How to communicate generative AI without generating false beliefs

by Giovanni Acerboni, June 28th 2024

Nicholas Belmore, a researcher at Harvard, asked me in a LinkedIn conversation what I think about finding a language for generative AI that is both reliable and accessible to non-experts. According to Belmore, many expressions used by data scientists create confusion among the general public. For instance, intelligence (artificial), reasoning, hallucinations, etc.

I told him that it would be difficult to adequately respond in a post and I committed to presenting my ideas in a more structured document. This is it.

1. New Words

When someone invents or discovers or creates something, they name it by creating a neologism. Neologisms are created in various ways, but what is relevant here is the **semantic neologism** (or resemanticization or redetermination), which consists of **attributing a new meaning to an existing word**.

L'INK SCRITTURA PROFESSIONALE S.a.s.
via Biella 24/A - 20143 Milano
telefono: ++39 02 36 53 48 71
fax: ++39 02 99 98 18 14
e-mail: info@scritturaprofessionale.it
web: www.scritturaprofessionale.it
Partita IVA e Codice Fiscale: 04305670962

Very often, the new meaning has some semantic relationship with the existing word. Some examples from the digital, hardware, and software fields: web, (web)page, home(page), (web)site, mouse, application, program.

The meaning of these **neologisms can cause confusion**, especially in the early days of their appearance. In fact, when the readers encounter '(web)page' for the first time, they immediately see that there are some analogies but also many differences with the meaning they are familiar with. For example, web pages are not numbered, do not appear one after the other in a bound sequence, can be modified and even removed.

When neologisms concern THINGS, whether virtual (page) or tangible (mouse), misunderstandings last a short time, new meanings stabilize, and no one gets confused anymore.

The situation is very different when we move **from THINGS to ABSTRACT CONCEPTS** such as, in our field, intelligence, reasoning, creativity, knowledge, consciousness, argumentation, persuasion, manipulation, etc.

Humanity has been questioning the meaning of these concepts for a couple of thousand years: philosophers, theologians, logicians, biologists, neurologists, chemists, cognitive scientists, linguists, etc.

A unique and shared meaning of such terms is rare, and in any case, research is always ongoing and continuously pushes the boundaries of knowledge. As a result, **terms like these have unstable meanings** which, within a specific specialized discourse, are well-defined and therefore clear to experts.

Non-experts, on the other hand, use these terms loosely, attributing a generic meaning valid for non-technical conversations where terminological errors do not matter much and misunderstandings are easily overcome.

However, whenever a non-expert encounters one of these terms in a specialized discourse, they cannot ponder how much the familiar meaning differs from the meaning attributed by the specialist, unless the specialist explains it very clearly (which is not easy and rarely happens).

In communication with non-experts, **the meaning of these terms is thus very often indeterminate and consequently ambiguous**, meaning it can be interpreted in various ways.

The effect of ambiguous communication is always negative. The recipient may:

- not understand
- understand something else
- perform incorrect behavior
- spread incorrect information
- doubt their own skills and knowledge
- doubt the reliability of the person speaking to them, even considering them hypocritical.

2. A Language for Generative AI

There is an original sin: Artificial Intelligence defined itself in 1956 with a term (Intelligence) of indeterminate meaning. The addition of Artificial narrows the scope of the discourse but does not remove the indeterminacy.

This likely happened due to the hypothesis that a machine could think and act like a human. Much AI research attempts to reproduce the mechanism of human reasoning and behavior. For example, the term Neural Network derives from this approach.

In any case, **it is pointless to oppose terms that have now stabilized**, such as Artificial Intelligence. This expression cannot be avoided.

It's a different story for terms that do not designate stabilized technical concepts but derive from belonging to the same functional and semantic area (in traditional meanings), such as reasoning, creativity, etc., which are often used in quotation marks (those who use them know and want to make it known that they are using them improperly, as a shortcut to easily understand each other).

All non-technical terms can avoid stabilization: just replace them with better ones.

However, as long as the discourse remains confined to experts, the issue would not even arise. The issue becomes **extremely critical when the discourse reaches non-experts**.

This is the case with generative AI, especially those of language, because they have taken away the human monopoly on the effective use of language. This is surprising, especially for non-experts in AI.

Non-experts in AI, however, know what language, communication, and writing are. They know that **language expresses always content** (despite Chomsky's opinion), and the choice of style functionally connotes that content. **Language is therefore the communicative product of reasoning, i.e., it expresses the intention to communicate something to someone**.

Using a software that writes, non-experts wonder how it does it. If they are answered with terms that apply to human communication, a great mess is made. Because **software is software, it has no intentions**, therefore it cannot conduct reasoning on what content to communicate and how. To write, **it follows other logics. To explain thme, metaphors are real traps**.

3. Technical standards

There are rules for precise and comprehensible communication of technical discourse. Long story short:

- **synthesis**: communicate only the necessary information, that is, useful to understand and usable to perform a behavior
- **clearness**: construct short sentences and use common and at the same time precise terms. For unavoidable technicalities, provide a translation into common terms.

These rules have been formalized in technical legislation.

In Italy, since 2013, we have Standard 11482:2013 *Structural elements and linguistic aspects of written communications of organizations* issued by UNI, the national branch of ISO.

ISO issued Standard 24495-1 *Plain language. Part 1: governing principles and guidelines* in 2023 and is developing *Part 2: Legal communication* and *Part 3: Science writing*.

4. Well-Tempered Cyber Reason

It is not always easy for a specialist to achieve an effective style for communication with non-experts. Specialists:

- fear of saying something different if they change style (a legitimate concern)
- wrongly believe that their content cannot be communicated differently
- struggle to get used to a different style

It always **depends on the purpose and motivations** (personal or corporate). Those who want, can. Especially since Artificial Intelligences support simplification. **Natural Language Processing works excellently** because it automates the identification of sentences and words to simplify. For the Italian language, I had created one (in collaboration with Alessandro Panunzi from the University of Florence), with patented algorithms used in machine learning and a database with thousands of terms. But it is currently not operational. There are others for other languages.

As for **generative AI, they are not very good at recognizing complex linguistic facts**, but it is undeniable that they write quite clearly, except for content errors, which are quite frequent and serious the more specialized the discourse.

I wrote "error." **Errors are errors, not mistakes, and certainly not hallucinations**. If there is an error downstream, there is an error upstream. It may be counterintuitive, but error **correction is done by reducing the dataset, not increasing it**.

A course in philology and in pragmatic would be useful in STEM degrees and companies that do machine learning.