

Noesis - Semantic Search

Rahul Ramachandran and Sunil Movva
Information Technology and Systems Center

Traditional Search Engines

- Provide syntax-based text string search.
- Keywords from the content of the webpages are indexed.
- The basis for search is term matching between the user's query and these databases.

An Example

The image shows a screenshot of a Mozilla Firefox browser window displaying a Google search for "Cyclone". The search results are categorized into "Web" and "Sponsored Links". Two green circles, labeled "Related Results" and "Unrelated Results", have arrows pointing to specific search results. The "Related Results" circle points to the "FAQ: HURRICANES, TYPHOONS, AND TROPICAL CYCLONES" result. The "Unrelated Results" circle points to the "Cyclone Microsystems: PCI Express Expansion Systems, PCI Express ..." result.

Web Results 1 - 10 of about 1,090,000 for **Cyclone** [definition]. (0)

FAQ: HURRICANES, TYPHOONS, AND TROPICAL CYCLONES
Dr. Chris Landsea's TC FAQ, NOAA's Atlantic hurricane outlook, hurricane shutter info., AOML climatological archive, NWS current weather conditions, ...
www.aoml.noaa.gov/hrd/tcfaq/tcfaqHED.html - 36k - [Cached](#) - [Similar pages](#)

AT&T Labs Research
Cyclone is a programming language based on C that is safe, meaning that it rules out programs ... **Cyclone** is a joint project of AT&T Labs Research, Harvard, ...
www.research.att.com/projects/cyclone/ - 23k - [Cached](#) - [Similar pages](#)

Cyclone Commerce
www.cyclonecommerce.com/ - 1k - [Cached](#) - [Similar pages](#)

Cyclone - Wikipedia, the free encyclopedia
Radar image of a tropical **cyclone** in the northern hemisphere. ... Tropical **cyclones** are distinguished from other cyclonic windstorms such as nor'easters, ...
en.wikipedia.org/wiki/Cyclone - 44k - [Cached](#) - [Similar pages](#)

Cyclone Microsystems: PCI Express Expansion Systems, PCI Express ...
Cyclone Microsystems: Developer of PCI Express Expansion Systems, PCI Express Intelligent I/O Processors, Customers Specific Boards and Systems, ...
www.cyclone.com/ - 13k - [Cached](#) - [Similar pages](#)

Cyclone Forwarding
This copy of the **Cyclone** home page has moved to Harvard. Now Forwarding... If forwarding doesn't work, click here: ...
www.cs.cornell.edu/Projects/cyclone/ - 1k - [Cached](#) - [Similar pages](#)

Iowa State
Iowa State News, scores, schedules, stats, live video, live audio, on-demand video.
www.cyclones.com/ - 96k - Nov 9, 2006 - [Cached](#) - [Similar pages](#)

Cyclone
C dialect, based partly on Popcorn, type-safe, polymorphic, gives C control over data representation and memory management, but prevents buffer overflows, ...
www.eecs.harvard.edu/~greg/cyclone/ - 1k - [Cached](#) - [Similar pages](#)

Sponsored Links

Cyclone
Free Shipping Both Ways
365 Day Return Policy
www.Zappos.com

Original Vortec Cyclone
Save Gas, Boost Horsepower
Back Guarantee. Now on Sale!
www.VortecCyclone.com

Live Cyclone Warning
Tornado Tracking, Hourly
Radar & More - The Weather Channel
www.weather.com

Dustvent Dust Collector
Cyclone dust collector
Quiet & efficient. Made in USA
www.dustvent.com

Cyclone For Less
Looking for **Cyclone**?
Find exactly what you want
www.eBay.com

Slumber Party-Adults
Empower & Entertain, It's
Shop Online or Host A Party
www.slumberparties.com

Cyclone
Cyclone at Wal-Mart™
Wide Selection, Always Low Prices
www.walmart.com

The Problem

- Search **query is broad** and not scoped.
- Only **one kind of resource** is searched (E.g., Web-documents).
- Semantics associated with the search string are not captured.

