

Entry

Tokenization in the Theory of Knowledge

Robert Friedman [†] 

Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA;
bob.network.science@gmail.com

[†] Retired.

Definition: Tokenization is a procedure for recovering the elements of interest in a sequence of data. This term is commonly used to describe an initial step in the processing of programming languages, and also for the preparation of input data in the case of artificial neural networks; however, it is a generalizable concept that applies to reducing a complex form to its basic elements, whether in the context of computer science or in natural processes. In this entry, the general concept of a token and its attributes are defined, along with its role in different contexts, such as deep learning methods. Included here are suggestions for further theoretical and empirical analysis of tokenization, particularly regarding its use in deep learning, as it is a rate-limiting step and a possible bottleneck when the results do not meet expectations.

Keywords: token; tokenization; neural network; deep learning; cognition; atomism; natural language

1. Tokens and Their Properties

In computer science, a token is an element of a programming language, such as the expression of a multiplication operation or a reserved keyword to exit a block of code. In a more general context, including from other areas of study, tokenization and tokens are useful to represent objects in general. In the case of a common computer program, tokenization is a procedure that creates tokens from a data source that consists of human readable text in the form of instructions (computer code). Given that the source code is a target to be compiled as a program, the idealized procedure is to begin with a tokenizer that transforms the text to a sequence of tokens, followed by delivery of the individual tokens to the subsequent step. Next, a parsing procedure transforms the tokenized sequence to a hierarchical data structure, along with validation and further processing of the code. Lastly, the machine code, a set of instructions as expected by a computer processing unit, is generated [1].

In an artificial environment, such as machine learning, there are other examples of tokenization apart from the idealized computer compiler, including the multi-layer artificial neural network and its weighted connections, in itself a type of computer program [2,3]. This network and its input data are likewise dependent on tokenization to create a sequence of tokens. As a neural network is expanded by adding layers of artificial neurons, this architecture may be referred to as deep learning.

In Nature, such as in the case of human speech, words are also divisible into a set of smaller elements, the phonemes, which are a type of set of tokens [4]. These phonemes are categorized by their occurrence along the natural boundaries of spoken words, emerging from the mechanics of speech and the cognitive processes. These elements, the phonemes, are further transformed, processed, and potentially routed to the other pathways of cognition. From an abstract perspective, these corresponding pathways encapsulate the internal representational forms, a type of intermediate language that originates with the tokenization process; these forms are expected to emerge from the cognitive processing of the phonemes. Otherwise, there is an expectation that the phonemes are unprocessed and uninterpretable along the downstream pathways of cognition.



Citation: Friedman, R. Tokenization in the Theory of Knowledge. *Encyclopedia* **2023**, *3*, 380–386. <https://doi.org/10.3390/encyclopedia3010024>

Academic Editors: Kamran Munir, José Raúl Romero, Khalid Hafeez and Raffaele Barretta

Received: 29 November 2022

Revised: 11 March 2023

Accepted: 16 March 2023

Published: 20 March 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Tokens are also the essential elements of text as observed in common documents. In this case, tokens are typically constrained in number, presumably a limitation of any human readable source that relies on the limits of human thought. However, in all cases, there is a potential for the generation of combinations and recombinations of tokens, and therefore the formation of new objects based on prior information, a process that corresponds to the mechanics of building knowledge. In deep learning, such as that implemented by the transformer architecture [5], the power is in collecting a large number of tokens as the essential elements for training a neural network [6–8]—a graph composed of artificial neurons (nodes) and their interconnections (edges) with weight value assignments [2,3].

2. Tokens as Elements of Nature

The basis of forming higher order objects from a set of constituent elements (atomism) has its origins in antiquity, including in the theories of Leucippus and Democritus [9,10]. Their theories of the natural world suggest the existence of a variety of elemental substances; each variant is associated with a unique set of properties, and the variants may combine and recombine into the more complex forms. From a modern perspective, observations in the natural world are consistent with these ideas of atomism, including atoms and molecules as the constituents of matter, the biological molecules of cells, the genetic elements of living organisms, and the protein receptors of vertebrate immunity [11].

