Virtual Observatories and Virtual Grids: The Interplay in Fully Exploiting Solar-Terrestrial Data

Mauro Messerotti

INAF-Astronomical Observatory of Trieste

and

Department of Physics, University of Trieste

Outline of the Talk

- **Historical evolution in S-T-Phys. data handling**
- **Requirements for S-T Physics & Space Weather**
- **The role of VOs & VGrids**
- **Knowledge Embedding via Cmaps**
- **Visions about an ideal VO**
- **Conclusions**

HISTORICAL EVOLUTION IN S-T PHYSICS DATA HANDLING

REQUIREMENTS FOR S-T PHYSICS AND SPACE WEATHER

JOSO Joint Organization for Solar Observations

Solar and Geophysical Databases

Solar and Geophysical Databases: The Tiles of a Planetary Meta-Archive

M. Messerotti

Trieste Astronomical Observatory

Data Organization

• Matter of Fact	Huge amount of space and g-b data
• Data Organization	Databases, Archives, Meta-Archives
• Data Indexing	Tables, Catalogs managed by RDBMS
• Data Access	FTP, TELNET, WWW via GUI
• Data Search	Local, Distributed over the net
• Data Analysis	Local

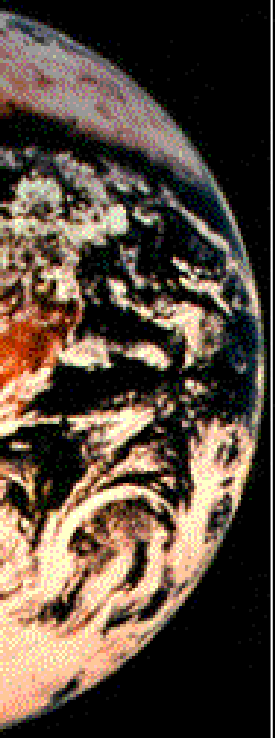
JOSO Joint Organization for Solar Observations

Scientific Requirements

- **Physical modelling** — MULTIWAVELENGTH DATA SEARCH
  MULTIWAVELENGTH DATA DISPLAY
  MULTIWAVELENGTH DATA ANALYSIS
  via a common unified, user-friendly interface
- **Space Weather** — SOLAR, SPACE, EARTH DATASETS
  MULTI-EVENT MODELLING
  LARGEST COVERAGE POSSIBLE
- **Event Prediction** — CROSS-SEARCH OVER ARCHIVES
  STATISTICAL ANALYSES
  REAL-TIME DATA AVAILABILITY

JOSO Joint Organization for Solar Observations

Scientific Motivations



- Some major Solar-Terrestrial Data Portals exist
- Mainly Resource Indexing is available
- Few resources partially allow complex, distributed data searching over limited subsets of databases
- Very few resources partially allow data analysis on inhomogeneous datasets

A PLANETARY META-ARCHIVE IS NEEDED TO EXPLOIT THE FULL SCIENTIFIC POTENTIALITIES OF MULTIWAVELENGTH MODELLING IN SOLAR-TERRESTRIAL PHYSICS

Solar-Terrestrial Physics Portals

CDS AstroWeb
http://cdsweb.u-strasbg.fr/astroweb.html

NASA Space Physics Data System (SPDS)
http://spds.nasa.gov/

NASA Space Physics Data Facility (SPDF)
http://nssdc.gsfc.nasa.gov/spdf/
Magnetospheric Yellow Pages
http://nssdc.gsfc.nasa.gov/spdf/yellow-pages/data-by-type.ht

NASA National Space Science Data Center (NSSDC)
http://nssdc.gsfc.nasa.gov/

Canadian Astronomy Data Center
http://cadcwww.dao.nrc.ca/

CDS AstroWeb - Astronomy on the Internet

http://cdsweb.u-strasbg.fr/astroweb/solar.html (A)

ARTHEMIS	GDC	Hiraiso Solar Terrestrial Research Center/CRL	SEC
BAse Solaire Sol 2000 (BASS2000)	GDC	IPS Radio & Space Services	SEC
Big Bear Solar Observatory (BBSO)	GSO	Imager for Magnetopause-to-Aurora Global Exploration	ESE
Birmingham Solar Oscillations Network (BiSON)	GSN	Institut d'Astrophysique Spatiale (IAS)	SPI
Boulder, Colorado – Dept. Astrophys. and Planet. Sciences	SRI	Instituto de Astronomia y Fisica del Espacio (IAFE)	SPI
Catania Astrophysical Observatory (OAC)	SOO	International Solar-Terrestrial Physics (ISTP)	STN
Centre de Prevision de l'activite solaire et geomagnetique	SEC	Joint Organization for Solar Observations (JOSO)	ISO
Cluster II, ESA's spacefleet to the magnetosphere	SSE	Kharkov multi-wave station of solar monitoring (KHASSM)	SRO
Cracow - Solar radio emission in dm wavelength	SRO	Kiepenheuer-Institut für Sonnenphysik (KIS)	SRI
Departement d'Astronomie Solaire (DASOP, Observatoire de Paris)	SRI	Laboratory for Atmospheric and Space Physics (LASP)	SRI
ETH Institute of Astronomy (ETH Zurich)	SRI	LASCO/SOHO	SSE
Estación de Observación Solar / Solar Observational Station (EOS)	SOO	MEDOC (Multi-Experiment Data Operations Center for SOHO)	SDC
European Incoherent SCATtter	GIN	MSU Solar Physics Group (Montana)	SRI
GALLEX	GPE	Mees Solar Observatory (MSO, Hawaii)	SOO
Global Oscillation Network Group (GONG)	GSN	Metsahovi Radio Research Station	SRO
Haleakala Observatories (Hawaii)	SOO	Mount Wilson Observatory	SOO
High Altitude Observatory (HAO)	SOO	NRL Solar Physics Branch	SRI
High Energy Solar Spectroscopic Imager (HESSI)	SSE	National Astron. Obs. of Japan - Solar Phys. Division	SRI

CDS AstroWeb - Astronomy on the Internet

http://cdsweb.u-strasbg.fr/astroweb/solar.html (B)

National Solar Observatory (NSO)	SOO
NSO Sacramento Peak, Sunspot, NM (NSO/SP)	SOO
Naval Research Laboratory Space Science Division (NRL SSD)	SPI
Observatoire Midi-Pyrenees (OMP)	SOO
Service d'Aeronomie	EDC
Soft X-Ray Telescope onboard Yohkoh Satellite, ISAS, Japan	SSE
Solar Data Analysis Center (SDAC)	SEC
Solar Extreme-ultraviolet Rocket Telescope and Spectrograph	SSE
Solar Flare Theory (NASA/Goddard Space Flight Center)	SRI
Solar Group of RATAN-600	SRO
Solar Physics Division - American Astronomical Society	NSO
Solar Physics at Stanford University	SRI
Solar Terrestrial Activity Report	SEC
Solar Terrestrial Dispatch (STD)	SEC
Solar UV Atlas from HRTS (HRTS data)	SDC
Solar and Heliospheric Observatory (SOHO)	SSE
Solar, Auroral, Ionospheric, ... Information (Lethbridge, Canada)	SEC
Solar-Terrestrial Physics Home Page (STP)	SEC
Space Environment Center	SEC
Stanford SOLAR Center	SEC
Sternberg Astronomical Institute (Heliophys. and Seismology)	SRI
THEMIS	SOI
The INTER-SOL Sun Observation Programme (ISP)	GSN
Transition Region And Coronal Explorer (TRACE)	SSE
Universitat de les Illes Balears - Solar Phys. at Dept. of Phys.	SRI
Wilcox Solar Observatory (WSO)	SOO
Yohkoh Public Outreach Project (YPOP)	SSE
Zurich Solar Radio Spectrometer	SRO

64 Entries

Goals

- Index observational resources in ST Physics
- Index theoretical resources in ST Physics
- Allow
  - User-transparent data access to distributed datasets all over the world
  - Complex data searching, retrieval and analysis via a simplified common GUI

PRESENT DATA ARCHIVING TECHNOLOGIES ALLOW THE ACHIEVEMENT OF SUCH GOALS PROVIDED THAT A GLOBAL COORDINATION AND COLLABORATION IS ESTABLISHED AS WELL AS THE ALLOCATION OF PROPER FINANCIAL RESOURCES BY THE PARTICIPATING ORGANIZATIONS

An Advanced System for Data Handling in SPW

- Title of contribution — STEVM (Solar-Terrestrial Environment Virtual Monitor)
- Proposer — M. Messerotti (& al.)
- Relevant MoU objectives — Data standardization and accessibility
- Relevant parts of MoU sci. programme — Aims of WG4
- Deliverables — Resources survey & architecture (& ?)
- Timetable — 3-5 years according to final goals
- Required Manpower — To be defined according to final goals
- Resources availability — Existing Data Archives community
- Expected collaborations — With main VO projects (e.g. EGSO)
- Previous experience in the field — SOLAR, SOLRA, SOLARNET, EGSO