Noesis

Noesis is an attempt to solve these issues

- Search only what you want (Search Scoping)
- Search everything that you want (Resource aggregation)

Noesis Approach

- Noesis provides topic-based searches, retrieving resources **conceptually related** to the user's need.
- Semantic information is captured in the **domain ontologies** and is used for providing context for the search process.

Noesis Approach (cont'd)

- Using domain ontologies, Noesis provides a **guided refinement of search query** producing successful searches and reducing the user's burden to experiment with different search strings.
- Semantics are captured by providing the option to expand the query terms to include suggested **related concepts**.

Noesis Algorithm

Noesis search engine uses a two step algorithm.

- Query Analysis: Search query is broken down to **identify concepts** that are defined in the domain.
- Query Expansion: **Related concepts** are added to the search string for Scoping the search.

Ontology

- From Machine Learning/Artificial Intelligence/Intelligent Systems perspective “an ontology is a formal, explicit specification of a shared conceptualization” (*Gruber, 1993*).
- An ontology captures and encodes **domain knowledge** of concepts, constraints and the relationships among them, for use in a machine-readable fashion.

Noesis Ontologies

Noesis uses two classes of ontologies.

- **Domain ontologies:** These are the ontologies used to describe concepts in a domain and their relationships. Noesis uses a set of core ontologies for describing concepts in Atmospheric Science.
- **Application ontologies:** Application ontologies describe a domain with respect to a function (E.g., A data archive). They enable flexible querying by bridging the gap between semantic concepts and the application specific vocabulary.

Providing Semantics - Specializations

- Adding **related concepts** to the search term as a way to provide semantic information.
- Ontologies are organized in tree like taxonomies, where the child nodes represent the **specializations** of a parent node.
- The first type of the related terms are specializations (SP). They are used to expand the search terms thus providing more detailed search. For example:
 - Search Term: “height
 - Additional term(SP): “free-board”

Providing Semantics - Synonyms

- Synonyms are different terms that have the same meaning. In ontological terms they are the **equivalent concepts**.
- owl:equivalent-class allows linking two **syntactically different terms** to one semantic concept (synonyms).
- The second form of related concepts are synonyms (SN). These include acronyms. For Example:
 - Search Term: 'Albedo'
 - Additional Term(SN): 'Reflectance'

Providing Context

- In reality, every concept has a set of **other concepts**, that are neither in the same inheritance hierarchy nor equivalent, that tend to co-exist.
- These are called the **associated concepts** and they are captured in the ontology through the property relationships.
- Associated concepts are the third type of related terms (RT) and they provide the context for the search query. For example
 - Search Term: 'Mesocyclone'
 - Associated Term (RT): 'Vorticity'

Semantics for Search Scoping

Specializations and related terms for the search term (“Cyclone”) presented to the user

Specializations

Refine Search

+Atmospheric Circula...

+Cyclone

- Hurricane
- Typhoon

Related Terms

+Phenomena

- Rain
- Storm
- Cloud
- Precipitation
- OccludedFront

+Property

- Temperature

+Earthrealm

- TropicalClimateZ...

+Process

- Rotation

+Substance

- Air

Add Additional Terms

Add

Related Terms

coexist, in commo
interchangeably. Also
inclement (often des
storms. See tropical

FAQ : HURRICANES,
Google
Dr. Chris Landsea's
shutter info., AOML cl
http://www.aoml.noaa

AT&T Labs Research
Google
Cyclone is a program
rules out programs
Harvard, ...
http://www.research.e

Cyclone Commerce
Google
http://www.cyclonecor

Cyclone - Wikipedia,
Google
Radar image of a t
cyclones are distir
nor'easters, ...
http://en.wikipedia.org

Cyclone Microsystem
Google

A user can start at a general topic and navigate to specific topics by selecting the related concepts of interest.

