Wiktor Rorot
Institute of Philosophy and Sociology, Polish Academy of Sciences &
Konrad Lorenz Institute for Evolution and Cognition Research

Concepts & texts in the practice of life science

Understanding the concept of “signaling” in life science with topic modeling

Introduction

Digital philosophy of science (e.g., Pence and Ramsey 2018; Pence 2022) harnesses computational methods of data mining for the study of digital outputs of science. While it allows for the analysis of large datasets, which go far beyond what can possibly be read during a human lifetime, the focus of this approach on the study of published papers has raised concern about the ability of such methods to track the actual practice of doing science. The issue stems from the fact that publications often offer an imperfect representation of how science is actually conducted in the laboratory or in the field. The literature tends to abstract away numerous vital details of the scientific process (see, for instance, Rouse 1990; Schickore 2008). Lean and colleagues (2021) call this a “justificatory gap” when attempting to draw conclusions regarding scientific practice from the study of published outputs of science.

Cognitive metascience bridging the gap

Recently, Miłkowski (2023) proposed to view theories and other products of scientific practice as “cognitive artifacts”, entities that are used to “maintain, display, or operate upon information in order to serve a representational function and that affect human cognitive performance” (Norman 1991, 11; cf. Callebaut 2013). Such cognitive artifacts of science include publications. Regardless of possible distortions they introduce, publications scaffold existing practices of science by assuming multiple distinct roles: offering justification (e.g., Reichenbach 1938), recruiting readers to the presented research programs (e.g., Knorr-Cetina 1981), or constructing narratives that make sense of scientists’ activities (e.g., Rouse 1990), among others. This parallels the individual ways in which, as Lean and colleagues suggest, the justificatory gap can be bridged. However, appreciating the diversity of those roles and the publications’ impact on science as a cognitive practice, requires appropriate methodologies.

Diffractive methodologies for Digital Philosophy of Science

An important benefit of the view of outputs of scientific practice as cognitive artifacts is that it underscores their active and constructive role in relation to the practices that have created them. Rather than being a simple passive representation, publications serve to reinforce particular values, highlight some of the aspects of scientific investigation, while hiding others (cf. tacit/explicit knowledge distinction in STS, Lynch 1988; and the concept of epistemic by-products, N. C. Nelson 2018; N. Nelson 2020). This distinction parallels Barad’s criticisms of reflexive methodologies, which partially stem from the observation that reflexivity assumes “that practices of representing have no effect on the objects of investigation” (Barad 2007, 87). As an alternative, Barad proposes “diffractive” methodologies, which center patterns of difference, eschew passive representationalism, and appreciate that “practices of knowing are specific material engagements that participate in (re)configuring the world” (Barad 2007, 91). As such, publications may serve as “diffraction gratings”, rendering obsolete the subject-object (publication-practice) distinction at the heart of the “justificatory gap”. In particular, investigating the use of scientific concepts in texts, and how this differs across contexts and research fields, highlights theoretical practices at play in scientific research. Below I propose a particular implementation of this methodology, centering on the use of the term “signaling” in life sciences.

For the PDF of the poster & Supplemental Information use the QR code or go to: https://wiktor.rorot.pl/blog/concepts-and-texts.html