Space Weather as Driver of Data Homogeneization

- Inhomogeneous and fragmented character of available observations

CAUSES

Difficulties in carrying out a posteriori modelling of complex phenomena

- These limitations are intrinsic to data acquisition mode

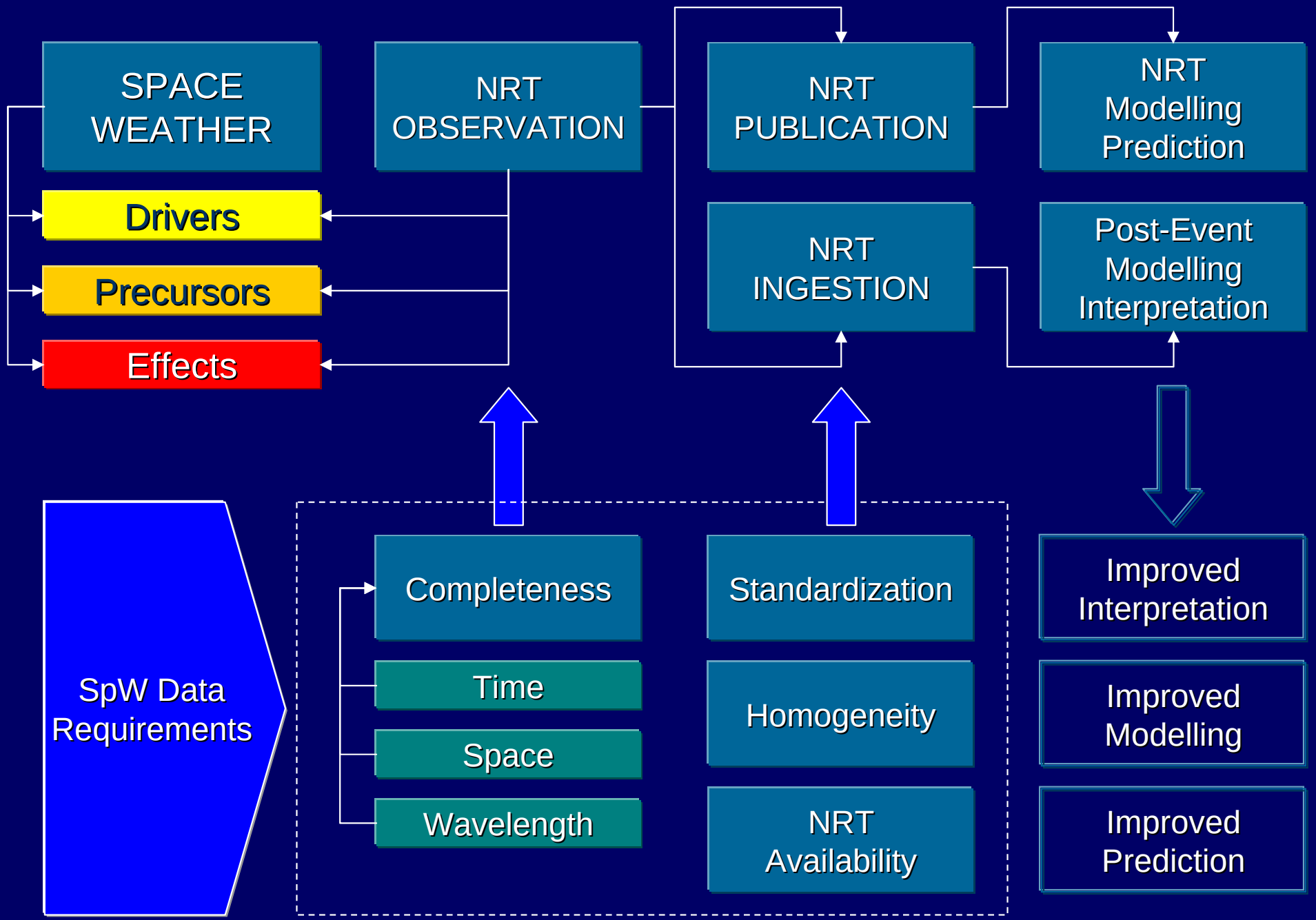
HENCE

Even advanced data search by Grid architectures cannot overcome

DRAWBACKS OF DATA INADEQUACY

- The outcomes are:
  - Inadequate modelling
  - Limited to a subset of phenomenological and physical aspects
  - Often neglects the complex interplays among different processes

Scheme of SpW Data Requirements



The Role of Observing Requirements for SpW

Observing requirements for SpW and SpW drivers observation in monitoring and nowcasting can play a primary role in providing:

1. homogeneization in observations
2. near real-time data ingestion in archives
3. unified data access via web through a user friendly GUI

capable to facilitate:

1. data availability in near real-time
2. full exploitation of the data information content by pointing out interrelationships in different datasets
3. self-consistent modelling

THE ROLE OF VOs AND V-GRIDS

SOLARNET
Italian Solar Archives Federation
Hosted by INAF/Osservatorio Astronomico di Torino

Nodes
- SOLAR
- SOLRA
- PSPT
- DISCO/VAMOS
- CATANIA

EGSO Resources
- EGSO/SEC
- EGSO/DSO

Availability Check

Welcome to the SOLARNET Portal for browsing and retrieving Italian Solar Archives data.

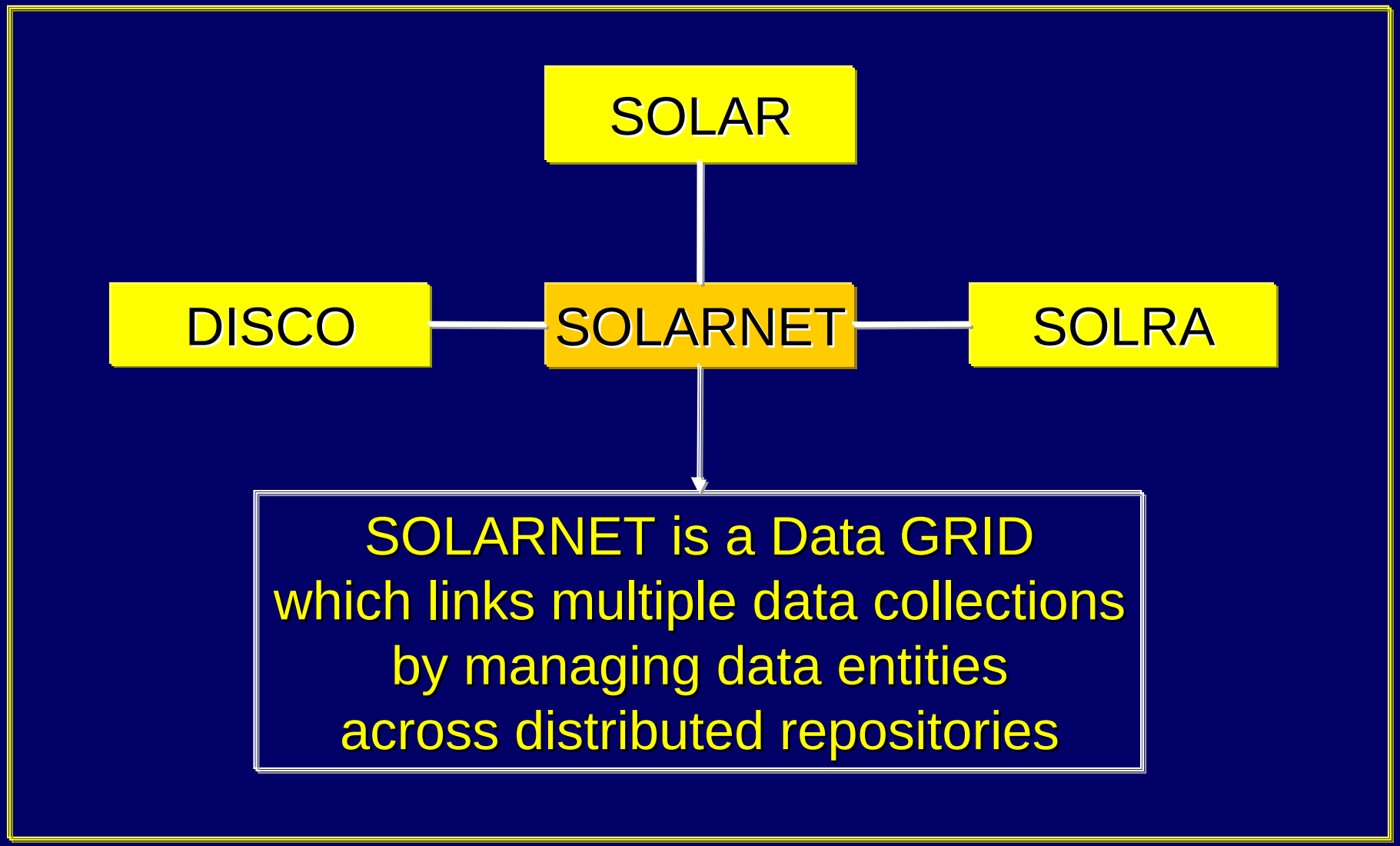
The national project SOLARNET (SOLar ARchive NETwork) aimed to federating all the Italian solar archives as a distributed database, is the first step toward an Italian Virtual Solar Observatory, which interconnects the distributed resources and available solar data in a unified database by a web user interfaces. Different user interfaces allow searches of all participating data services using different input parameters. Currently there are 5 data providers in SOLARNET:
SOLAR , SOLRA , PSPT , DISCO/VAMOS and CATANIA archives, plus two services that the Italian solar comunity has developed for the EGSO project:
SEC (Solar Events Catalog) and DSO (Database for Solar Observatory).

http://solarnet.to.astro.it:8080/portal/

Search by:

Date/Time Date/Time,Instrument

DATA GRID Architecture



European Grid of Solar Observations

Site Navigation

- Home
  - Overview
  - Status
  - Services & Software
  - Help
  - Documents
  - Papers & Presentations
  - Organisation & Contacts
  - Workpackages
  - Communities
    - Physics Communities
    - Grid Community
    - Public & Students
  - Related Projects
  - Internal

Introduction

EGSO, the "European Grid of Solar Observations", is a Grid test-bed that will lay the foundations of a "Virtual Solar Observatory".

EGSO addresses the problem of combining heterogeneous data from scattered archives of space and ground-based observations into a single "virtual" dataset. The project will also create catalogues of solar features and observation data to enable innovative searching, and provide visualisation tools for user-friendly data browsing. EGSO will be a unique resource for the solar physics community, while also serving as an interface to solar data for the Space Weather, Climate Physics and Astrophysics communities.

EGSO is funded under the Information Society Technologies (IST) thematic programme of the European Commission's Fifth Framework Programme. The project is one of many partners from across Europe that co-operate through the EU GRIDSTART initiative. EGSO is also working closely with the Virtual Solar Observatory (VSO), Collaborative Sun-Earth Connector (CoSEC) and the Virtual Space Physics Observatory (VSPO) projects, all funded by NASA.

Information Society Technologies

GRIDSTART

EGSO News

Try out EGSO
The current capabilities of EGSO can be tested, including the main GUI and SEC, SFC and DSO

April 2005
EGSO presented at the European Geophysical Union in Vienna

April 2005
EGSO presented at the UK National Astronomical Meeting in Birmingham

December 2004
EGSO demonstrated at the American Geophysical Union in San Francisco, CA

