

Globally predefined and defined information

Every digital information is a number sequence. The numbers select from a set of possibilities resp. **domain**.

Definition of information:

Information = selection from a set or **domain**.

Preconditions for precise transfer of information are:

- (1) **Well predefined domain** (a priori, for all participants of conversation)
- (2) **Ordered domain** (so that its elements are selectable by numbers)
- (3) **Transfer of the number sequence which shows the selection in the domain**

Up to now (2018) the domain is predefined locally (by context). The internet makes much more possible:

We can predefine the **domain globally!**

Information is predefined after predefinition of its domain.

→ The new purposefully unifying information carrier ("Domain Vector" resp. DV) has the form:

UL (of online predefinition) plus number sequence

In the DV information is **predefined (by UL)** and **defined (by numbers which select)**

UL = "Uniform Locator" on the internet (used for globally unique identification and predefinition),
- can be the **URL** or an **abbreviated equivalent**, e.g. a short numeric pointer (e.g. to a local URL table),
- identifies the kind of the following **number sequence** (like a file ending can identify the kind of data in the file),
- is also pointer to the predefinition of the number sequence, so it **predefines the domain** of a selection (globally).
The **number sequence** after the UL defines the **selection** in the domain and **represents the variables (data)**.
The online predefinition of the domain is located (globally uniquely) by the UL. The online predefinition contains a bijective mapping from every number sequence to every element of the predefined domain. Therefore further definitions (metadata) mixed with the **data (number sequence)** are not necessary. The central (globally uniform) predefinition of the domain allows (as additional "bonus") **similarity search**. The domain can be predefined as multidimensional metric space. It is called "Domain Space" (DS).
Nomenclature: Every DV (resp. Domain Vector) is element of a DS (Domain Space).

The DV: **UL (of predefinition) plus number sequence**

- enables the combination:
 - **Maximal competence (predefinition by all internet users)**
 - **Maximal efficiency (number sequence)** (allows maximal entropy)
- can represent every definable information - **precisely** from **simple** word to **complex** multidimensional information e.g. in science, medicine, industry.
- is internationally uniform and comparable for similarity search. The **users create** the search criteria.
- **<http://numericsearch.com>** demonstrates searchability.
- is **language independent** (if translation of predefinition is possible) so that the online predefinition can be **multilingual**.

Existing online predefinitions can be reused (linked together and recombined) in new predefinitions. For search (and before providing data) users can select the best (most meaningful) predefinitions. This motivates to create better and better predefinitions.

Details: <https://arxiv.org/abs/1801.03106>

Contact: Wolfgang Orthuber, Kiel University, Germany

Email: orthuber@kfo-zmk.uni-kiel.de