

The *Petroleum Abstracts*[®] Exploration and Production Thesaurus

Position Paper for the W3C Workshop on Semantic Web in Energy Industries. Part I: Oil & Gas

John A. Bailey, Thomas E. Burchfield and Bethany J. Crow
Petroleum Abstracts

Since 1965, The Petroleum Abstracts (PA) Service at The University of Tulsa has developed and maintained an Exploration & Production thesaurus used to index and classify the technical articles abstracted by the service. Originally the Exploration & Production Thesaurus, the taxonomy consisted of words and terms representing technical concepts important to the industry. In 1970, the Geographic Thesaurus was created that contained terms describing sedimentary basins, named geologic features and geographic areas. The taxonomy also contains company names and specific chemical names.

The terms are interrelated on the concept-oriented (faceted) principle, with the taxonomy containing eleven facet headings. Most terms in the taxonomy are classified into a hierarchy of concepts and numerous cross references are included to aid the indexer and searcher in finding the proper term to describe a particular concept. In addition, many synonyms are structured to point the way to preferred terms. The taxonomy is updated daily by a team of engineers and earth scientists.

The main categories covered include alternate fuels and energy sources; business and economics; drilling; geochemistry; geology; geophysics; health, safety and environment; pipelining, shipping and storage; production of oil and gas; reservoir engineering and recovery methods; science and engineering; well completion and servicing; and well logging and surveying.

For many years, the taxonomy was published as a series of books and hard copy lists. In 2003, the taxonomy was developed into the *Bricks* Web based system that provides the user with a number of tools useful in locating specific terms. There are currently some 70,000 terms and identified synonyms contained in the *Bricks* system maintained by PA.

Terms are currently limited to a maximum length of 26 characters, including spaces between words. For longer terms, abbreviations are required. Under each entry term, the following information is displayed: whether the term is valid or invalid, indexible or non-indexible, autopostable or non-autopostable, the entry date by month and year, and the number of times the term has been used in the Petroleum Abstracts database.

Under each entry term, one or more of the following entries may also be found:

- Complete spelling of an abbreviated entry term.
- SN (Scope Note) entries are used to restrict the scope of an index term or to define its meaning, to indicate previous (now invalid) scope notes, to tell when the term was first available for indexing, to show what terms were used previously to describe this concept, and to indicate changes in hierarchical relationships.
- US (Use) entries indicate the valid term used for indexing instead of the entry term.

- PLS (Plus) indicates the second term of a two-term synonym; used with the USE statement.
- UF (Used For) indicates an invalid term that is directed to the valid term under which it is listed.
- WTH (With) indicates the second term of a two-term synonym; used with the UF statement.
- NT (Narrow Term) designates a term which is a more specific subdivision of the entry term.
- BT (Broad Term) designates one or more hierarchically related terms, of which the concept is a logical subdivision.
- SA (See Also) usually designates terms that are related in meaning but not directly connected in a vertical hierarchical relationship. It may be used also to show terms of alternate or opposite meaning.

Bricks display example of a valid E&P term:

VISCOUS OIL RECOVERY

valid: yes

index: yes

autopost: yes

entered: 4/1965

used: 11038/6397

UF HEAVY OIL RECOVERY

BT OIL RECOVERY

RECOVERY

PHYSICAL OPERATION

PROCESS

SA COLD PRODUCTION

SA IN SITU COMBUSTION

SA OIL MINING

SA OIL SAND

SA TAR SAND OIL RECOVERY

SA THERMAL RECOVERY

SA THERMIT STIMULATION

SA UNCONVENTIONAL OIL RECOVERY

SA VISCOSITY REDUCING

SA VISCOUS CRUDE OIL

Historical Notes:

BT (67---) OIL RECOVERY

Bricks display example of an invalid term:

HEAVY OIL RECOVERY

valid: no
index: no
autopost: no
entered: 1/1965
used: /

US VISCOUS OIL RECOVERY

Bricks display example of Geographic valid term:

RED SEA

valid: yes
index: yes
autopost: yes
entered: 2/1965
used: 1410/764

UF RED SEA AREA
UF RED SEA BASIN

NT ALBATROSS DEEP
NT AQABA GULF
NT ATLANTIS II DEEP
NT BAB EL MANDEB STRAITS
NT CHAIN DEEP
NT DAHLAK ARCHIPELAGO
NT DISCOVERY DEEP
NT KEBRIT DEEP
NT NEREUS DEEP
NT PORT SUDAN DEEP
NT SARSO ISLAND
NT SUAKIN DEEP
NT VALDIVIA DEEP

BT SEAS AND OCEANS
WORLD

SA FARASAN ISLANDS
SA GREAT RIFT VALLEY
SA MERSA ALEM AREA
SA ZULA GULF

Historical Notes:
ADDED FEBRUARY 1965

Position

For the last 43 years PA has developed a thesaurus of terms representing technical concepts of the Petroleum Exploration and Production industry. The terms in the thesaurus have been selected based on abstracting and indexing more than 900,000 articles selected from a diverse body of technical literature including 300 journals, 250 conference proceedings, patent gazettes and full-text patents, government documents and reports, dissertations, books, maps and other sources. This thesaurus reflects the efforts of engineers, earth scientists and information professionals and should be a valuable tool in the development of Semantic Web Technologies for the Oil and Gas Industry.

Reference

Bailey, J.A., Martinez, S.J. and Wattenbarger, D.W.; Interactive System for Controlled Vocabulary Maintenance; Annual Assoc. Computing Mach. Conf. Proc., pp. 79-85, 1977.