The tokens of deep learning also represent a set of elements, the discretized units of information. Therefore, the representational forms of the natural world and those of cognition and neural networks are expected to originate from the most basic of elemental forms; the open question, at least in Nature, is whether any set of elements is truly indivisible into a set of smaller forms. A modern corollary to atomism is that natural processes are non-deterministic, and the divisibility of any hypothesized elemental form is indefinite, a problem that may extend beyond the limits of knowledge [12].

Nature provides examples on the limits of the potentiality of elements to combine into the more complex forms. An example from biological evolution shows limits in the exploration of all possible genetic variants, as originating by mutation and recombination, where these processes are restricted in practice by the high probability of lower fitness with the occurrence of each genetic change [13]. Therefore, the evolutionary search space of all possible life forms is largely unexplored. Another example is in the chemistry of biological processes. The molecules of cells are a highly restricted subset of arrangements as compared to the world of all possible molecules; this is partly a result of the evolutionary process and the unexplored space of novelty in molecular processes. These observations support the dictates of probabilism in Nature and its potential for construction of new forms, including restriction by spatial and temporal factors. In other words, Nature is highly dynamic at all scales as measured by spatial extent and time.

Likewise, in an artificial setting, a computer compiler is dependent on a restricted number of tokens, although this observation may be considered a limitation of the programmer's capability to memorize a very large number of tokens that correspond to the keywords, expressions, and other elements of a computer language. The artificial neural network is less or not necessarily restricted by biological constraints, so it may include a very large number of tokens. As an example, the transformer-based GPT-2 model of natural language [14], as implemented by Huggingface [15], is dependent on approximately fifty thousand tokens. In the case of a natural language corpus, the natural language model is instead restricted by a multitude of transcriptional and spoken forms [4], a phenomenon best sampled at the population level, so the number of tokens in the GPT-2 model may be compared against a survey of a common dictionary of words, such as in the English language [16,17]. This survey results in a half-million words. However, in the context of tokenization, a word may be classified as overlapping with other words if they share a major component. Another issue in tokenization is common versus rare words, where their distribution affects the reliability of a neural network trained on a number of tokens. For

example, if a word appears just a few times in a corpus, then its absence during tokenization and training may expectedly have little effect on the robustness of the model.

3. Tokens as Scientific Data

However, the many types of scientific data expressed in sequence format, an expression typically incompatible with the natural languages, are commonly encoded for the attributes of precision, reliability, and machine readability, such as available in the one letter amino acid code of proteins [18] and in the text-based SMILES code (simplified molecular input line entry system) that represents the two dimensional arrangement of atoms of a molecule [19–23]. The SMILES code is also used to represent the pathways of chemical reactions [20,22]. In these cases, there are limits on the atoms for forming chemical compounds, and, in the case of biology, the arrangement of amino acids that can form proteins. Therefore, tokenization based on the individual elements of these sequences results in a smaller number of tokens than can potentially occur [20,24]. However, if there is a generation of tokens in which the tokens may consist of more than one element, such as in the case of an amino acid sequence [25], then the token count may exceed, even significantly exceed, that observed in the natural languages, as the latter are further constrained by phoneme variety, a product of speech and its perception, and human memory.

Furthermore, as tokenization in deep learning affects the subsequent steps of the methodology, the distribution of potential tokens is a bottleneck for the performance of a model, and is a crucial area of study in theoretical and empirical contexts. This practice would lead to a better expectation in the design of a deep learning system that blends natural language with the many types of scientific data [23,26]. Another possible constraint in deep learning (for example, when using the highly efficient transformer architecture) is modeling the long-range associations of tokens along a sequence, particularly in a very large document where associations between words or symbols are distant.

As the sequence databases of biology and chemistry are vast, along with their appended descriptions in human readable format, deep learning methods are highly competitive with traditional search methods for the retrieval of scientific data and journal articles [26]. The Galactica model by Meta [26] further captures the information of mathematical expressions (LaTeX) and computer code. Moreover, Taylor and others [26] include a detailed discussion of the importance of tokenization in achieving their promising results, which shows improvement over prior work.

