for groups of related documents, which can greatly expedite the patent examination process and assist with IP valuation.

Our classification tools provide novel ways to navigate and bridge various patent classification systems, enabling more precise classification and integration with additional proprietary classifications.

These individual software components have been integrated into a single platform that can support a variety of document analysis needs. Our software may be deployed in a web service container, as a desktop application, or extended/integrated with third party software via our developer API.

Backed by the Fairview Research Alexandria platform (CLAIMS® Global Patent Database), this analysis suite has access to the full text of the worldwide patent literature.

The Semantic Desktop (see reverse) is a document classification application built upon this platform. We demonstrate the ability to reverse-engineer the diagnostic phrases that human indexers use to classify large corpora of technical documents, and to measure both the quality of previously-annotated documents and the cohesion of individual document classifications.

## About N4L

**NamesforLife, LLC is a Michigan based software company that has pioneered semiotic classification methods and semantic annotation technology for the life science, technical and medical literature.**

**The company provides taxonomic and analytical services, data, software and technology licensing for the academic publishing industry, life sciences research, commercial partners and federal laboratories.**

**Our technology arose from a need to support thesaurus construction, vocabulary integration and ontology development. As a result, we have created bridges between document analytics and important industry standards for knowledge representation.**

**Our patented technology produces high-quality data sets from scientific, medical and legal literature via its partnerships with the academic publishers, and Fairview Research, LLC.**

NamesforLife semiotic classification and semantic annotation technology are covered by patents US 7,925,444, US 8,036,997 and US 8,903,825. Other US and PCT patents pending.

CLAIMS® is a registered trademark of IFI CLAIMS Patent Services.

# N4L Semantic Analysis Platform

Patent Databases



CLAIMS® Global Patent Database

Classification Standards

USPC CPC IPC ECLA

Content



Documents (XML, Word, ODF)

## N4L PLATFORM

### Enrich Documents

The N4L::Scribe annotation tool enriches documents with related information and embeds semantic links to additional resources.

atmospheric nitrogen in association with annual and perennial pasture in agriculture. Some of the most important associations in temperate are the *Eschscholium-Medicago* symbioses that produce edicago is a genus within tribe *Trifolieae*, which is included in the "hionoid" Inverted Repeat Lacking Clade (IRLC) legumes [1, 2]. ally specific: nearly all studied species are nodulated by strains of *Eschscholium* ATCC700748 (GenbankID ATTR00000000) or the *Eschscholium* [4, 5].

**Comment [N4L2]:** *Eschscholium* is the current name for this taxon. The current name is *Eschscholium*. Consider revising the document to refer to the current name.

**Comment [N4L3]:** External link: <https://www.ncbi.nlm.nih.gov/taxonomy/1200000000>

### Entity Recognition

Our custom grammars accurately recognize complex entities in text, such as chemical compounds, scientific names and specialized identifiers, which can then be linked to appropriate resources or used for indexing, query or thesaurus construction.

Information: 1 occurrence of "3'-O-(4'-phosphoryl)-D-glucopyranose" in document 21,790/38,645

Entity: 3'-O-(4'-phosphoryl)-D-glucopyranose

Entity Type: Chemical Compound

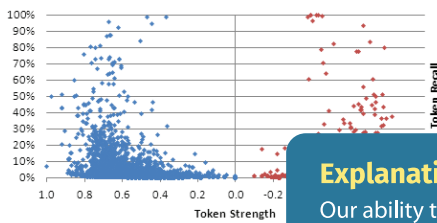
Entity Score: 0.999

Entity Link: [https://pubchem.ncbi.nlm.nih.gov/compound/3-O-\(4-phosphoryl\)-D-glucopyranose](#)

### Visualization

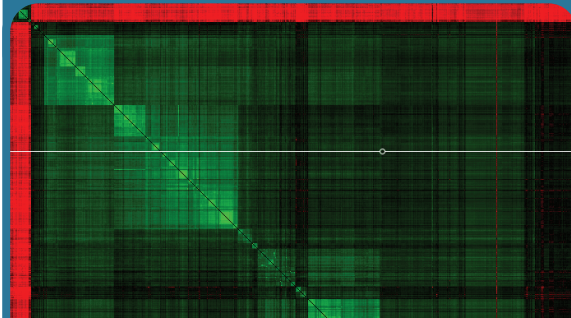
Multidimensional analysis of language used in groups of documents provides a means to visualize the precision of manually classified documents.

#### 4227 (Polymerization Catalyst)



### Taxomatic

Our self-organizing heatmap component clusters documents based on semantic (term) and semiotic (concept) similarity.



### Semantic Similarity

A semantic similarity measure provides a consistent way to compare documents.

Document Classifications	11297	1298	362	2658	103	224	2655	8
US-7521165-B2	0.033	0.23	0.014	0.011	0.012	0.023	0.012	0.023
US-4407984-A	0.055	0.37	0.009	0.009	0.015	0.019	0.038	0.022
US-4892478-A	0.05	0.506	0.014	0.013	0.012	0.011	0.025	0.025
US-488	0.014	0.009	0.012	0.009	0.017	0.017	0.032	0.032
US-604	0.008	0.007	0.01	0.01	0.01	0.009	0.018	0.018
US-493	0.01	0.007	0.015	0.014	0.013	0.037	0.037	0.037
US-516	0.009	0.007	0.01	0.01	0.01	0.03	0.035	0.035
US-031	0.013	0.005	0.022	0.009	0.016	0.098	0.098	0.098
US-248	0.007	0.012	0.011	0.03	0.016	0.016	0.016	0.016
US-379	0.009	0.004	0.142	0.008	0.011	0.022	0.022	0.022
US-304	0.008	0.006	0.12	0.01	0.047	0.051	0.051	0.051
US-449	0.04	0.008	0.123	0.014	0.016	0.044	0.044	0.044
US-402	0.007	0.017	0.012	0.009	0.036	0.018	0.018	0.018
US-248	0.005	0.011	0.011	0.01	0.013	0.019	0.019	0.019
US-041	0.01	0.011	0.023	0.011	0.027	0.029	0.029	0.029
US-556	0.011	0.014	0.01	0.012	0.023	0.02	0.02	0.02
US-262	0.01	0.012	0.012	0.015	0.024	0.024	0.024	0.024

### Explanations

Our ability to reverse-engineer human classifications enables automatic multiple-classification, with explanations.

Document Scores:

- US-4732843-A
- US-4812489-A
- US-7368484-B2
- US-4269741-A
- US-6034174-A
- US-6989225-B2
- US-4696990-A
- US-5143992-A
- US-4956140-A
- US-3944512-A
- US-5516920-A
- US-6232433-B1
- US-4845175-A
- US-5539071-A

Token List:

- catalyst components
- the catalyst
- catalyst component
- catalyst and
- ant catalyst
- catalyst composition
- catalytic
- preparation of
- hydrogen
- carried out
- hydride
- hours
- preparation
- was added

Document Title: Additive system to improve adhesion and hydrolytic stability of silicone elastomers

### Diagnostic Terms

Our analysis tools can extract and chart the diagnostic terms for groups of documents.

