

Wittgensteinian Grammar and OED Dictionaries for AI

A foundation of semantic understanding of scientific texts

Gerd Graßhoff 

Humboldt-Universität zu Berlin, Deutsches Museum
gerd.grasshoff@hu-berlin.de

ABSTRACT

Meaning in scientific prose is a documented inventory of rules of use, scoped by time, person, and discipline, rather than similarity of tokens. Semantic understanding of scientific texts is conformity to such rules: what an expression means is shown by the grammar of its employment; whether a formulation hits that meaning is decided by whether it follows a rule or contradicts it — and a violation can be stated only by naming the specific rule whose authority is breached. This paper develops that claim from Wittgenstein’s later philosophy (rule-following, the standard metre, perspicuous representation) and from the entry form of the Oxford English Dictionary. It then shows how, after OCR and linguistic analysis, those rules can be compiled as LWgrammar entries from a declared corpus and consulted as a reference library. The argument is carried by cases: length without a measurement standard; Kepler’s *anima* beside *virtus motrix*; the phrase *lex naturae* attested in Euler as a nominative proposition and in Kepler as an ablative mode; three grammars of *vis motrix* in Kepler, Newton, and Euler. A formulation score operationalises the thesis: no verdict of violation without a cited rule and an anchor. The instrument is built and measured: an ai-dictionary whose senses are scoped by author, work, and period, with every attestation addressed to a page and a sentence; a rule-submission path with a revision lifecycle; a conformity check that assigns each rule the verdict engine its content demands (dependency syntax, lexis, or natural-language inference), calibrated to 16/16 on a seeded golden test; and a logical-relation scan that finds, among the 663,972 sentences of an archival collection, those expressing propositions that entail or contradict a given proposition. The same rules guard what language models write: every revision is measured against its source, each measurement names its rule, and the guard rejects 133 of 136 revisions in which a lectoring run had inserted evaluations of the author’s text; the author’s accept and reject decisions become evidence for or against the rules.

KEYWORDS Wittgenstein · rule-following · Oxford English Dictionary · LWgrammar · historical scientific Latin · semantic conformity · Kepler · Euler

Table of contents

1 The problem: token similarity is not understanding	2
2 Following a rule	3
2.1 The paradox and its resolution as practice	3
2.2 The negation test: grammatical sentences are not empirical	4
2.3 The standard metre: a role, not a property	4
2.4 The same sentence, a different logical function	4
2.5 Where conventions live, and why they move	5
3 The dictionary as a grammar of use	6
4 From page to grammar	8
5 Microscopic semantic history	9
5.1 Case A — Kepler: from soul to force, without substitution	9
5.2 Case B — <i>lex naturae</i> as an emergent phrase	9
5.3 Case C — <i>vis motrix</i> across three authors	10
6 The ai-dictionary: scoped senses and submitted rules	11
6.1 The entry form	11
6.2 The submission path	12
6.3 From annotation to rule	12

6.4 Services	16
6.5 Modes: the same store under different contexts	17
6.6 The consultation hub: three application forms	17
6.7 Lectoring for maximal linguistic precision	17
7 Measuring semantic conformity	18
7.1 Four verdicts	18
7.2 The formulation score	19
7.3 Context axes and style profiles	19
8 Two services, measured	20
8.1 The conformity check	20
8.2 Logical relations across a collection	21
9 Guarding what a model writes	21
9.1 Task forms, context, and measurements	22
9.2 Measured on the damaged revisions	22
9.3 Decisions become rule evidence	23
9.4 Senses in time: the store behind Case A	23
10 Relation to fine-tuning	23
11 Conclusion	24
Bibliography	24

1 The problem: token similarity is not understanding

The Latin lemma *vis* occurs in Kepler's *Astronomia nova* (1609), Newton's *Principia* (1687), and Euler's mechanics of the 1740s. A retrieval system that ranks passages by embedding similarity treats these three passages as near neighbours. They are not: Kepler models *virtus motrix* on light spreading from the Sun; Newton defines *vis motrix* as the product of an accelerating force and a quantity of matter; Euler defines *vis motrix* as the cause that of itself always produces motion. One lemma, three language-games. Token similarity collapses the distinction a historian of science seeks to recover.

This paper therefore asks not how to embed scientific Latin more cleverly, but what it means to *follow* or *contradict* a meaning-rule — and how, after OCR and linguistic parsing, to build an instrument that makes that judgement checkable on scientific texts.