Semantics for Search Scoping

Synonyms for the search term (“Reflectance”) presented to the user

Synonyms

Refine Search

+Radiation- Medium I...

+Reflectance

Synonym

Albedo

Add Additional Terms

Add

bidirect

Near In
(PTFE)

Yahoo

... is c

hemisp

of PTFE

http://nv

Near In

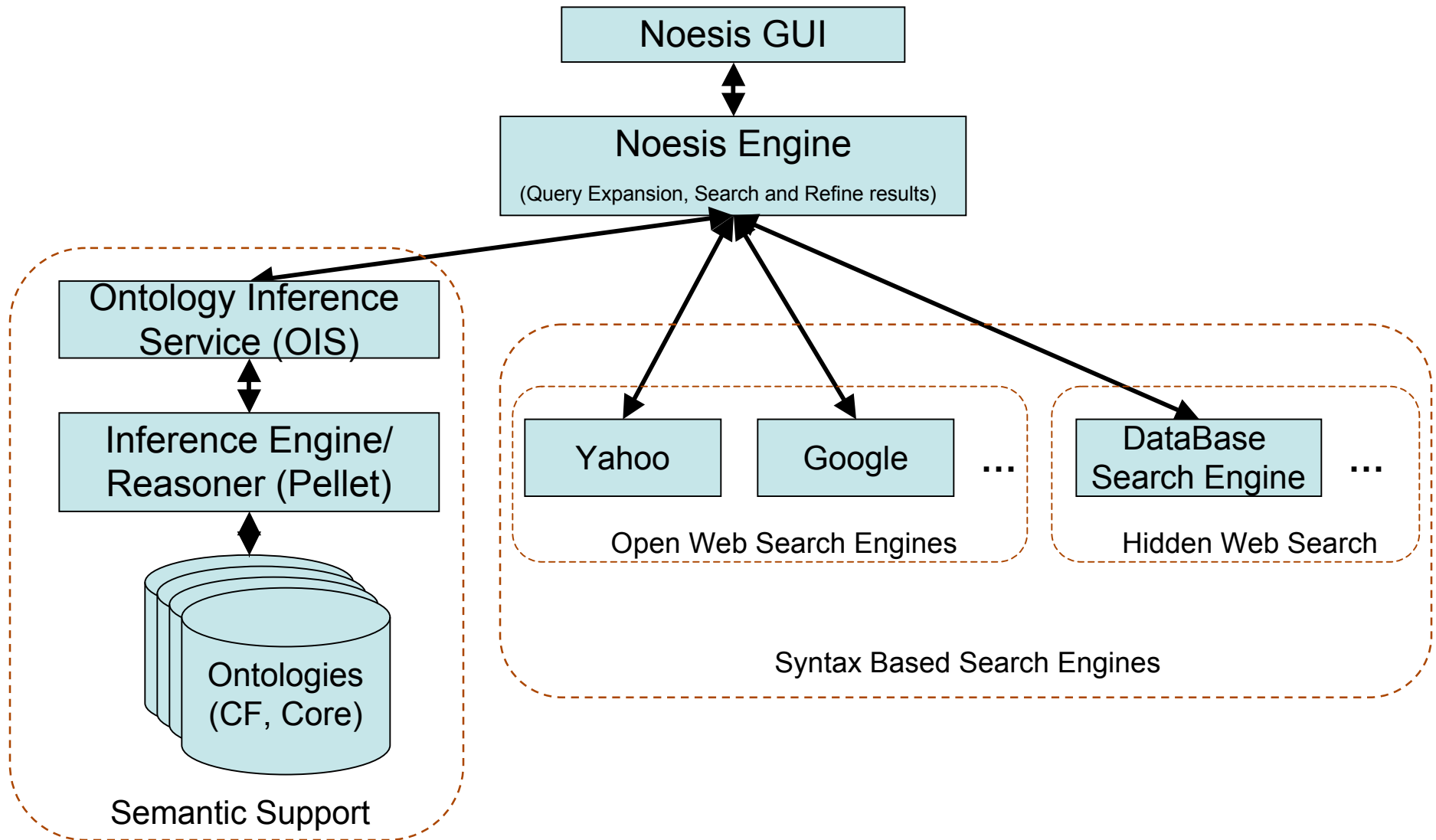
(PTFE)