It is possible to further explore the many varieties and combinations of tokens of scientific data and the expectation on prediction of the higher order structures, such as those observed in the correspondence between a protein sequence and its tertiary structure [27–29]. Where the algorithms of Nature are not amenable to these approaches, then other methods of reinforcement learning [30] and use of a decision transformer [31] are powerful tools for the discovery of the higher order forms from the low dimensional information in scientific data.

These methods of curating and identifying syntactical structure within data, particularly in the case of automatic curation, as in the use of keyword tagging in a corpus for the identification and processing of the data types, is a synergistic approach for extending the foundational language models [32]. There is also an important distinction between natural language and scientific data. The latter are not a product of persuasive speech, at least not ordinarily, so the data quality is more verifiable. However, as expected, the natural language corpus is replete with examples of mere opinion and persuasion [33,34], so it is difficult, at best, and arguably not possible, to truly verify the validity of statements in common samples of natural language.

4. Tokens as Elements of Logic

The constraints on validating samples in the corpus as coded by natural language are partly a byproduct of phylogenetic dependence [35,36] on this form of communication, a descendant of similar forms of animal communication, including their underlying physical mechanisms, as observed in the morphological characters that relate to aural

communication. These include the mechanical and anatomical features of the larynx and also the bones of the middle ear. Furthermore, they contribute as drivers for the intermediate steps in the pathways of speech and its perception, along with their downstream influence on other cognitive pathways involved in the construction and interpretation of sensory data. This imperfect process of knowledge transfer has led to the ancient and modern human pursuit of rhetoric as an art form in itself, along with the support of vast academic and public enterprises of titled interpreters and associated publications based on opinion, and corresponding interpretation and reinterpretation of current and past written knowledge [33,34,37,38]. Altogether, this is presumed a byproduct of an evolutionary design for animal communication and behavior, as opposed to any engineered design purposed for the idealized flow of information and knowledge [39].

The lack of consistency and permanence in conveying an idea in any natural language, and that natural languages are not purposed for the construction of knowledge, led philosophers to pursue logic as a reliable method of communication of concepts [33,34,39,40]. This pursuit has been essentially a reference to a set of logical primitives that describe objects and their interrelationships as perceived in the world, a type of tokenizing of material and abstract entities, regardless of whether the origin of the object is in Nature or the Mind [33]. Further, for this conjecture, these basic primitives are considered the elemental constituents of any of these objects, the idealized forms. These forms are expectedly robust, precise, and low in dimensionality, properties that lend themselves to efficiency in computation (Figure 1). These properties are also ideal in the formation of reliable knowledge with resistance to perceptual perturbation, such as those observed across the various visual perspectives or the variation of percepts as constructed among individuals. The philosophers of antiquity [33] theorized about the permanence and indivisibility of these primitives (the elemental forms of logic), a general concept revisited and diminished by modern theorists and empiricists [12,39,40].

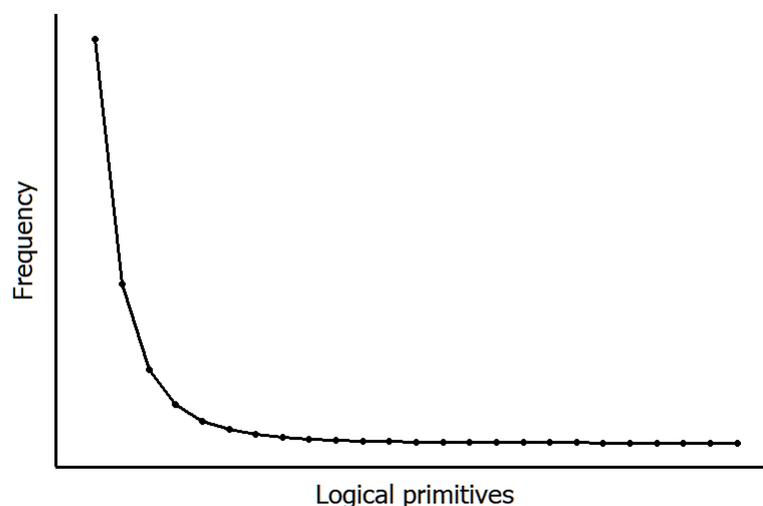


Figure 1. A long-tail distribution of logical primitives. This distribution shows the hypothesis that a few of the logical primitives are frequently used in cognition, but the vast majority are rarely used. This hypothesis is consistent with the emergence of new abilities in large language models [41].