Five contributions follow from that question:

1. **Rule-following as a measure of conformity.** A formulation is semantically in order when it follows an identifiable rule of use; a lector may report a violation only by naming and citing the rule.
2. **The OED entry as perspicuous representation.** The dictionary that begins from dated quotations and lets the definition summarise the evidence is the lexicographical form of Wittgenstein's *übersichtliche Darstellung*.
3. **LWgrammar.** The same entry type, compiled from one declared corpus after OCR and Stanza analysis, with the rule of use written out as `usage_rule`.
4. **A casebook of microscopic semantic history.** Who used *anima*, *vis*, *lex naturae* when, and under which rule — with every quotation anchored at `(doc_id, pidx, sent_idx)`.
5. **Context-specific judgement.** Time, person, and discipline scope the same engine; publication-format profiles (the analogue of a LaTeX document class) configure style dimensions without replacing the rule inventory.

CONCEPT The thesis in one sentence

Semantic understanding is rule-conformity with a citable rule, not latent similarity. A verdict of violation that cannot name the rule it invokes has not yet made a judgement.

This paper works immanently on the *Philosophische Untersuchungen* and *Über Gewißheit*, pairing them with a built instrument and with corpus evidence that the reader can open at a page. It does not rehearse secondary literature on Wittgenstein. Companion papers carry what this one assumes: the PDAI page format [1], the linguistic index [2], the dictionary as a scholar's operation [3], the Partitur as the score behind an agent's answer [4], and the golden-query testbed that holds those operations to a declared oracle [5]. Computational epistemology treats the validity of *arguments* [6]; this paper treats the grammar of *words*. The two tasks are neighbours, not substitutes.

■ 2 Following a rule

■ 2.1 The paradox and its resolution as practice

Wittgenstein states the paradox of rule-following in PU §201:

Unser Paradox war dies: eine Regel könnte keine Handlungsweise bestimmen, da jede Handlungsweise mit der Regel in Übereinstimmung zu bringen sei. Die Antwort war: Ist jede mit der Regel in Übereinstimmung zu bringen, dann auch zum Widerspruch. Daher gäbe es hier weder Übereinstimmung noch Widerspruch.

[7, PU §201]

If every action can agree with a rule, then every action can contradict it. Agreement and contradiction would both vanish.

The example that makes the paradox concrete is the pupil of PU §185. He has mastered, by the ordinary criteria, the series of cardinal numbers and the command “+ n”. The exercises have been in the number space up to 1 000. Asked to continue “+ 2” beyond 1 000, he writes 1 000, 1 004, 1 008, 1 012. Told that he was to add *two*, he does not understand. One can always offer a further interpretation of “+ 2” (add two up to 1 000, add four thereafter) for what he did [7, PU §185]. No finite sample of applications determines the next step by itself. That is why the answer cannot be another interpretation. The misunderstanding, §201 continues, shows itself in the habit of placing interpretation behind interpretation, as if each reading of the rule calmed us until the next one unsettled it. What that chain reveals is that there is a way of taking a rule which is *not* an interpretation, but which expresses itself, from case to case of application, in what we call “following the rule” and “going against it”.

That taking is a practice. PU §202 draws the consequence:

Darum ist ›der Regel folgen‹ eine Praxis. Und der Regel zu folgen *glauben* ist nicht: der Regel folgen.

[7, PU §202]

To follow a rule is not to have a private conviction that one is following it. And one cannot follow a rule privately, because then believing one was following it would be the same as following it. The grammar of “following a rule” is the grammar of a custom. PU §199 makes the point as a remark on that grammar: following a rule, giving an order, playing a game of chess are *Gepflogenheiten* — customs, institutions. A language is a technique one has mastered, not a hypothesis one entertains about signs.

For semantic judgement this is not atmosphere. It is the difference between two kinds of report. “This sentence is off” is an impression. “This sentence violates rule *R*, whose statement is ... and whose authority is cited at PU §50 / CGPM 1983 / Kepler kgw-01 Bl. 83 (S. 70)” is a move in a practice. The second report can be contested, because the rule can be looked up. The first cannot.

METHOD Following versus contradicting

To contradict a rule is not to use a word vaguely or poorly. It is to go against a *nameable* rule of a practice. The operational form of that demand: a verdict of violation cites a rule identifier and a documentary anchor. Without both, the verdict is not a judgement but a mood.

2.2 The negation test: grammatical sentences are not empirical

PU §251 distinguishes a grammatical sentence from an empirical one by what happens under negation. “I cannot imagine the opposite” is not a report on the limits of one’s imagination. It is a defence against something that *looks* like an empirical claim and is in fact a rule of the language-game. The example is “Every rod has a length”:

Beispiel: »Jeder Stab hat eine Länge«. Das heißt etwa: wir nennen etwas (oder *dies*) »die Länge eines Stabes« — aber nichts »die Länge einer Kugel«.