Last updated 18th September 2005

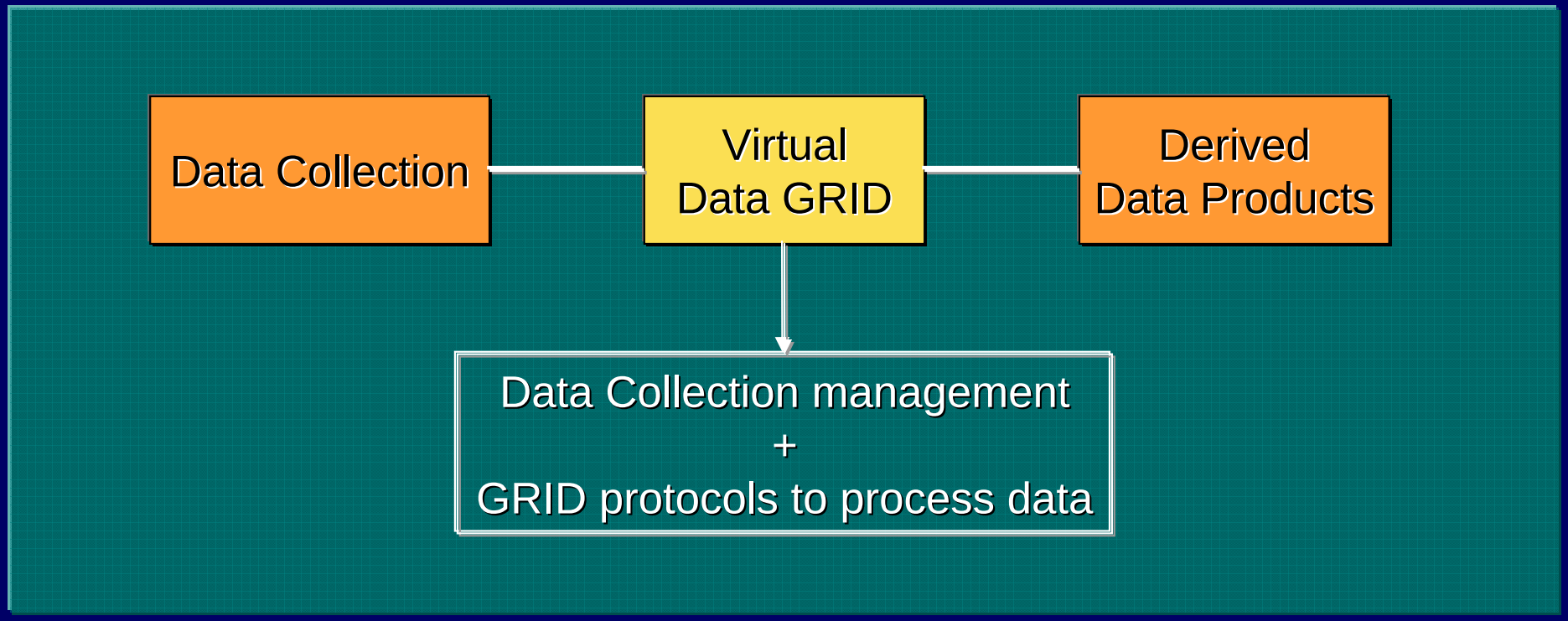
Maintained by Bob Bentley

http://www.egso.org/

VIRTUAL DATA GRID Architecture

A consistent modelling requires a multi-instrument multi-wavelength approach

DATA ANALYSIS must be provided in addition to **SEARCH and RETRIEVAL**

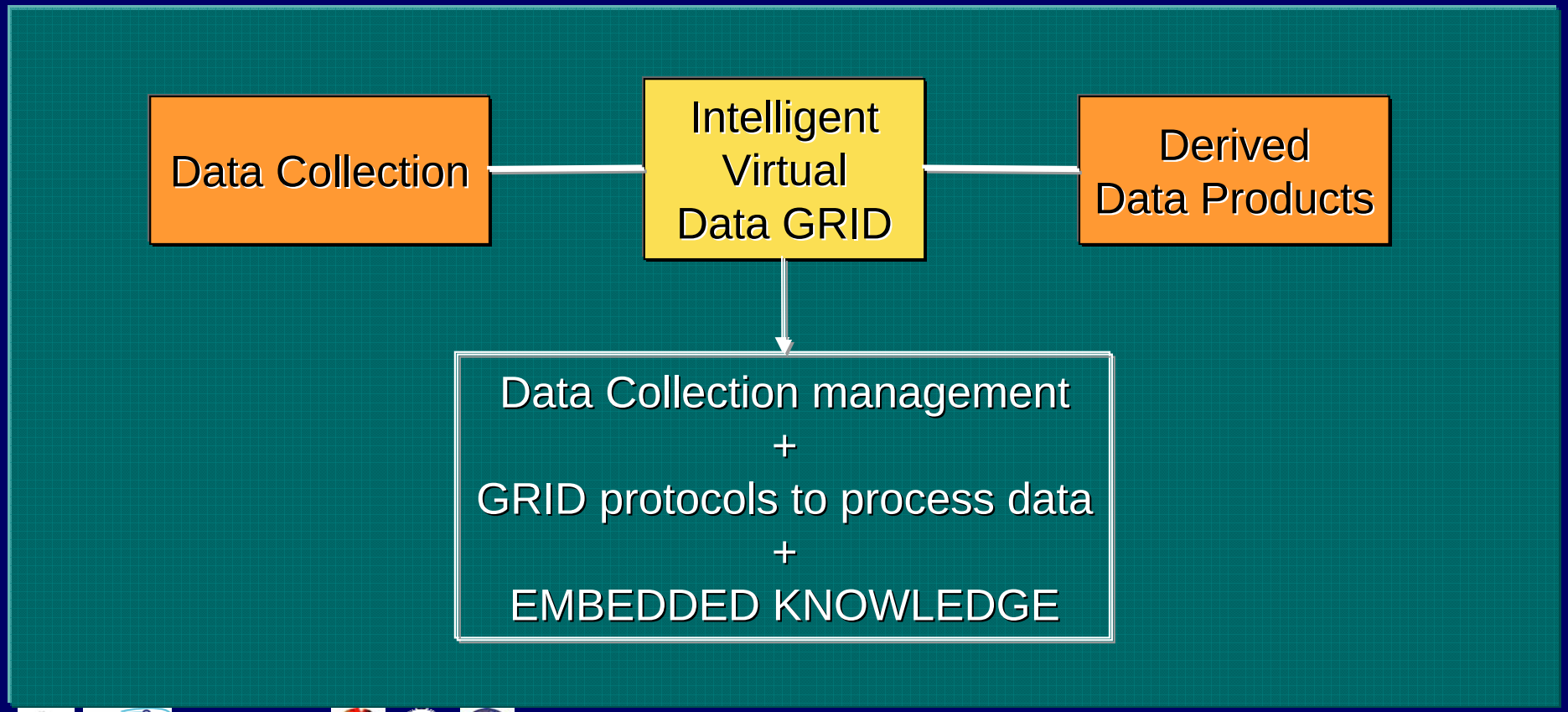


EGSO European Grid of Solar Observations

INTELLIGENT VIRTUAL DATA GRID Architecture

KNOWLEDGE DISCOVERY IN DATABASES (KDD)

DATA ASSOCIATION AND GUIDED PROCEDURES are EMBEDDED



ADVANCED GOAL

- Pointing out the physical associations in multi-wavelength datasets is the basis of interpretative scientific research
- Concept association is the kernel of knowledge
- Automated storage and search of knowledge in databases is possible through advanced techniques and is called

  Knowledge Discovery in Databases (KDD)

- Advanced techniques are based on Artificial Intelligence (AI) and Expert Systems (ES) embedding

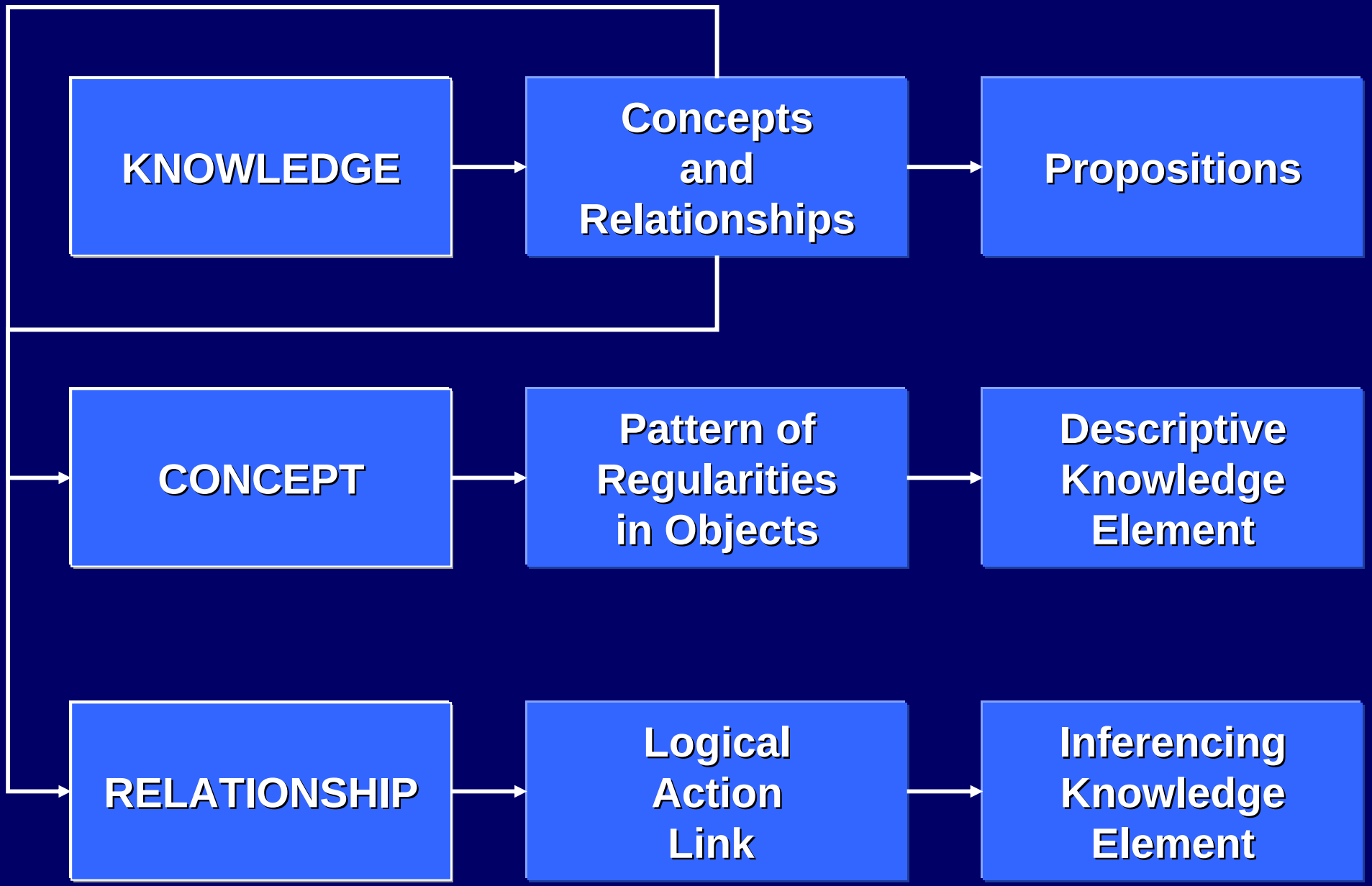
THE EMBEDDING OF AI-ES TECHNIQUES IN THE GRID ARCHITECTURE REPRESENTS THE NEXT GENERATION IN DATA SEARCH, RETRIEVAL, PROCESSING AND ANALYZING

KNOWLEDGE EMBEDDING VIA CONCEPT MAPS

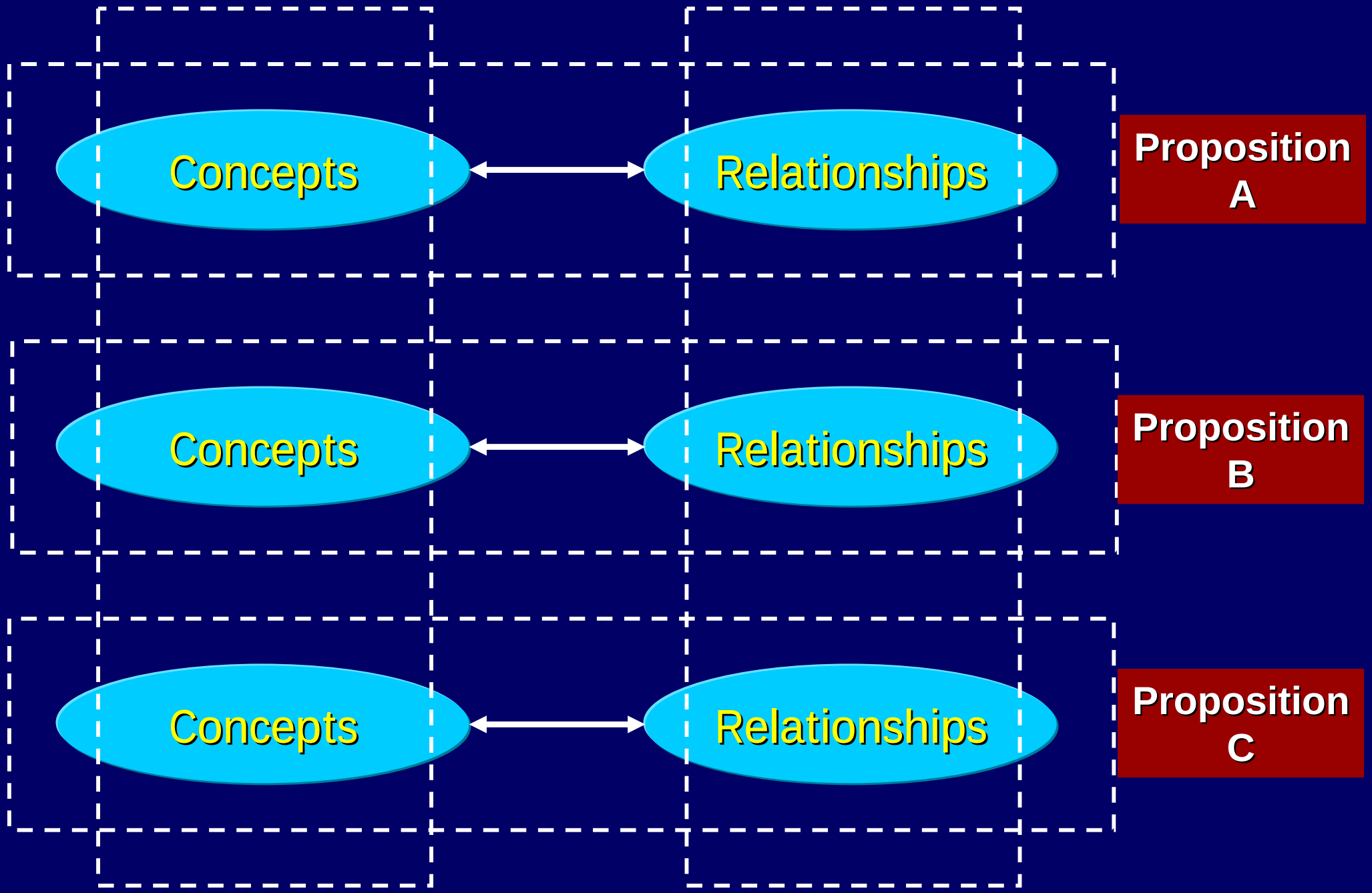
A Foundation Ontology for Space Meteorology