You can contact me at:
w.rorot@uw.edu.pl

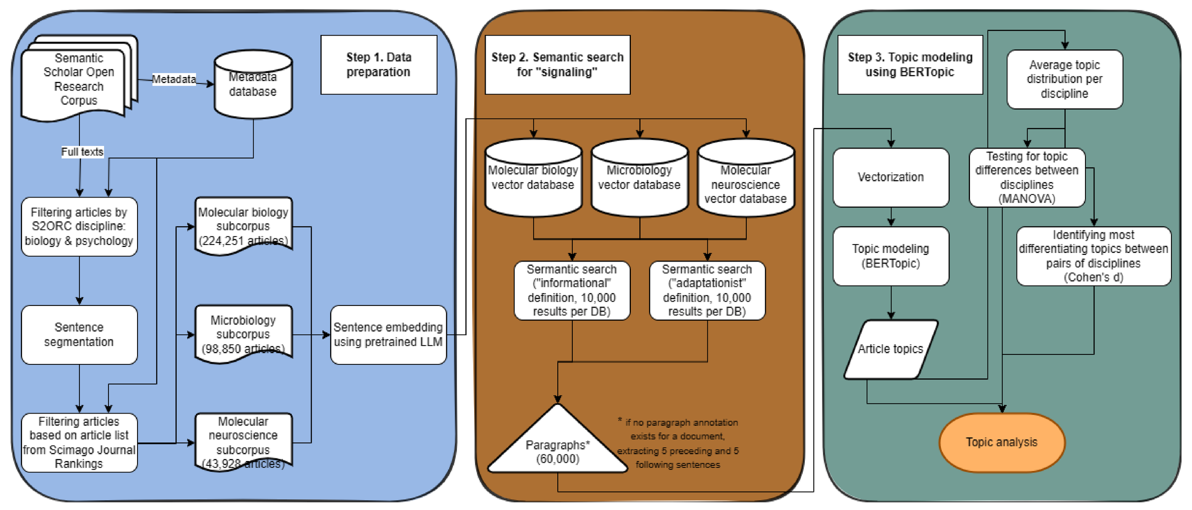

Figure 1. Research design.

Box 1. Adaptationist definition of signaling

“A signal is any act or structure that affects the behaviour of other organisms, evolved because of those effects; and which is effective because the effect (the response) has evolved to be affected by the act or structure.” (Scott-Phillips 2008, 388)

Identified mentions (manually picked from top 10):
A-CMN1. “A signal can be defined as a stimulus which carries a message to a receiver.” (Corpus ID: 18782890)
A-CMN2. “In this paper, we define ‘signal’ as the part of the observations that can be well predicted from the past history of the time series, and ‘noise’ as what is completely unpredictable and produces the variation across realizations.” (Corpus ID: 8027578)
A-CMN3. “Typically, a signal is sensed at a specific point of the network (input) and is propagated to modulate the activity or abundance of other network components (output).” (Corpus ID: 15079581)

A-MoB1. “According to Venturi and Keel [1], a signal is a mobile compound whose occurrence leads to one or more cellular responses by the receiving cell that are not limited to catabolism, transformation, or other aspects of this compound (e.g., resistance to its toxicity).” (Corpus ID: 269894057)
A-MoB2. “It was deemed reasonable to define a signal as a message that may be generated by any interaction between a cell and its environment, such as the engagement of a membrane receptor by a specific ligand molecule.” (Corpus ID: 256260998)
A-MoB3. “The signal is a measure of the difference between patterns of activity between conditions.” (Corpus ID: 208563566)

A-MiB1. “On the other hand, a signal is a substance, produced at a lower cost by a sender, to communicate information for the benefit of both the sender and the receiver of the signal (Maynard-Smith et al., 2003).” (Corpus ID: 246829152)
A-MiB2. “The main distinction with a signal is that the biological response did not evolve for that purpose, which benefits only the receiver (Keller and Surette, 2006; Diggle et al., 2007; Stacy et al., 2012).” (Corpus ID: 8536262)
A-MiB3. “Molecules are generally classed as ‘signals’ if they are produced by a dedicated pathway at a specific stage of growth, and elicit responses in the receiver organism that are distinct from those required for the processing of the molecule.” (Corpus ID: 13793404)

