



Role and Importance of Natural Language in Geomatics Engineering

Dr. Dafer Ali Algarni

Associate Professor

Surveying Engineering Program

King Saud University

P. O. Box 800 Riyadh, Saudi Arabia

Dalgarni@ksu.edu.sa

Outlines



- **Objective**
- **Type of data of Geomatics**
- **Major issues**
 - ◆ **Interpretation**
 - ◆ **Integration**
 - ◆ **Automation**
 - ◆ **Education**
- **Human Language Limitations**
 - ◆ **General Limitations**
 - ◆ **Specific Limitations**
- **Conclusion**

OBJECTIVE



Addressing the role and importance of natural language in Geomatics to prove the need of incorporating some linguistic courses into the discipline, and to see how far can we go with automation.

Feature, Object, Entity, Concept, Hypothesis, etc



- Spatial data
- Quantitative
- Location
- Graphic
- etc

- Non-spatial data
- Qualitative
- Non-location
- Non-graphic
- Attribute
- Textual
- Description
- Footprint
- Metadata
- Semantic
- etc

WHERE

WHAT

WHAT is WHERE

Notice



- **Machines and humans seem to be opposite to each other in perceiving and handling these two types of data.**
- **Machine handles the numerical and graphical data efficiently, but faces obstacles with semantic data. Human mind faces no difficulties in understanding semantic information, and finds it difficult to deal with numerical and graphical data with the same degree of easiness and efficiency.**
- **Scientists are being affected by technology and are willingly concentrating on the “where” part in which the machine is predominant. The “what” part is ignored.**

Major issues



- **Interpretation**
- **Integration**
- **Automation**
- **Education**

Interpretation



- **Interpretation is the process of identifying image features and communicating the findings with other users.**
- **Interpretation depends on natural language, resolution and scale.**
- **Resolution and scale seem not to pose any serious problem.**

Characteristics of some selective sensors/satellites.



Type	Coverage (km ²)	Spectral Bands	Radiometric Resolution	Temporal Resolution (days)	Spatial Resolution (m)
MSS	185X185	4	$2^8 = 256$	16	80
TM	185X185	7	$2^8 = 256$	16	30
SPOT	60X60	3 color, 1 B/W	$2^8 = 256$	26	20 color, 10 B/W
IRS-ID	141X141 color, 70X70 B/W	3 color, 1 B/W	$2^8 = 256$	24 color, 12 B/W	23.5 color, 5.8 B/W
IKONOS	11X11	4 color, 1 B/W	$2^{11} = 2028$	1.5	1

Integration



- **Specialization goes in accordance with human capabilities and performance. However, specialization proves to us that all types of knowledge are integrated.**
- **Integration of spatial data is possible, but what about the non-spatial data.**
- **Homogeneity in spatial data implies some sort of homogeneity in terms and vocabulary that describe such data.**

Lack of homogeneity



- **Different terms of the same functions and concepts (not multi-names). Ex. area-based and Feature-based in digital photogrammetry vs field-based and entity-based in computer vision which lead to raster and vector GIS.**
- **Abbreviations: GPS (global positioning system or global problem solving)**
- **The same term for different concepts: example, Generalization in mapping & data-bases.**

Automation



- **Automation is based on integration.**
- **Automation in the “where” part is good.**
- **Automation in the “what” part is poor, unless it is processed separately.**
- **Both parts have to be processed together.**
- **Automation obsession**

Education



- **Concentration in geomatics has been on the “where” part.**
- **The quality of education is affected by ignoring the “what” part.**
- **Role of language in education is increasing due to changes in education approaches (internet).**
- **Through education changes could occur.**

“What is where” status wrt the four issues.



Issue	Objective	
	What	Where
Characteristic	Attributes, non-spatial, footprints, qualitative, etc.	Feature, entity, concept, object, quantitative, etc.
Treatment	Semantic Models	Mathematical Models
Interpretation	Poor	Good
Integration	Poor	Good
Automation	Not Automated	Almost Automated
Education	Not emphasized	Emphasized

Language limitations



- **General limitations**

- ◆ **Environment – village experience.**

- ◆ **Effect of previous knowledge.**

- ◆ **Machine limitations**

- **Words that are nouns, verbs, and adjectives.**

- **Words with the same spelling or pronunciation**

- **Long words: deinternationalization.**



Language limitations

- **Specific limitations**
 - ◆ **Lack of conciseness.**
 - ◆ **Lack of names of features:**
 - Problem of sub-feature.
 - Long names: Ex. RealStateObject, NaturalTransportationLink, etc.
 - ◆ **Sense and pronunciation of terms: Ex.**
 - Pixel, voxel, groundel, grxel.

Conclusion



- ◆ Geomatics is built of two equally important types of data that can be summarized by the “what is where” slogan.
- ◆ The concentration has been on the “where” part, the “what” part is almost abandoned.
- ◆ Inappropriate separation between these two parts is motivated by separating knowledge to art and science.

Conclusion (Cont.)



- ◆ Knowledge of geomatics can be increased by incorporating courses from relative fields such as computational linguistics, cognitive science, geomorphology, ontology, etc.
- ◆ Language is the indicator of human power (or strength) and weakness at the same time. If this is true, automation will always be amazing but with some sort of deficiency.



THANK YOU