The world of mathematics and of computer code are more closely related to the above formal systems than to natural language. Computer code and mathematics both consist of a formal syntax of keywords, symbols, and rules, and an adaptation for transformation to machine level code, such as by a computer compiler or interpreter. Use of an intermediate language for representing math or computer problems, such as LaTeX, or in the code of Python, has led to successes in deep learning and inference on advanced math, such as math questions as posed in a natural language format [26]. This is a concrete example of the importance of data, their tokenization, and signal processing [42] in the expectations

regarding building a robust model and recovery of the basic elements (primitives) from a data source.

As it is possible to acquire general knowledge of an academic topic with dependence on a structured and formalized language, such as in the mathematics of trigonometry or in the discipline of classical mechanics, these languages result in information that is hypothetically very compressible and adapted to the processes of learning and cognition. These languages are correspondingly distinct in their operations, syntax, and definitions. If this hypothesis is not true, then knowledge would not be generalizable (an extreme form of an hypothesis, as proposed in antiquity by Heraclitus, in which change is rendering all or nearly all forms as distinct and unique). The natural world instead has narrow paths of change that are mapped along physical pathways, at least within the confines of human experience and measurement, so a categorization of forms and its associated knowledge are favored against the product of miscategorization and a lack of commonality in Nature. This is another limit to knowledge of the natural world, as specifically documented in the evolutionary sciences, a discipline firmly based on an essential and formalized language for discerning the true and false categories of characters and their states [35,36].

5. Conclusions

In summary, there are a variety of ways to study a deep learning system, including the data collection step, the neural network architecture, and the dependence of these systems on a tokenization procedure for achieving robust results. Of these parts of a system, tokenization, is as crucial a factor as the design of the deep learning architecture (Table 1), although the architecture is often displayed as the main feature of deep learning. Moreover, ensuring the quality and curation of source data introduces efficiency and robustness in the training and inference stages of deep learning [26].

Table 1. Summary of concepts in this entry.

Concept	Description
Tokens in general	A token can be considered the elementary unit of a text document, speech, computer code, or another form of sequence information.
Tokens in Nature	In natural processes, a token can represent the elemental forms that matter are composed of, such as a chemical compound or the genetic material of a living organism.
Tokens as scientific data	In scientific data, a token can represent the smallest unit of information in a machine readable description of a chemical compound or a genetic sequence.
Tokens in math & logic	In the structured and precise languages of math and logic, a token can represent the elementary instructions that are read and processed, as in the case of an arithmetical expression, a keyword in a computer language, or a logical operator that connects logical statements.

Apart from data quality, it may be possible to increase the extraction of salient information, particularly in a scientific data set, by training the neural network on tokens as formed from different perspectives. As an example, biological data which has information as extracted from the individual amino acids of a sequence also has information at the subsequence level as reflected by the structural and functional elements of a protein. This possibility emerges from the observation that natural processes are influenced at various scales, as in the combinatorics of natural language with its definite biological limits, as observed by the dictionary of available phonemes and constraints in the associated cognitive processes. To capture the higher order information in protein sequence data, AlphaFold [27] appended tokens from other sources; however, it is ideal to capture these rules of Nature as they emerge from the sequence data itself and their associated tokens [28].

Furthermore, tokens, as the representative elements of knowledge, are not likely uniform in status, but instead are expected to have an uneven distribution in their applicability as descriptors (Figure 1). The commonality of forms and objects that lead to knowledge are

sometimes frequent in occurrence, but it can be conjectured that these commonalities are often rare. This hypothesis, even if proposed in posterior to recent findings from the study of large language models, supports an expansive search for data in the construction of deep learning models, even if to merely validate the importance of rare patterns that have not yet been sampled from the overall population of patterns in the world of data.