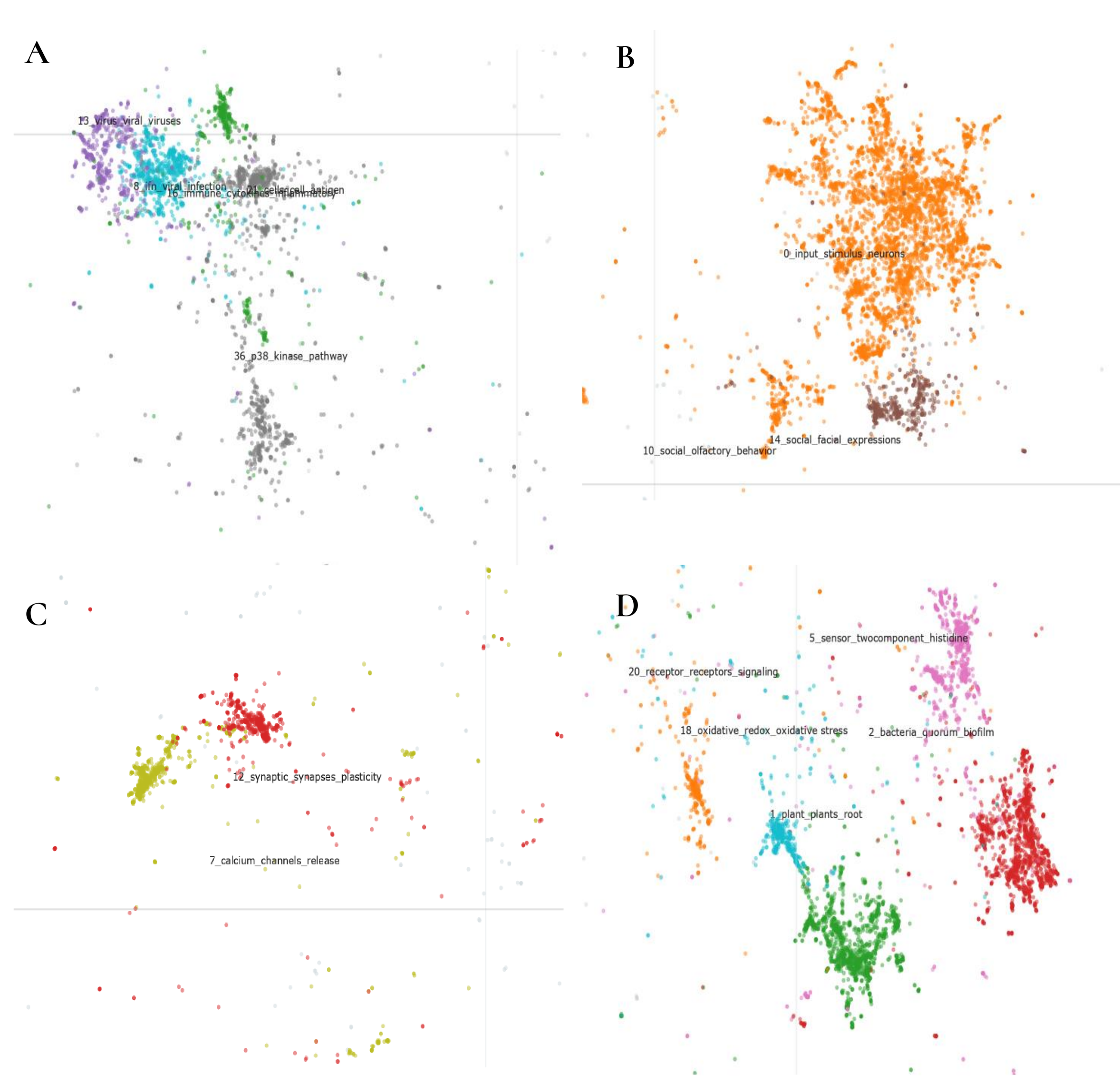

Figure 3. Document clusters for topics from top 20 topics, manually selected for topics with higher document-level coherence. A – topics related to viruses and immune system; B – “social” topics and their close proximity to the largest topic on sensory system; C – topics central to molecular-level neuroscience; D – topics discussing signaling in bacteria and plants

Box 2. Informational definition of signaling

“Signaling is a type of behavior that enables the sharing of information between interacting individuals.” (Scott-Phillips 2008, 391)

Identified mentions (manually picked from top 10):
I-CMN1. “There is compelling evidence that [oxytocin] signaling plays an important role in mediating multiple aspects of social behaviors—encompassing social investigation, social motivation, social memory, and social threat.” (Corpus ID: 253598169)
I-CMN2. “It is worth noting at this point, individuals only interact when a signal has been given.” (Corpus ID: 262355364)

I-MoB1. “The vast majority of interactions are unlikely to be related to signaling.” (Corpus ID: 195890841)
I-MoB2. “From this point of view, a signaling network communicates and conveys signals from its inputs to the outputs.” (Corpus ID: 202554259)
I-MoB3. “Signaling networks are rather assemblies of intricate highly connected modules controlling key biological processes in a context-dependent manner.” (Corpus ID: 233457609)

I-MiB1. “Various signaling pathways play their respective roles and can interact with each other.” (Corpus ID: 248003769)
I-MiB2. “Signaling is a series of chemical and/or energetic transmissions from an external stimulus to the cell.” (Corpus ID: 3343152)
I-MiB3. “A common feature of many signaling systems is ”recruitment” of signaling proteins into complexes by specific interaction with another protein in the complex.” (Corpus ID: 244479137)