For a Structured Organization of the Knowledge on the Subject

A Semantic Model for Knowledge



Knowledge Representation Scheme

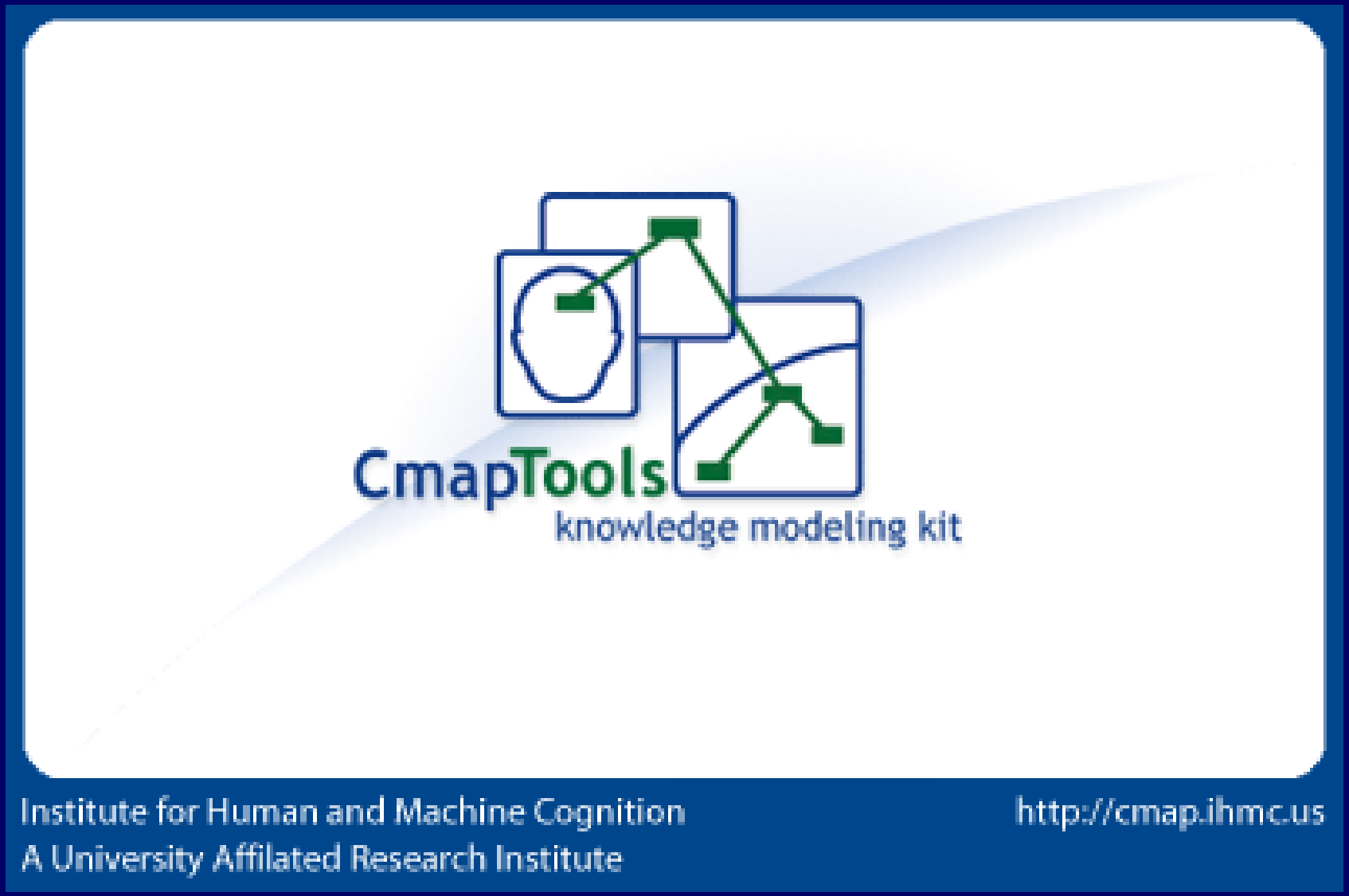


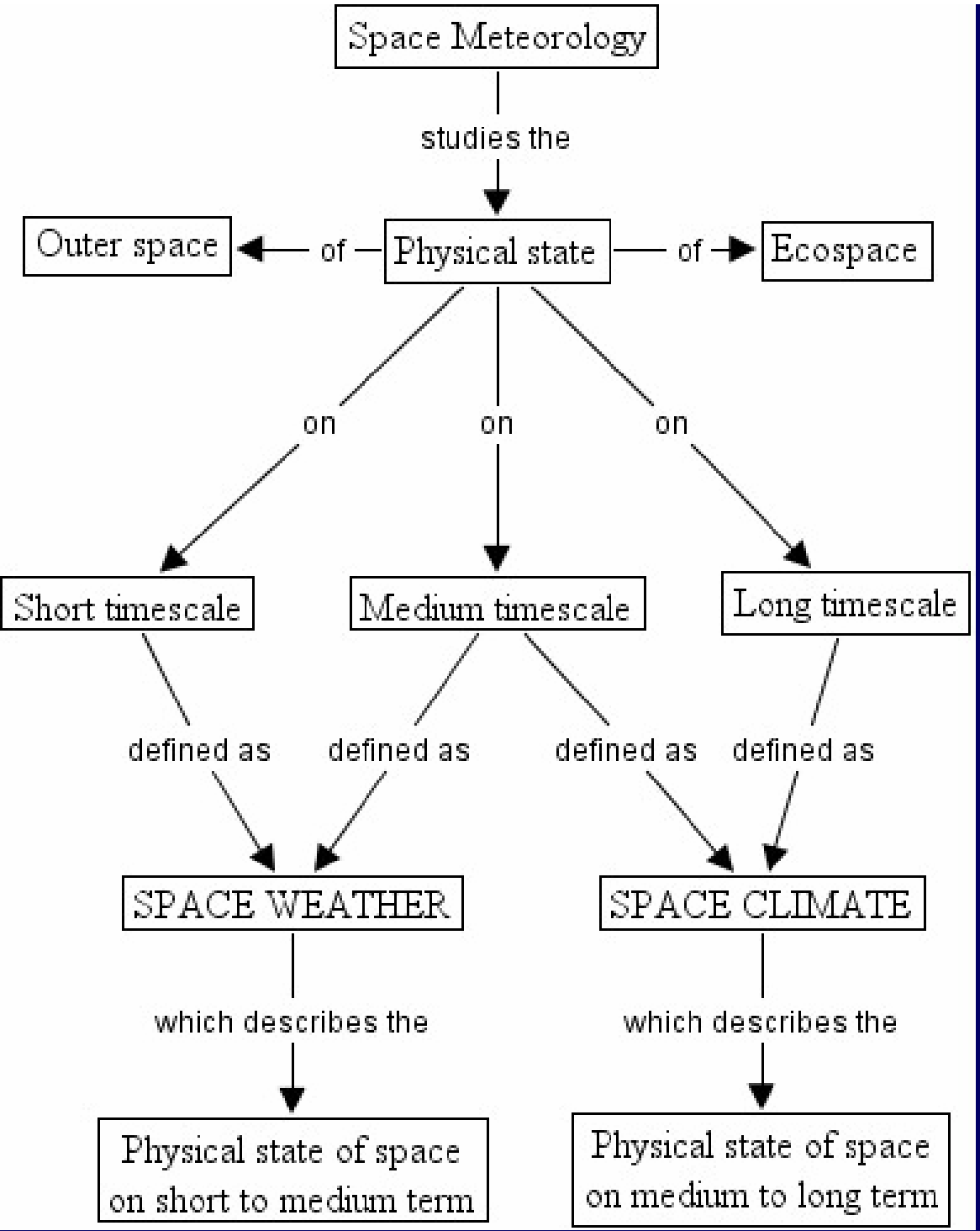
IHMC[1] Concept Map Features

- **Represent a graphical scheme of knowledge in organized form**
- Are interactively generated by means of a multi-platform software tool
- **Are implementable as XHTML/XML documents**
- External resources can be associated to concepts (e.g. scripts, hyperlinks, etc.)
- **A CXL (Connection Mapping) XL is implemented)**