Yahoo

The Search

- The identified concepts from the query string are used to search the Ontology inference service (OIS) to get the related concepts (SP,SN and RT).
- The obtained terms are used to search other resources like Google, Yahoo etc
Search Query:(SP1.. SPn SN1..SPn RT1..RTn).

The Search (Cont'd)



DataSet Interoperability

- Over Years, data formats have evolved from different communities, thus making them limited to the community and context.
- Semantically homogeneous data is archived as **syntactically heterogeneous data** sets. For example:
 - DataSet1: Albedo
 - DataSet2: Reflectance
- **Semantic Interoperability** between data formats is required to make them usable across communities.

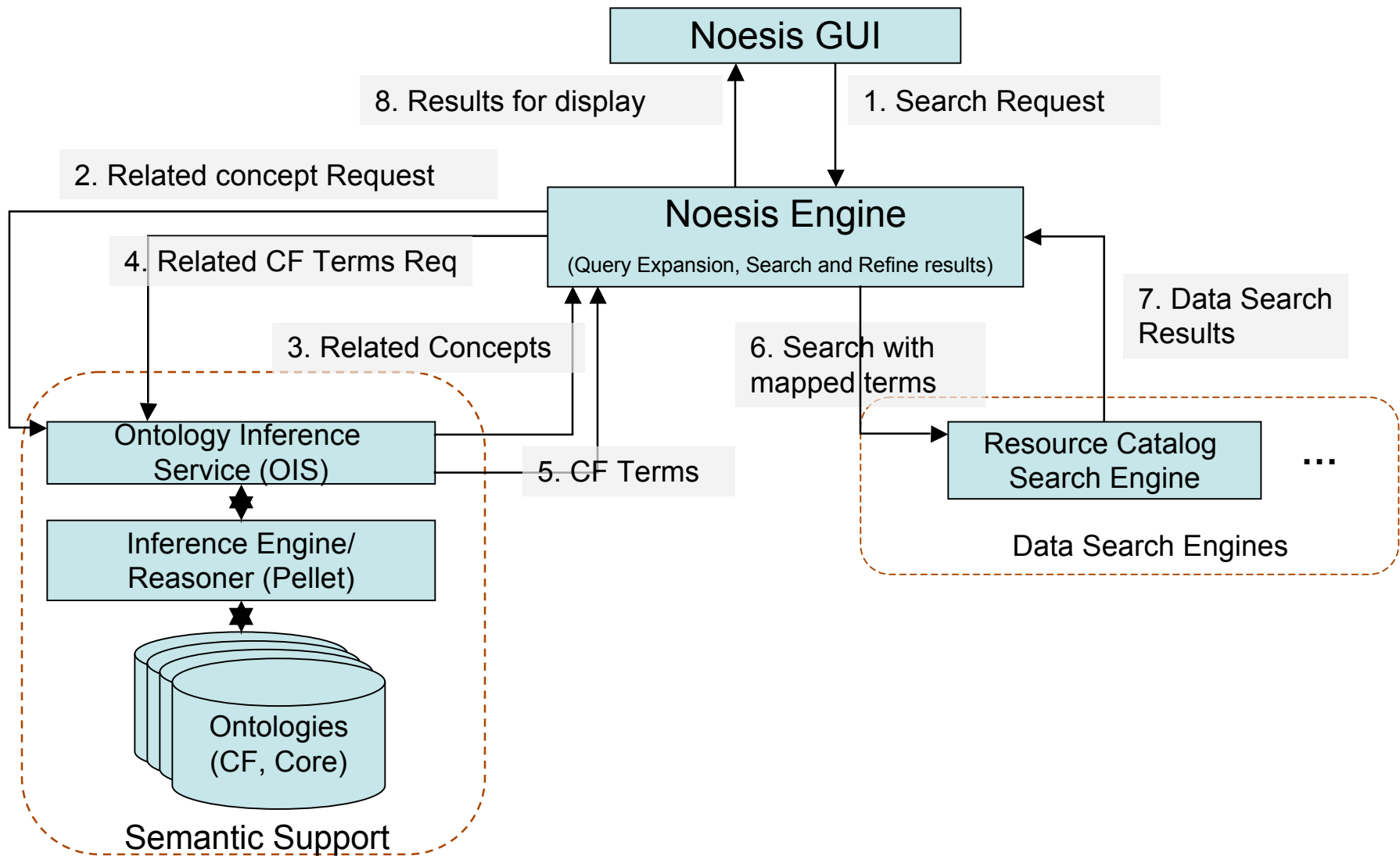
Application Ontologies

- Data archives are often catalogued with content meta data to enable searching. But, users still need to search for syntactically matching keywords.
- **Application ontologies describe a domain with respect to function** (i.e. Data archive).
- Application ontologies enable flexible querying by bridging the gap between semantic concepts and the keywords.

Application Ontologies (Cont'd)

- An application ontology is added for every new catalog (E.g.: CF-Ontology).
- This ontology annotates the terms used in the catalog with the conceptual meanings.
- The concepts in this ontology are linked with Core ontology (Modified SWEET) through the owl:equivalentClass.

Noesis - Data Search



Resource Aggregation

Noesis is a meta search engine

- Meta Search Engines simultaneously search multiple **Open Web and Hidden Web resources** to provide increase search coverage, efficiency and effectiveness.
- Searches for **web-pages, data, education material and Publications.**

Resource Aggregation (Contd.)