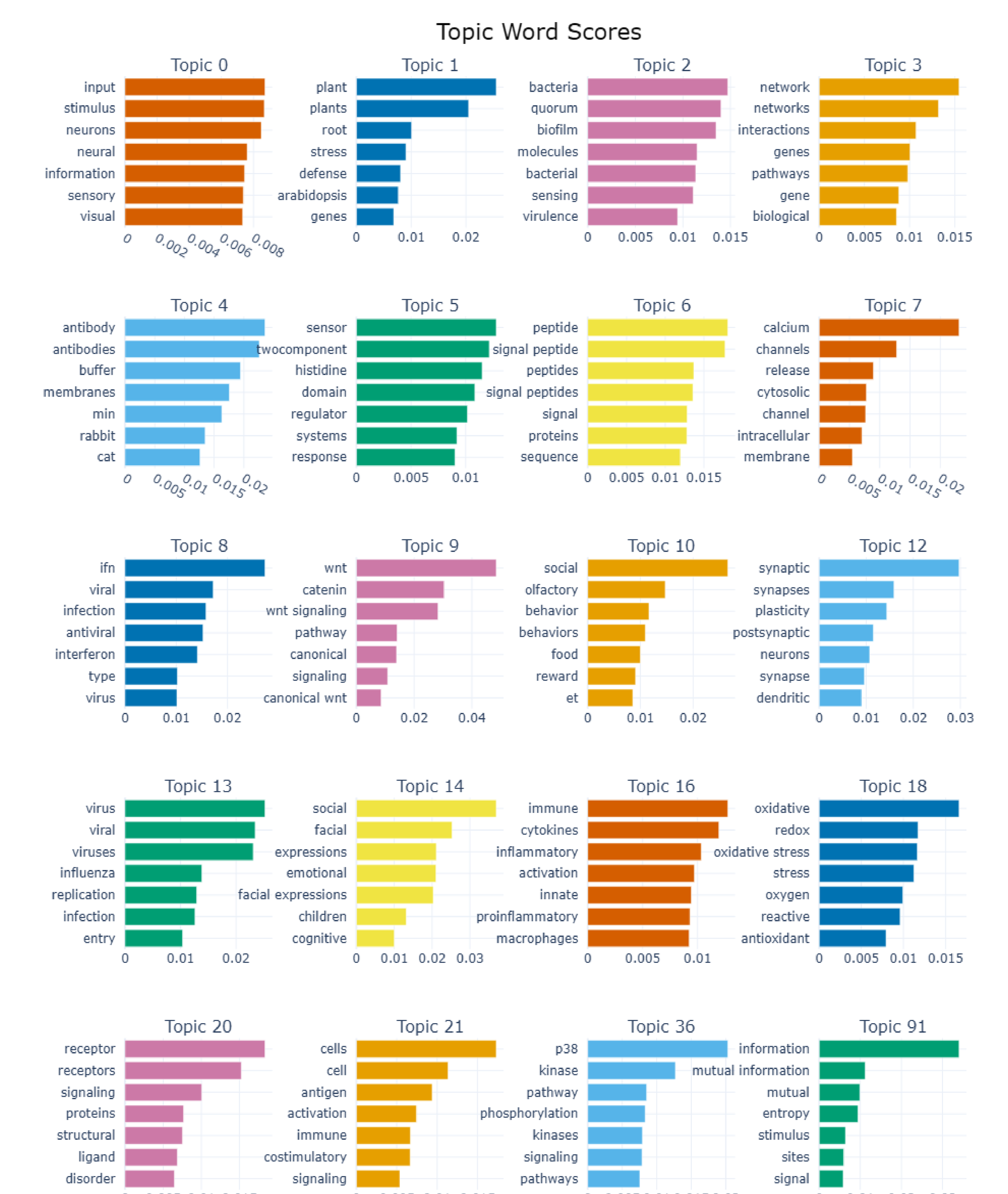

Figure 2. Contribution of 7 top terms for 20 most frequent topics.

Box 4. Most distinctive and common topics for disciplines

Cellular and Molecular Neuroscience:
1. 12_synaptic_synapses_plasticity_postsynaptic: synaptic, synapses, plasticity, postsynaptic, neurons, synapse, dendritic, synaptic plasticity, presynaptic
2. 0_input_stimulus_neurons_neural: input, stimulus, neurons, neural, information, sensory, visual, activity, network, noise
3. 82_reward_learning_dopamine_aversive: reward, learning, dopamine, aversive, salience, appetitive, value, dopaminergic, neurons, outcomes
4. 14_social_facial_expressions_emotional: social, facial, expressions, emotional, facial expressions, children, cognitive, individuals, face, autism

Molecular Biology:
1. 139_et_regeneration_nerve_injury: regeneration, nerve, injury, axon, growth, neurons, factor, factors, cell
2. 77_signaling pathways_pathways_signaling_pathway: signaling pathways, pathways, signaling, pathway, signaling pathway, notch, cancer, wnt, cell, genes

Microbiology:
1. 2_bacteria_quorum_biofilm_molecules: bacteria, quorum, biofilm, molecules, bacterial, sensing, virulence, formation, species, production
2. 23_microbial_species_community_microbiome: microbial, species, community, microbiome, interactions, communities, taxa, diversity, microbial communities, gut