[1]Institute for Human and Machine Cognition, FL, USA

IHMC Concept Map Development Tool



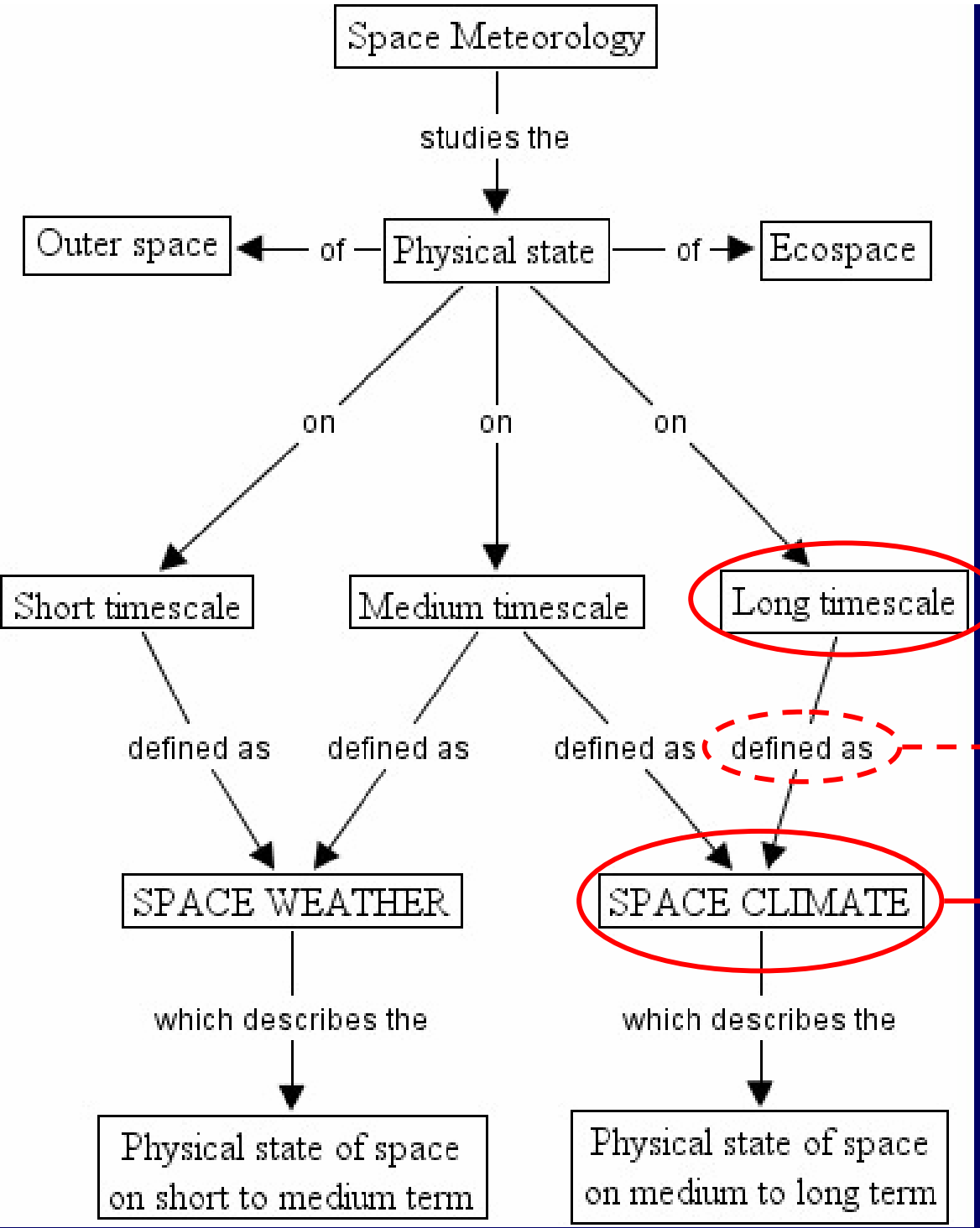


A Sample Cmap

2-D REPRESENTATION OF A SET OF CONCEPTS AND THEIR INTERRELATIONSHIPS

CmapTools knowledge modeling kit

Cmap Structure



Proposition

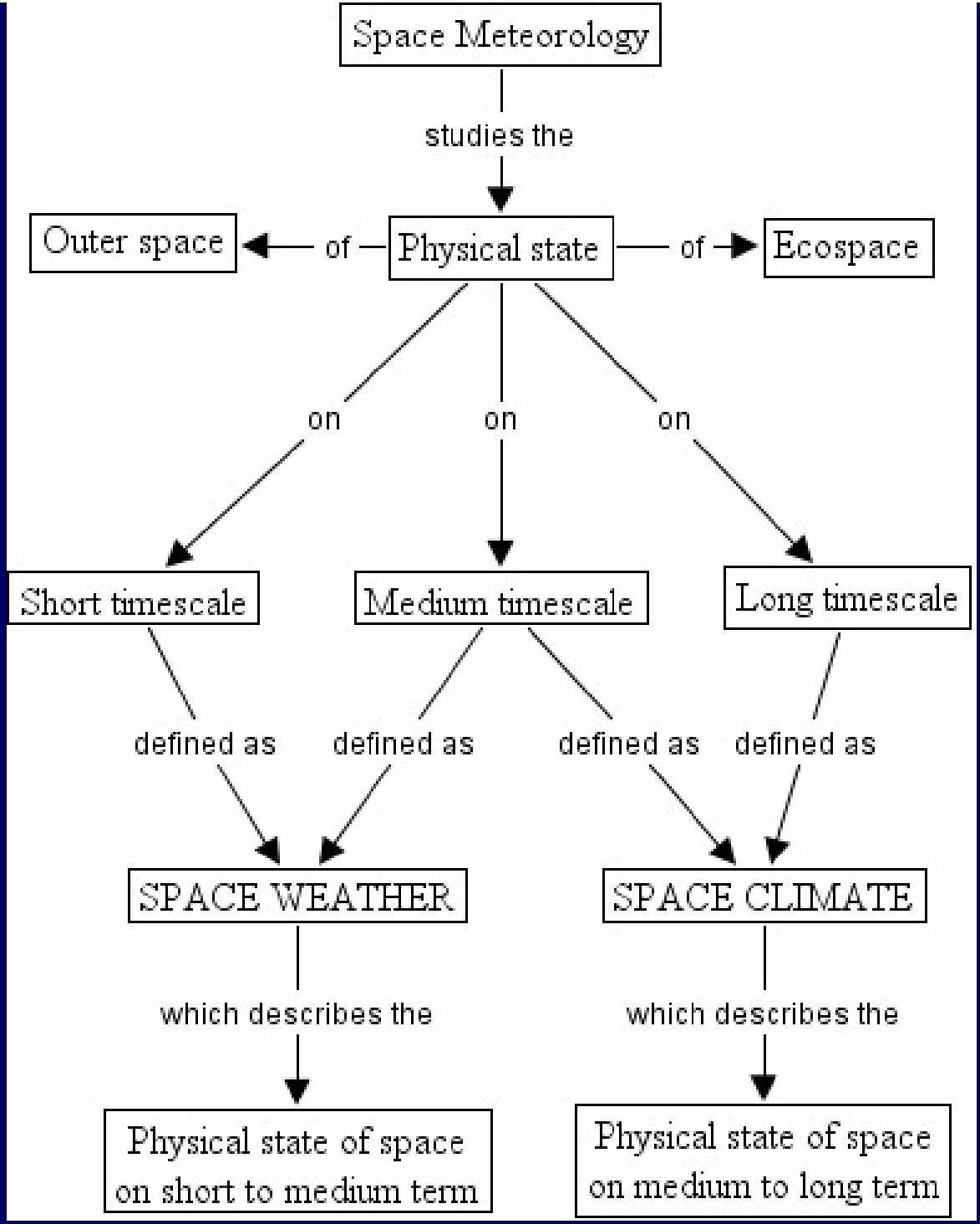
CONCEPT

RELATION

CONCEPT

KNOWLEDGE

HIERARCHY



Most inclusive

CONCEPT FRAMEWORK

Least inclusive

GENERALIZATION

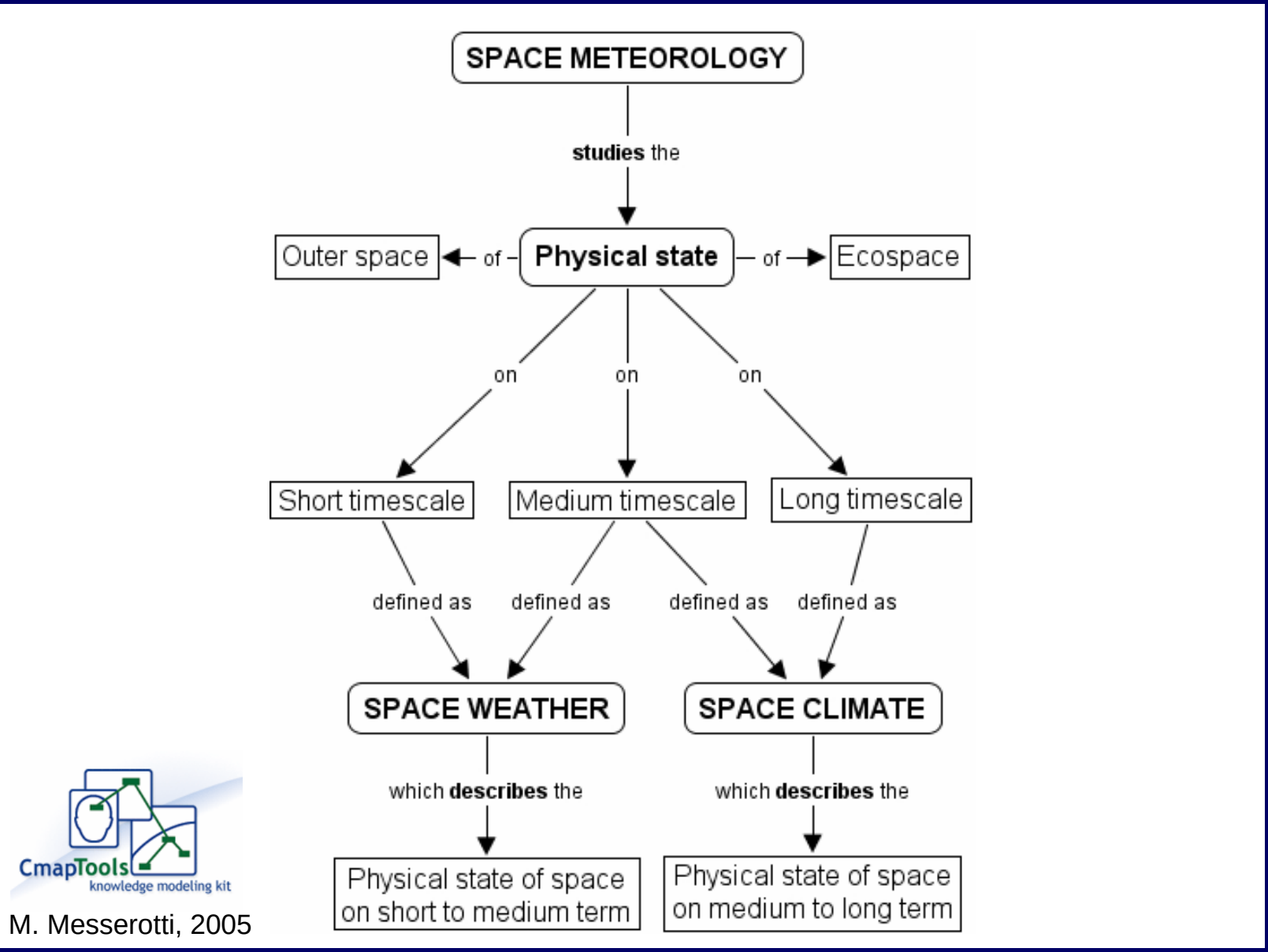
What is a Foundation Ontology?

- **ONTOLOGY** describes knowledge and it is the **formulation of a conceptual schema about a domain** constructed by:
  - Defining the precise meaning of domain entities (**semantics**)
  - Identifying the relationships between entities (**associativity**)
  - Stating the rules between entitites/set of entities (**operativity**)

Why do we need a Foundation Ontology?

