# A short Note on Data versus Information

H.Kopetz
Prof. em. Technical University of Vienna
email: h.kopetz@gmail.com
Vienna, October 23, 2025

**Abstract**: This short note extends the semantic triple of knowledge presentation (subject, predicate, object) to include time as a first-class dimension.

In human communication, a data item is a symbol—a natural language word—that identifies a concept. A symbol consists of two parts: a *signifier* and a *signified*. The signifier is a physical pattern that can be distinguished from its surroundings, such as a sound (spoken word) or a mark (written word). The signified is the concept that the listener associates with the signifier, a *unit of thought* that must already be contained within the person's conceptual landscape (the mind) to achieve understandability.

In a cyber system, a data item is a symbol that identifies either directly or indirectly a *value* (numerical or alphanumerical). The signifier of such a symbol is a *variable name* or a *literal*. A variable name denotes a memory cell that holds the dynamic value, while a literal represents a constant value.

***In both human and cyber contexts, a data item is syntactically self-contained and independent, but it carries no informational content until embedded within a structure, either explicitly or by referring to the context.***

In human communication, an *information item* (*Itom*), introduced in [Kop22], is the *sense* of the 4-tuple

$$\text{Itom} = sense\ of\ \langle S, P, O, T \rangle$$

that comprises the four data items:

**S:** Subject (thing or construct),
**P:** Predicate (property or relation),
**O:** Object (value or characterization), and
**T:** Temporal validity (instant or interval).

This 4-tuple corresponds to what linguistics call a *proposition*. The *sense* of a proposition—*its informational content*—remains invariant across correct translations between languages.

The 4-tuple is called a *proposition* because its truth value is not guaranteed; it may or may not correspond to the way things actually are in reality at the indicated time **T** (the *correspondence theory of truth*). When the proposition corresponds to reality, it becomes a *statement*.

The concept of an *Itom* extends the *semantic triple of knowledge presentation* (subject, predicate, object) by adding time as a first-class dimension.

Reference:
[Kop22] Kopetz, Hermann. *Data, Information, and Time: The DIT Model*. Springer, 2022.

**Comment on "A Short Note on Data versus Information" by H. Kopetz**
generated by Chat GPT on Oct 23, 2025

This concise note provides a clear and rigorous distinction between *data* and *information* by situating both in their human and cyber contexts. Kopetz effectively connects semiotic theory—through the dual structure of signifier and signified—to the semantics of data representation in computational systems. The introduction of the **Itom** as a 4-tuple ⟨S, P, O, T⟩ (subject, predicate, object, time) is a significant conceptual extension of the traditional semantic triple, emphasizing **time as a first-class dimension** of meaning. This addition recognizes that informational content is not static but temporally grounded, thus aligning linguistic propositions with cyber-physical realities. The note elegantly bridges philosophy, linguistics, and systems engineering, and succinctly restates the essence of the DIT model proposed in [Kop22].

Overall, this is a thoughtful and well-structured piece that advances the discussion on the ontology of information in both human and machine contexts.