As a final comment, overall deep learning methodology may be considered a variant of the computer compiler and its architecture, with tokenization of input data (tokens) and transformation of tokens into a graph data structure, along with a final generation of code to form a program (neural network and weight values). Each step of this process is ideally subjected to hypothesis testing for better expectations and designs in the practice of the science of deep learning.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Wirth, N. *Compiler Construction*; Addison Wesley Longman Publishing, Co.: Harlow, UK, 1996.
2. Hinton, G.E. Connectionist learning procedures. *Artif. Intell.* **1989**, *40*, 185–234. [CrossRef]
3. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
4. Michaelis, H.; Jones, D. *A Phonetic Dictionary of the English Language*; Collins, B., Mees, I.M., Eds.; Daniel Jones: Selected Works; Routledge: London, UK, 2002.
5. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. In Proceedings of the 31st Conference on Neural Information Processing System, Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008.
6. Zand, J.; Roberts, S. Mixture Density Conditional Generative Adversarial Network Models (MD-CGAN). *Signals* **2021**, *2*, 559–569. [CrossRef]
7. Mena, F.; Olivares, P.; Bugueño, M.; Molina, G.; Araya, M. On the Quality of Deep Representations for Kepler Light Curves Using Variational Auto-Encoders. *Signals* **2021**, *2*, 706–728. [CrossRef]
8. Saqib, M.; Anwar, A.; Anwar, S.; Petersson, L.; Sharma, N.; Blumenstein, M. COVID-19 Detection from Radiographs: Is Deep Learning Able to Handle the Crisis? *Signals* **2022**, *3*, 296–312. [CrossRef]
9. Kirk, G.S.; Raven, J.E. *The Presocratic Philosophers*; Cambridge University Press: London, UK, 1957.
10. *The Stanford Encyclopedia of Philosophy*; Stanford University: Stanford, CA, USA. Available online: <https://plato.stanford.edu/archives/win2016/entries/democritus>; <https://plato.stanford.edu/archives/win2016/entries/leucippus> (accessed on 27 November 2022).
11. Friedman, R. A Perspective on Information Optimality in a Neural Circuit and Other Biological Systems. *Signals* **2022**, *3*, 410–427. [CrossRef]
12. Godel, K. *Kurt Godel: Collected Works: Volume I: Publications 1929–1936*; Oxford University Press: New York, NY, USA, 1986.
13. Kimura, M. The Neutral Theory of Molecular Evolution. *Sci. Am.* **1979**, *241*, 98–129. [CrossRef]
14. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language Models Are Unsupervised Multitask Learners. Available online: https://d4mucfpsywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf (accessed on 5 September 2022).
15. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-Art Natural Language Processing. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Online, 16–20 November 2020; pp. 38–45.
16. *Merriam-Webster Dictionary*; An Encyclopedia Britannica Company: Chicago, IL, USA. Available online: <https://www.merriam-webster.com/dictionary/cognition> (accessed on 27 July 2022).
17. *Cambridge Dictionary*; Cambridge University Press: Cambridge, UK. Available online: <https://dictionary.cambridge.org/us/dictionary/english/cognition> (accessed on 27 July 2022).
18. IUPAC-IUB Joint Commission on Biochemical Nomenclature. Nomenclature and Symbolism for Amino Acids and Peptides. *Eur. J. Biochem.* **1984**, *138*, 9–37.
19. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36. [CrossRef]
20. Bort, W.; Baskin, I.I.; Gimadiev, T.; Mukanov, A.; Nugmanov, R.; Sidorov, P.; Marcou, G.; Horvath, D.; Klimchuk, O.; Madzhidov, T.; et al. Discovery of novel chemical reactions by deep generative recurrent neural network. *Sci. Rep.* **2021**, *11*, 3178. [CrossRef]
21. Quiros, M.; Grazulis, S.; Girdzijauskaite, S.; Merkys, A.; Vaitkus, A. Using SMILES strings for the description of chemical connectivity in the Crystallography Open Database. *J. Cheminform.* **2018**, *10*, 23. [CrossRef]
22. Schwaller, P.; Hoover, B.; Raymond, J.L.; Strobelt, H.; Laino, T. Extraction of organic chemistry grammar from unsupervised learning of chemical reactions. *Sci. Adv.* **2021**, *7*, eabe4166. [CrossRef] [PubMed]
23. Zeng, Z.; Yao, Y.; Liu, Z.; Sun, M. A deep-learning system bridging molecule structure and biomedical text with comprehension comparable to human professionals. *Nat. Commun.* **2022**, *13*, 862. [CrossRef] [PubMed]