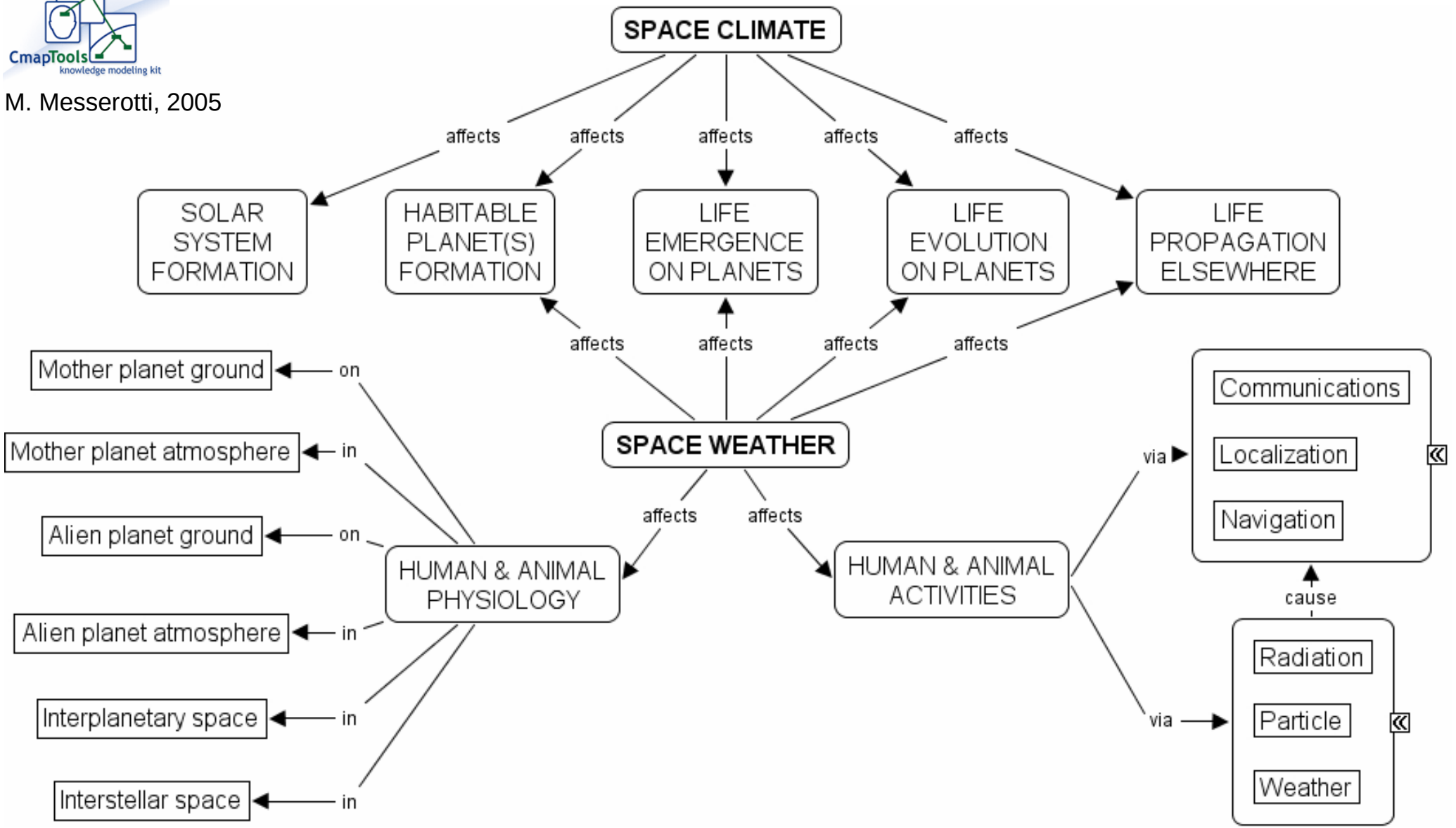
- No clear definition of the terminology
- No clear definition of the physical domains
- Interrelationships defined only on a fragmentary basis and limited to sub-domains