Noesis uses the scoped search string to fetch resources, through **search web services** provided by third parties:

- **Web:** Yahoo, Google, Ask.com
- **Data:** NCDC, NCAR, NOAA, NASA GCMD, LEAD, IPCC Model Output, LDEO
- **Publications:** AMS, Elsevier, Springer, RMS, Blackwell, AGU
- **Education:** DLESE

Resource Aggregation (Contd.)

- Filter by Engine
- + All
 - + Web
 - Google
 - Yahoo
 - Ask.com
 - + Data
 - LEAD
 - NCDC
 - LDEO
 - NCAR
 - NOAA
 - NASA GCMD
 - IPCC Model Output
 - + Publications
 - AMS
 - Elsevier
 - Springer
 - RMS
 - Blackwell
 - AGU
 - + Education
 - DLESE

Collates resources such as web pages, data sets, publications etc.

Example – Step 1 (User Input)

noesis

[Home](#) | [About](#) | [FAQS](#) | [Contacts](#) | [ITSC](#) | [Disclaimer](#)

User types a search term ('Cyclone')

Enter a term to search

Getting started: Enter an atmospheric science related term such as hurricane or pressure.

Example – Step 2 (Query Refinement)

User selects the related terms to add to his search query

Specializations

Related Terms

The screenshot shows a search interface for the term 'Cyclone'. At the top, there is a 'Refine Search' section with a green header. Below it, there are several expandable categories, each with a '+' sign and a list of terms with checkboxes. The 'Cyclone' term is highlighted in red. The categories and their terms are:

- +Atmospheric Circula...**
 - +Cyclone** (highlighted in red)
 - Hurricane
 - Typhoon
- Related Terms** (highlighted in green)
 - +Phenomena**
 - Rain
 - Storm
 - Cloud
 - Precipitation
 - OccludedFront
 - +Property**
 - Temperature
 - +Earthrealm**
 - TropicalClimateZ...
 - +Process**
 - Rotation
 - +Substance**
 - Air
- Add Additional Terms** (highlighted in green)
 -
 -