24. Friedman, R. A Hierarchy of Interactions between Pathogenic Virus and Vertebrate Host. *Symmetry* **2022**, *14*, 2274. [[CrossRef](#)]
25. Ferruz, N.; Schmidt, S.; Hocker, B. ProtGPT2 is a deep unsupervised language model for protein design. *Nat. Commun.* **2022**, *13*, 4348. [[CrossRef](#)] [[PubMed](#)]
26. Taylor, R.; Kardas, M.; Cucurull, G.; Scialom, T.; Hartshorn, A.; Saravia, E.; Poulton, A.; Kerkez, V.; Stojnic, R. Galactica: A Large Language Model for Science. *arXiv* **2022**, arXiv:2211.09085.
27. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, K.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)]
28. Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; et al. Evolutionary-scale prediction of atomic level protein structure with a language model. *Science* **2023**, *379*, 1123. [[CrossRef](#)]
29. Wu, R.; Ding, F.; Wang, R.; Shen, R.; Zhang, X.; Luo, S.; Su, C.; Wu, Z.; Xie, Q.; Berger, B.; et al. High-resolution de novo structure prediction from primary sequence. *bioRxiv* **2022**. [[CrossRef](#)]
30. Fawzi, A.; Balog, M.; Huang, A.; Hubert, T.; Romera-Paredes, B.; Barekatin, M.; Novikov, A.; Ruiz, F.J.R.; Schrittwieser, J.; Swirszcz, G.; et al. Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **2022**, *610*, 47–53. [[CrossRef](#)]
31. Chen, L.; Lu, K.; Rajeswaran, A.; Lee, K.; Grover, A.; Laskin, M.; Abbeel, P.; Srinivas, A.; Mordatch, I. Decision Transformer: Reinforcement Learning via Sequence Modeling. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 15084–15097.
32. Bommasani, R.; Hudson, D.A.; Adeli, E.; Altman, R.; Arora, S.; von Arx, S.; Bernstein, M.S.; Bohg, J.; Bosselut, A.; Brunskill, E.; et al. On the opportunities and risks of foundation models. *arXiv* **2021**, arXiv:2108.07258.
33. Waddell, W.W. *The Parmenides of Plato*; James Maclehose and Sons: Glasgow, UK, 1894.
34. Lippmann, W. *Public Opinion*; Harcourt, Brace and Company: New York, NY, USA, 1922.
35. Hennig, W. *Grundzüge einer Theorie der Phylogenetischen Systematik*; Deutscher Zentralverlag: Berlin, Germany, 1950.
36. Hennig, W. Phylogenetic Systematics. *Annu. Rev. Entomol.* **1965**, *10*, 97–116. [[CrossRef](#)]
37. Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **2014**, *5*, 1093–1113. [[CrossRef](#)]
38. Friedman, R. All Is Perception. *Symmetry* **2022**, *14*, 1713. [[CrossRef](#)]
39. Russell, B. The Philosophy of Logical Atomism: Lectures 7–8. *Monist* **1919**, *29*, 345–380. [[CrossRef](#)]
40. Turing, A.M. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc. Lond. Math. Soc.* **1937**, *s2-42*, 230–265. [[CrossRef](#)]
41. Wei, J.; Tay, Y.; Bommasani, R.; Raffel, C.; Zoph, B.; Borgeaud, S.; Yogatama, D.; Bosma, M.; Zhou, D.; Metzler, D.; et al. Emergent Abilities of Large Language Models. *arXiv* **2022**, arXiv:2206.07682.
42. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 21–24 June 2022; pp. 10684–10695.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.