Definition of Space Meteorology

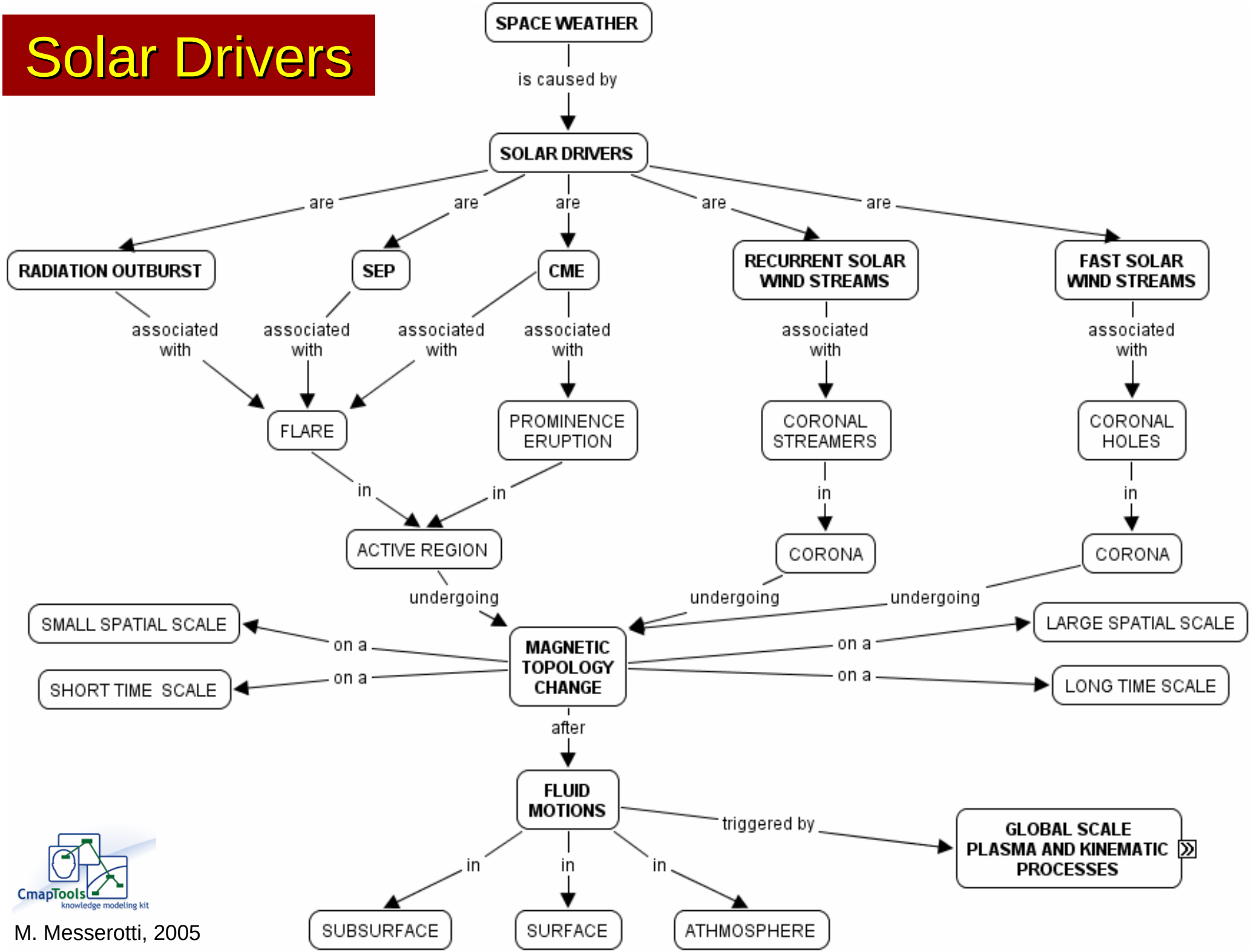


Technological and Environmental Effects

M. Messerotti, 2005



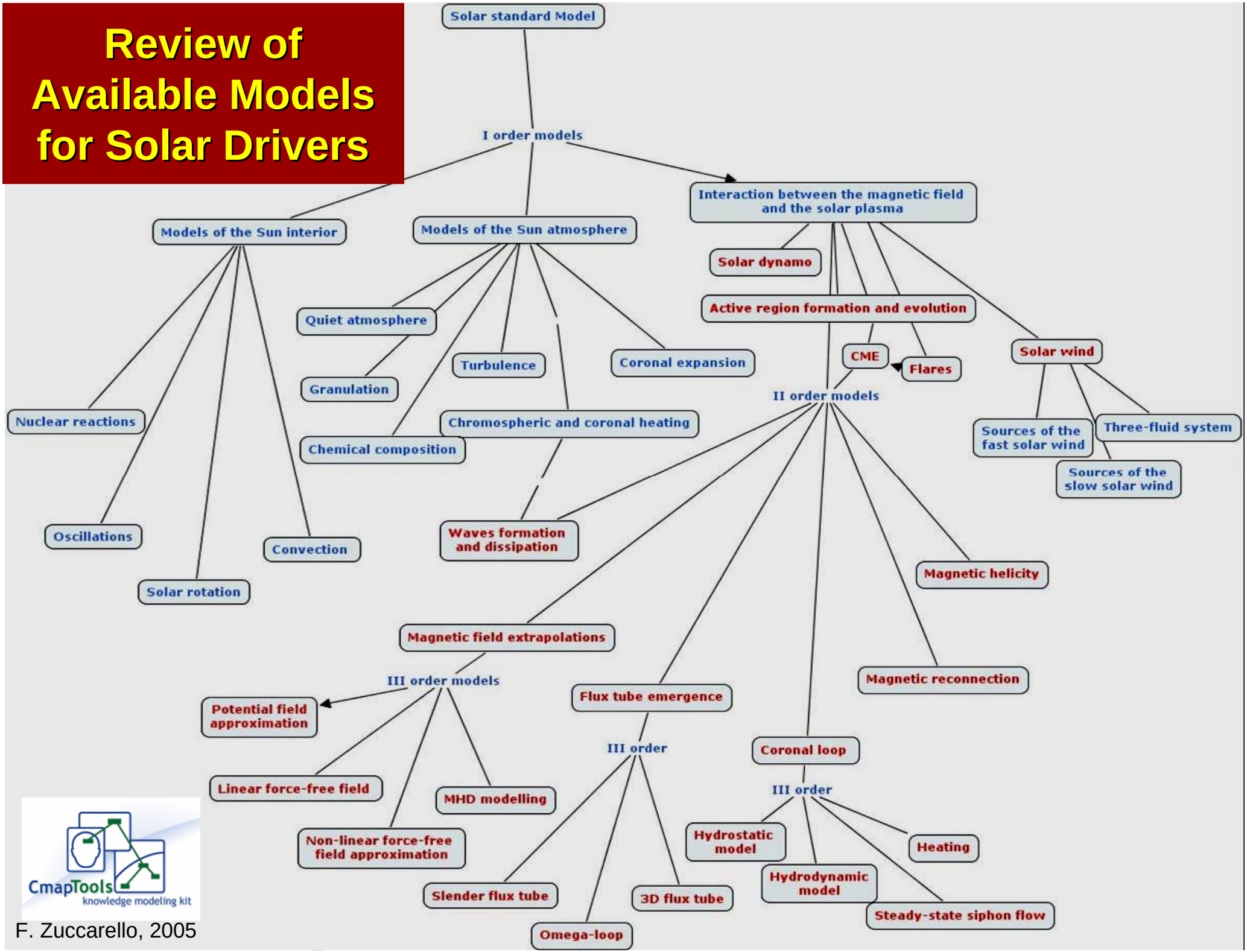
Solar Drivers



CmapTools knowledge modeling kit

M. Messerotti, 2005

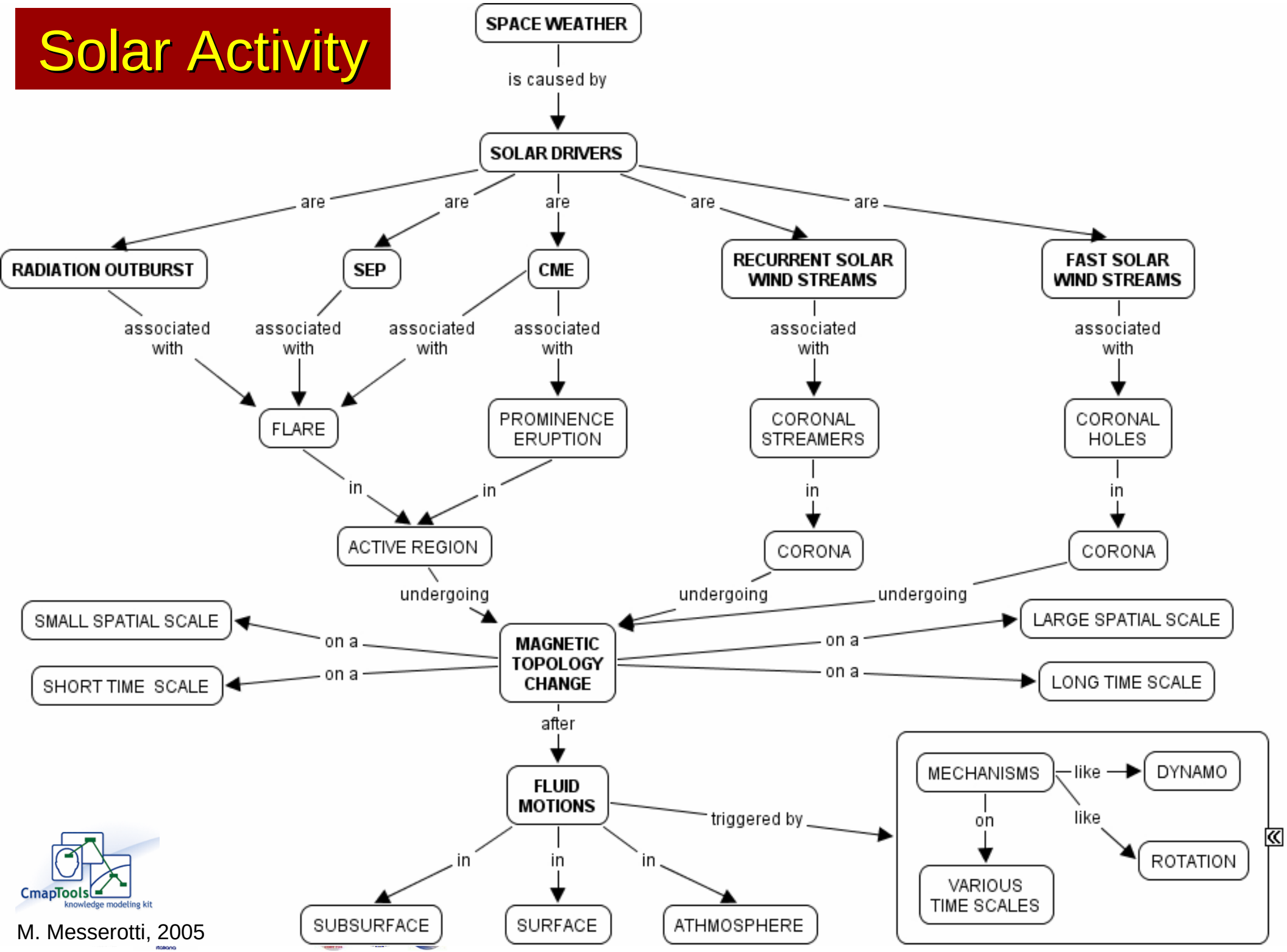
Review of Available Models for Solar Drivers



CmapTools knowledge modeling kit

F. Zuccarello, 2005

Solar Activity



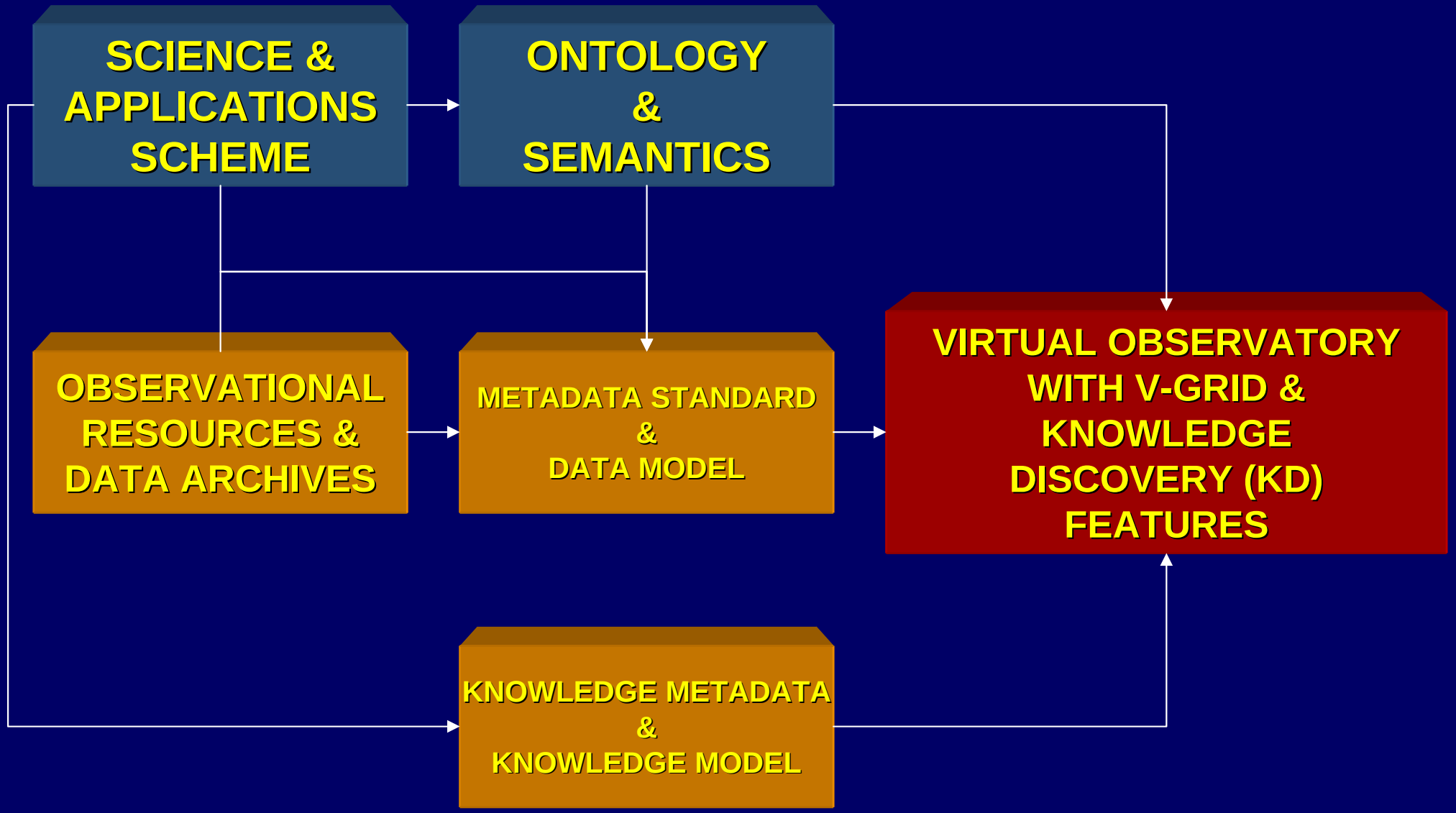
CmapTools knowledge modeling kit

M. Messerotti, 2005

VISIONS ABOUT THE IDEAL VO

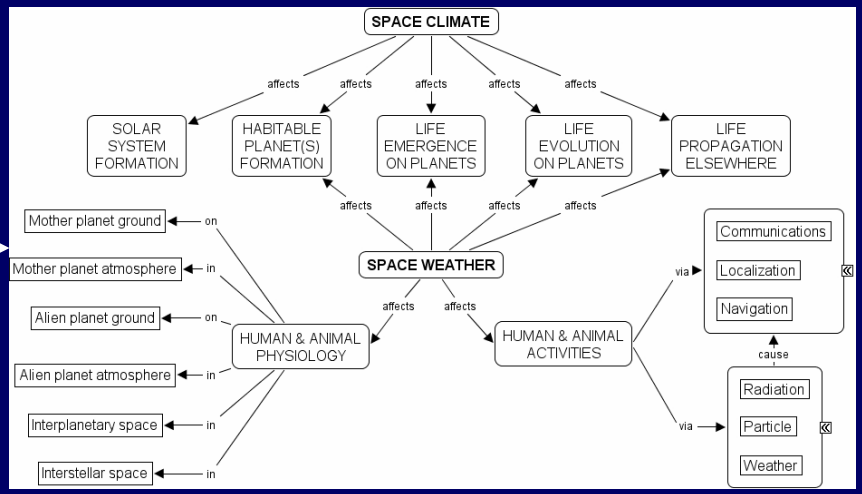
THE SEMANTIC VO
FOR
SUN-EARTH CONNECTIONS