On the right side of the interface, there are search results for 'Cyclone'. The first result is 'FAQ : HURRICANES, Google' with a URL 'http://www.aoml.noaa.gov/faq_hurricanes/'. The second result is 'AT&T Labs Research, Google' with a URL 'http://www.research.att.com/~larry/cyclone/'. The third result is 'Cyclone Commerce, Google' with a URL 'http://www.cyclonecommerce.com/'. The fourth result is 'Cyclone - Wikipedia, Google' with a URL 'http://en.wikipedia.org/wiki/Cyclone'. The fifth result is 'Cyclone Microsystems, Google'.

Example – Web Results (Step 3)

noesis

Home | About | FAQs | Contacts | RSS | Disclaimer

Search

Refined query

Number of Results:
10

Definition:
An atmospheric cyclonic circulation, a closed circulation. A cyclone's direction of rotation (counterclockwise in the Northern Hemisphere) is opposite to that of an anticyclone. While modern meteorology restricts the use of the term cyclone to the so-called cyclonic-scale circulations, it is popularly still applied to the more or less violent, small-scale circulations such as tornadoes, waterspouts, dust devils, etc. (which may in fact exhibit anticyclonic rotation), and even, very loosely, to any strong wind. The first use of this term was in the very general sense as the generic term for all circular or highly curved wind systems. Because cyclonic circulation and relative low atmospheric pressure usually coexist, in common practice the terms cyclone and low are used interchangeably. Also, because cyclones are nearly always accompanied by inclement (often destructive) weather, they are frequently referred to simply as storms. See tropical cyclone, extratropical cyclone; compare trough.

Query: Cyclone + Hurricane + Storm

Relevant Results

Querying Web Search Engines

Refine Search

- + Atmospheric Circula...
- + Cyclone
 - Hurricane
 - Typhoon

Related Terms

- + Phenomena
 - Rain
 - Storm
 - Cloud
 - Precipitation
 - OccludedFront
- + Property
 - Temperature
- + Earthrealm
 - TropicalClimateZ...
- + Process
 - Rotation
- + Substance
 - Air

Add Additional Terms

Add

Search Results

- worldwide tropical cyclone names
Google
<http://www.nhc.noaa.gov/aboutnames.shtml>
- Tropical cyclone - Wikipedia, the free encyclopedia
Google
A hurricane or typhoon (sometimes simply referred to as a tropical cyclone, as opposed to a depression or storm) is a system with sustained winds greater ...
<http://en.wikipedia.org/wiki/Hurricane>
- Navy hurricanes, typhoons, satellite images, forecasts, tracks
Google
NRL Monterey Marine Meteorology Division (Code 7500) Tropical Cyclone Page (Ver.4.13.0). Development Team ... Forecast by: National Hurricane Center ...
http://www.nrlmry.navy.mil/tc_pages/tc_home.html
- Tropical Weather : Weather Underground
Google
CIRA's experimental tropical cyclone formation products · University of Wisconsin CIMSS tropical products ... Storm Advisories / Hurricane RSS feed! ...
<http://www.wunderground.com/tropical/>
- Historical Hurricane Tracks Home - NOAA Coastal Services Center
Google
Tropical Cyclone Reports: Access storm reports written by hurricane specialists at the National Hurricane Center. Reports are available for the Atlantic and ...
http://www.csc.noaa.gov/hurricane_tracks/

Filter by Engine

- + All
- + Web
 - Google
 - Yahoo
 - Ask.com
- + Data
 - LEAD
 - NCDC
 - LDEO
 - NCAR
 - NOAA
 - NASA GCMD
 - IPCC Model Output
- + Publications
 - AMS
 - Elsevier
 - Springer
 - RMS
 - Blackwell
 - AGU
- + Education
 - DLESE

Example – Publications (Step 3)