Most common:
1. 42_light_blue light_stomatal_photoreceptors: light, blue light, stomatal, photoreceptors, photosynthetic, chromophore, phytochromes, signal, blue
2. 145_resource_conspecifics_recruitment_flows: resource, conspecifics, recruitment, flows, uptake, electric, signaller, fields, organisms, search (note: this topic is shared by neuroscience with micro- and molecular biology, but not by micro- and molecular biology)

The case study of “signaling”

“Communication is so widespread in the biological world, that we probably can use it to characterize life” (Hasson 1997). Yet, it is far from clear what “communication” actually is. In this study, I investigate how the concept of “signaling” specifically is invoked in the corpus of articles from microbiology, molecular biology, and cellular and molecular neuroscience (see Supplemental Information for detailed description of data and methods, as well as results). Using two alternative definitions of “signaling”, adaptationist (Box 1) and informational (Box 2; see Scott-Phillips 2008), I have identified paragraphs which mention “signaling”. Next, I have constructed a topic model of those paragraphs (Box 3), which enabled a detailed comparison between different contexts and disciplines (Box 4).

Discussion

Semantic search using the adaptationist definition seems to be picking out more interesting contexts, compared to the informational definition which is much more generic, and in result picks out mostly generic mentions of the term. Yet, selection, which plays a central differentiating role in adaptationist definition, is mentioned directly only in the microbiology subcorpus (e.g., A-MiB2, A-MiB3). Interestingly, the adaptationist definition more readily picks out contexts related to experimental protocols and data analysis (e.g., A-CMN2, A-MoB3). In all contexts signals seem to be characterized as a relation from an input (sensation, stimulus) to an output (behavior, response). While the adaptationist definition contains this structure explicitly, the informational focuses more on the continuous, processual character of signaling and doesn’t presuppose well-delineated outputs. In result, this definition picks out more mentions of interaction, whether in reference to social behaviors (in the neuroscience subcorpus) or to the complexity of signaling networks, presenting significantly less directed account of signaling compared to the adaptationist definition.

While there are significant differences of topic distributions between the disciplinary subcorpora (see Supplemental Information), the effect is not strong, even when looking at topics which differentiate the subcorpora the most. These topics refer primarily to distinctive areas of investigations of the disciplines—with the references to social cognition, dopamine and reinforcement learning, and synaptic plasticity in the context of cellular neuroscience, regeneration and cancer-related notch signaling in molecular biology, and bacterial communities and quorum sensing in microbiology. Interestingly, the topic that is least distinctive between the disciplines refers to light, and the role it plays whether in vision (in neuroscience subcorpus), for other forms of photosensitivity (in microbiology) or for optogenetics (in molecular biology). At the same time, the five broader categories identified within the topics (see Box 3) are represented in all subcorpora. While “signaling” is used to account for different phenomena, the underlying concept seems largely the same across the different disciplines.

Box 3. Topic modeling results

The topic model identified 163 topics, which can be grouped into several broad categories:

1. **Modeling**: topics which discuss signaling on a relatively abstract level, often in connection to computational modeling, e.g., 3_network_networks_interactions_genes, 33_boolean_logical_networks_node, 87_theory_information_phase_synchronization, 98_interspecies_chemical_communication_signal molecules, 158_cooperation_game_evolve_cheaters
2. **Mechanistic**: topics which focus on specific signaling compounds and physical structures, e.g., 9_wnt_catenin_wnt signaling_pathway, 25_exosomes_vesicles_cells_extracellular vesicles, 81_striatum_cocaine_dopamine_neurons, 90_mechanical_mechanical signals_forces_mechanotransduction, 159_notch_notch1_differentiation_infection
3. **Functional**: topics which focus on the biological processes signaling contributes to, e.g., 2_bacteria_quorum_biofilm_molecules, 8_ifn_viral_infection_antiviral, 54_circadian_clock_rhythms_clocks, 60_bone_differentiation_osteogenic_osteoblasts, 82_reward_learning_dopamine_aversive, 110_immune_innate_immune response_recognition
4. **Measurement**: topics which refer to signals as measurements, e.g., 24_fluorescence_nm_imaging_image, 26_data_genes_gene_variants, 37_signals_metabolites_mass_spectra, 89_phylogenetic_phylogenetic signal_tree_phenotype, 124_cm_biosensors_biosensor_detection
5. **Artifacts**: topics which result from not theoretically motivated mentions of terms related to “signaling”, e.g., 69_communication_social media_science_people, 79_panels_probe_supplementary_individual signals