[7, PU §251]

The negation “Some rods have no length” is not a false physical hypothesis. It is nonsense in this game: it withdraws the very technique by which “length of a rod” is said. That is the test this paper will use on scientific prose. A formulation that drops the standard by which a quantity is defined does not make a weaker empirical claim. It leaves the game in which that quantity can be said.

2.3 The standard metre: a role, not a property

PU §50 is the load-bearing example, and the wording matters. Wittgenstein writes *aussagen*, not *sagen*:

Man kann von *einem* Ding nicht aussagen, es sei 1 m lang, noch, es sei nicht 1 m lang, und das ist das Urmeter in Paris. — Damit haben wir aber diesem natürlich nicht irgend eine merkwürdige Eigenschaft zugeschrieben, sondern nur seine eigenartige Rolle im Spiel des Messens mit dem Metermaß gekennzeichnet.

[7, PU §50]

One cannot *predicate* of the Paris prototype that it is, or is not, one metre long. The sentence would seem to measure something; instead, it stipulates the game. The prototype has no occult property; it has a role. When one measures, all other length-talk in that game multiplies that role quantitatively.

The scientific consequence is immediate. Before each redefinition, the then-current standard is the reference against which length-statements are moves. A length given without reference to the standard and the procedure is already a violation of the grammar of measuring: it is unclear what the figure is a multiple of. The violation is not that the number is inaccurate. The violation is that the number has not yet been given a move in the game.

2.4 The same sentence, a different logical function

The metre has changed its definition without the word “metre” disappearing. Three epochs, three stipulations [8]:

Epoch	Definition	What is held fixed
1799–1960	Platinum–iridium Paris	prototype, An artefact as reference standard

Epoch	Definition	What is held fixed
1960–1983	1 650 763.73 wavelengths of Kr-86 radiation	11th CGPM, Resolution 6 [9]
1983–	Distance light travels in vacuum in 1/299 792 458 s	17th CGPM, Resolution 1 [10]

Since 1983 the sentence “The speed of light is 299 792 458 m/s” is no longer an empirical report. It *defines* what a metre is. The wording is almost unchanged; the logical function has switched. The sentence has frozen into the riverbed (the image is taken up below). One does not test it against a measurement. One tests the measurement against it.

The OPERA neutrino-velocity episode is the counter-check. In September 2011 the collaboration announced an apparent arrival of CNGS muon neutrinos earlier than light-travel time would allow [11]. The physics community did not, as a body, treat “nothing travels faster than c ” as a hypothesis now in doubt. It treated the timing chain as the thing to be examined. A fibre-optic connection in the GPS synchronisation path and a clock oscillator were identified as the sources of the anomaly; the published measurement is consistent with c [12]. The measurement was checked against the sentence, not the sentence against the measurement — for as long as that sentence plays the role of a norm of description.

RULE R-SI-LENGTH-REF

A quantitative length without reference to the metre definition then in force, and without a measuring procedure, violates the grammar of measuring: it remains unclear to which standard and which procedure the figure is attached. Authority: BIPM SI Brochure, 9th ed. (2019). Anchor: PU §50.

Two everyday and two scientific violations of this family:

1. “The rod is 2 m long,” with no standard and no procedure. Violates R-SI-LENGTH-REF and the grammar of PU §50.
2. The same sentence in a scientific paper, still without the standard. Same rule, tightened by the explicitness demanded of scholarly prose.
3. A word used against the sense documented in the governing dictionary, without being marked as a coinage. Violates the documented rule of use; the verdict cites the entry.
4. Kepler 1596 (*anima motrix, anima mundi*) read with Newton’s grammar of *vis*. An anachronism: it violates the temporally scoped rule of the corpus. The case is worked in Section 5.

2.5 Where conventions live, and why they move

If following a rule is a practice, conventions live where the practice is kept: in a metrology conference, a dictionary office, a laboratory, a discipline’s way of writing. The OED, the Duden, Le Robert and the Académie, Treccani, the Historical Dictionary of the Academy of the Hebrew Language, Grimm’s *Deutsches Wörterbuch* — these are not oracles. They are movable authorities. A speech community can violate them, and they change historically. That is the dynamic Wittgenstein returns to in *Über Gewißheit*.

Certain sentences of the form of empirical sentences can *harden* and function as a channel for the ones that remain fluid; the relation changes with time, as fluid sentences harden and fixed ones liquefy [13, ÜG §§96–97]. The mythology can flow again; the riverbed of thoughts can shift. Wittgenstein distinguishes the motion of the water from the shift of the bed, while refusing a sharp boundary between them. That is the right picture for the metre and for a dictionary sense alike.