Refine Search	Search Results	Filter by Engine
<p>+Cyclone +Hurricane</p> <p>Related Terms</p> <p>+Phenomena</p> <ul style="list-style-type: none">- <input type="checkbox"/> Wind- <input checked="" type="checkbox"/> Storm- <input type="checkbox"/> Cloud- <input type="checkbox"/> Rain- <input type="checkbox"/> Precipitation <p>+Property</p> <ul style="list-style-type: none">- <input type="checkbox"/> Temperature <p>+Earthrealm</p> <ul style="list-style-type: none">- <input type="checkbox"/> Atmosphere <p>+Process</p> <ul style="list-style-type: none">- <input type="checkbox"/> Rotation <p>+Substance</p> <ul style="list-style-type: none">- <input type="checkbox"/> Air	<p>Increased SpringerLink ...sample. In terms of storm tracks, the high CO2 sample is quite similar to the control. The mean radius of hurricane force winds is 2... http://www.springerlink.com/content/xhbfv25acjxufnf0/?p=461e...</p> <p>Vorticity asymmetries in Hurricane Josephine (1984) RMS Author(s): ARNO GLATZ; ROGER K SMITH http://oberon.ingentaconnect.com/vl=883209/cl=21/nw=1/rpsvf~...</p> <p>The detection of hurricane asymmetries from aircraft reconnaissance flight data: Some simulation experiments RMS Author(s): ROGER K SMITH; ARNO GLATZ http://oberon.ingentaconnect.com/vl=883209/cl=21/nw=1/rpsvf~...</p> <p>Symmetrization of lopsided vorticity monopoles and offset hurricane eyes RMS Author(s): R PRIETO; JP KOSSIN; WH SCHUBERT http://oberon.ingentaconnect.com/vl=883209/cl=21/nw=1/rpsvf~...</p> <p>Angular momentum variation in a trans... RMS Author(s): STEVEN SKUBIS; JOHN MOL... http://oberon.ingentaconnect.com/vl=883209/cl=21/nw=1/rpsvf~...</p> <p>A numerical simulation of Hurricane Bret on 22-23 August 1999 initialized with airborne Doppler radar and dropsonde data RMS Author(s): Olivier Nuissier; Robert F. Rogers; Frank Roux http://oberon.ingentaconnect.com/vl=883209/cl=21/nw=1/rpsvf~...</p> <p>A theory for vortex Rossby-waves and its application to spiral bands and</p>	<p>+ <input type="checkbox"/> All</p> <p>+ <input type="checkbox"/> Web</p> <ul style="list-style-type: none">- <input type="checkbox"/> Google- <input type="checkbox"/> Yahoo- <input type="checkbox"/> Ask.com <p>+ <input type="checkbox"/> Data</p> <ul style="list-style-type: none">- <input type="checkbox"/> LEAD- <input type="checkbox"/> NCDC- <input type="checkbox"/> LDEO- <input type="checkbox"/> NCAR- <input type="checkbox"/> NOAA- <input type="checkbox"/> NASA GCMD- <input type="checkbox"/> IPCC Model Output <p>+ <input checked="" type="checkbox"/> Publications</p> <ul style="list-style-type: none">- <input checked="" type="checkbox"/> AMS- <input checked="" type="checkbox"/> Elsevier- <input checked="" type="checkbox"/> Springer- <input checked="" type="checkbox"/> RMS- <input checked="" type="checkbox"/> Blackwell- <input checked="" type="checkbox"/> AGU <p>+ <input type="checkbox"/> Education</p> <ul style="list-style-type: none">- <input type="checkbox"/> DLESE
<p>Add Additional Terms</p> <input type="text"/> <input type="button" value="Add"/>	<p><i>Refined query used for publications search</i></p>	

Example – Data Search (Step 3)