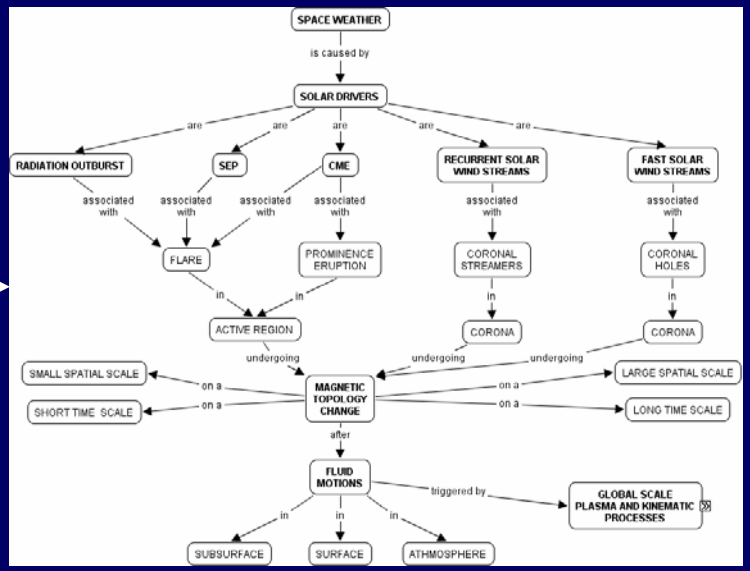
The Semantic VO



ONTOLOGY & SEMANTICS

CMAPS





Metadata Standards & Data Model

MULTI-DATA METAMODEL

Describes

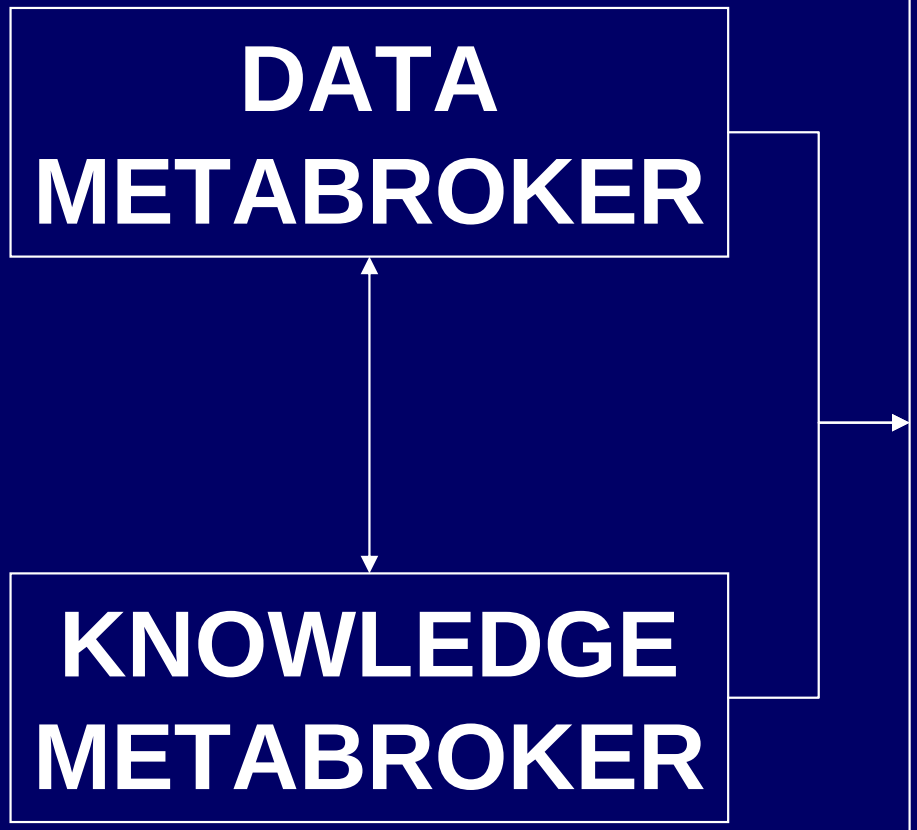
- Multi-disciplinary
- Multi-domain
  - Multi-wavelength
  - Multi-instrument

Data

Knowledge Metadata & Knowledge Model

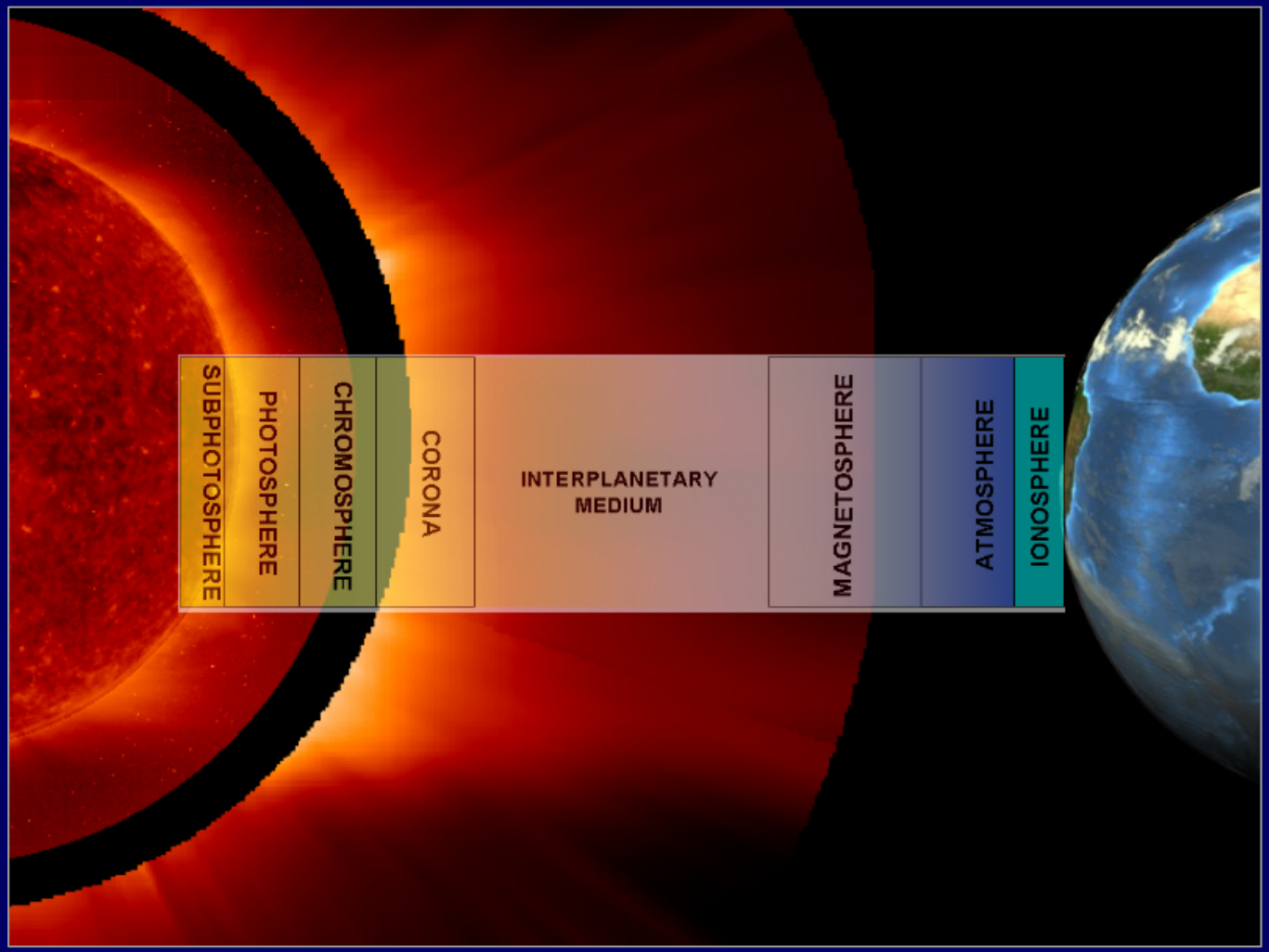
KNOWLEDGE METAMODEL → Describes
- Concepts
- Interrelationships
- Case-Based Reasoning

Advanced Brokering System



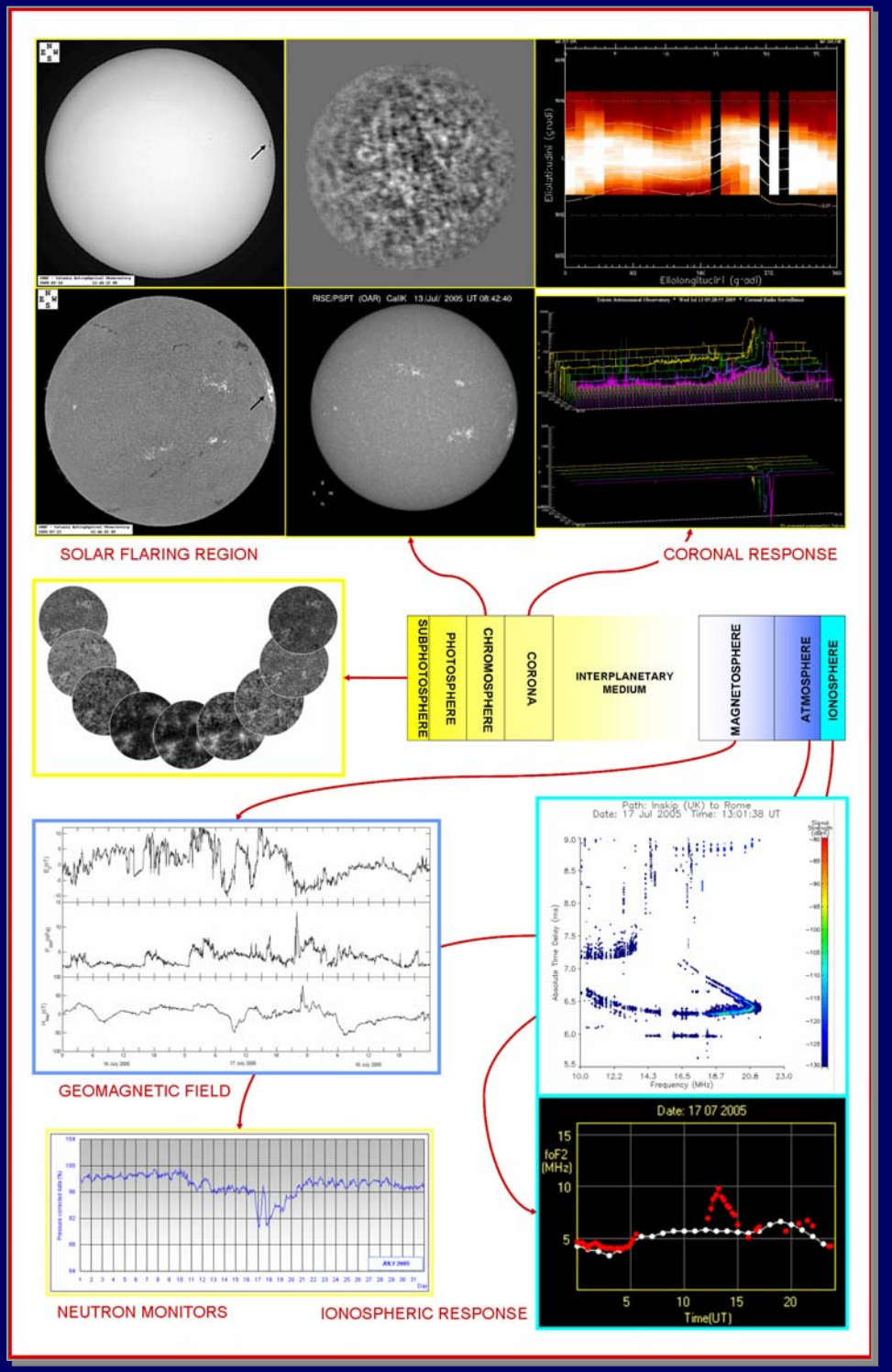
- **OWL** Web Ontology Language
  - Web Services
  - Description on Demand
- **KEA** Knowledge Exchange Architecture
  - Web Services
  - Cmaps
  - CXL Concept Mapping XL

MULTI-DOMAIN SPATIAL GUI



- **AJAX** Asynchronous Javascript & XML
  - XHTML + CSS for visuals
  - DOM Document Object Model

GRAPHICAL DISPLAY of a MULTI-DOMAIN DATA SEARCH



From IVOSEC prototyping

CONCLUSIONS

- **COMMONS TO SPACE- AND GROUND-BASED MONITORS:**
  - HUGE NUMBER OF DATA SETS
  - LARGE NUMBER OF DATA STANDARDS
  - LIMITED DATA AVAILABILITY
  - NON-REAL-TIME AVAILABILITY
  - LIMITED DATA ACCESSIBILITY
  - NON-USER-FRIENDLY SEARCH AND RETRIEVAL
  - DIFFICULT DATA CALIBRATION
  - COMPLEX DATA ANALYSIS
  - LIMITED CROSS-DATA AN.

- **POSSIBLE SOLUTIONS TO MOST ISSUES:**
  - NONE: WILL INCREASE TO PBs
  - COORDINATION ON COMMON STANDARDS
  - AGREEMENT ON DATA POLICIES
  - DEVELOPMENT OF VIRTUAL MONITORS
  - IMPROVEMENT IN WEB ACCESSIBILITY
  - ADVANCED DATA HANDLING
  - INCORPORATION OF S/W LIBRARIES
  - DEVELOPMENT OF VIRTUAL OBSERVATORIES

HS NETWORKING, HPC, I-GRID

Messerotti & Lundstedt (2004)