Refine Search	Search Results	Filter by Engine
<p>+Cyclone +Hurricane</p> <p>Related Terms</p> <p>+Phenomena</p> <ul style="list-style-type: none">- <input type="checkbox"/> Wind- <input checked="" type="checkbox"/> Storm- <input type="checkbox"/> Cloud- <input type="checkbox"/> Rain- <input type="checkbox"/> Precipitation <p>+Property</p> <ul style="list-style-type: none">- <input checked="" type="checkbox"/> Temperature <p>+Earthrealm</p> <ul style="list-style-type: none">- <input type="checkbox"/> Atmosphere <p>+Process</p> <ul style="list-style-type: none">- <input type="checkbox"/> Rotation <p>+Substance</p> <ul style="list-style-type: none">- <input type="checkbox"/> Air <p>Add Additional Terms</p> <input type="text"/> <p>Add</p>	<p>NASA Atmospheric Infrared Sounder (AIRS) Imagery of Hurricane Rita [NASA_AIRS_RITA] GMCD of Hurricane Rita. AIRS will show how the storm looks through an AIRS Infrared window channel. Window channels measure the temperature of the cloud tops or the surface of the Earth in cloud-free http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&K...</p> <p>National Climatic Data Center's Weather Event Case Study Archive (1991-2006) [NCDC_WX_EVENT_ARCH] GMCD 1992 ---- - Hurricane Andrew--August 1992... The March 1993 "Storm of the Century" - Winter Precipitation in SW Arizona--1992-1993 http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&K...</p> <p>COMET Case Study 003: Hurricane Erin 1995 Data e UCAR/JOSS/NOAA/CODIAC [COMET003_UCAR_JOSS_NOAA_CODIAC] GMCD This case contains data from Hurricane Erin as it affected... States in early August 1995. It includes data from Florid... For more http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Port...</p> <p>A Case Study of Hurricane Andrew (August 16-27, 1992) [UIUC_HURRIC_ANDREW] GMCD tropical depression on August 16, 1992 and became a hurricane 6 days later. Andrew reached landfall in Southern Florida on August 24. http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&K...</p> <p>CD-ROM with METEOSAT Images and Sequence Images of Meteorological Situations; Hurricane, Tropical Storm, Summer and Perturbed Spring Situation [CD_METEOSAT_WEATHER_IN_MOTION_2] GMCD Following situations... summer situation (20 August-1 September 1994) - The hurricane ANDREW over the Caribbean http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&K...</p> <p>1998 Atlantic Tropical Storms: Views from the NOAA Satellites [C00524] GMCD preliminary summary of the 1998 Atlantic Hurricane season. It provides a synopsis of each named tropical storm, using textual information obtained from the National http://gcmd.nasa.gov/KeywordSearch/Metadata.do?Portal=GCMD&K...</p>	<p>+ <input type="checkbox"/> All</p> <p>+ <input type="checkbox"/> Web</p> <ul style="list-style-type: none">- <input type="checkbox"/> Google- <input type="checkbox"/> Yahoo- <input type="checkbox"/> Ask.com <p>+ <input type="checkbox"/> Data</p> <ul style="list-style-type: none">- <input type="checkbox"/> LEAD- <input type="checkbox"/> NCDC- <input checked="" type="checkbox"/> LDEO- <input checked="" type="checkbox"/> NCAR- <input checked="" type="checkbox"/> NOAA- <input checked="" type="checkbox"/> NASA GMCD- <input checked="" type="checkbox"/> IPCC Model Output <p><input type="checkbox"/> Publications</p> <ul style="list-style-type: none">- <input type="checkbox"/> AMS- <input type="checkbox"/> Elsevier- <input type="checkbox"/> Springer- <input type="checkbox"/> RMS- <input type="checkbox"/> Blackwell- <input type="checkbox"/> AGU <p>+ <input type="checkbox"/> Education</p> <ul style="list-style-type: none">- <input type="checkbox"/> DLESE

Mapped query used for data search

Reference

Noesis – Web Interface:

<http://noesis.itsc.uah.edu>