The hinge image gives the other half. The questions we ask, and our doubts, rest on certain sentences being exempted from doubt — the hinges on which they turn. If the door is to turn, the hinges must stay put [13, ÜG

§§341, 343]. A check presupposes rules that are not themselves being checked in the same move. Semantic conformity is such a check. It is not a demand that every rule be proved before it is used.

Two further remarks in *Über Gewißheit* decide the relativism question that the riverbed image otherwise leaves open. First, a world-picture is not held because one has been convinced of its correctness. It is the inherited background against which one distinguishes true and false [13, ÜG §94]. Second, that background can still be revised, and the revision can be given reasons. §92 asks whether someone could have a good ground for believing the earth had existed only since his birth; Moore, meeting a king raised in that belief, might convert him — but it would be a conversion of a special kind, the king brought to look at the world differently. Simplicity and symmetry can move one to a view; one then says “It must be so.” §§608–612 press the same point from the other side. Acting on the propositions of physics *is* what we call having a good ground. One would “combat” a person who did not; one would give reasons; and at the end of reasons stands persuasion. Conventionality is not arbitrariness. Rules can be revised; a revision that cannot give reasons has left the practice.

The instrument below takes that pair of thoughts as a design constraint. Rules are scoped (they have a time, a person, a discipline). They are citable (they have an identifier and an anchor). They are revisable (an LWgrammar entry has a revision log). They are never silent: a missing rule is reported as a status, not as an empty search.

■ 3 The dictionary as a grammar of use

PU §43 gives the method, with the restriction Wittgenstein himself writes in:

Man kann für eine *große* Klasse von Fällen der Benützung des Wortes »Bedeutung« — wenn auch nicht für *alle* Fälle seiner Benützung — dieses Wort so erklären: Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache.

[7, PU §43]

For a large class of cases, though not for all, the meaning of a word is its use in the language. The claim is methodological, not a reduction of meaning to behaviour. If one wants to know what a term means, one does not hunt an essence. One surveys the uses. PU §373 states what that survey yields: grammar tells us what kind of object anything is. PU §122 names the form of the survey. A principal source of our lack of understanding is that we do not *overlook* the use of our words; our grammar lacks perspicuity. The perspicuous representation (*übersichtliche Darstellung*) is the understanding that consists in seeing connections.

The Oxford English Dictionary is already that form, in lexicography. It does not start from a definition and hunt for illustrations. It starts from dated quotations and lets the definition summarise the evidence. An entry collects real sentences, sorts them into groups that share a way of using the word (branches and numbered senses), and gives each group a short label read off the examples. Around that core it records spellings (Forms), origin (Etymology), and marks the uses that have died (†). The quotations are the evidence; the definitions are summaries of that evidence. The entry never says what the word *is*. It shows the family of language-games the word plays.

The entry for English *time* is the structural model — not reproduced here as OED text, which remains the Dictionary’s, but as an anatomy. Forms and etymology sit above the senses. Branch I treats time as a bounded stretch (a while, a lifetime). Branch II treats time *when*: a juncture, an appointed moment. Later branches generalise (time as a dimension) and record combinations. Newton’s scholium on absolute time in the *Principia* belongs in the generalised branch: time as a container independent of matter, subject to mathematical operations [14]. A relational use, in which time is derivative of motion, is a different game. The dagger marks what has dropped out of the living language. That is perspicuity: one page on which the connections can be seen.

LWgrammar writes the same object from **one declared corpus**. The corpus sentences become the quotations; the branches capture the distinct language-games of that corpus; every quotation keeps its address (`doc_id`, `pid`, `sent_idx`). The one thing the OED leaves implicit — the actual rule of use: which verbs the term takes, what it is contrasted with, the construction it sits in — LWgrammar writes as `usage_rule`, because that rule is what a search engine and a conformity judge need.

Two live entries from the Eneström corpus make the anatomy concrete. Both were read from `enestroem_lwgrammar.duckdb` (short-lived read-only connection) for this paper.

tempus (lemma `tempus`, Latin noun). Corpus frequency 5,548 tokens across 252 documents — 4,547 Latin sentences in 251 of them; four senses in two branches; first attestation in the compiled entry at E1. A scoped sibling `enestroem:tempus:la#le1748` compiles the same lemma on documents up to 1748 (84 documents, 2,749 tokens, three senses); the rule of use is read off a different evidence set, so the entry is allowed to differ.

- **Branch I — physical and mathematical time.** Sense I.1: duration of physical motion, especially descent under gravity. Rule of use: a variable to be minimised (`tempus descensus`, `tempore brevissimo`) or equated (`tempore aequali`), typically governed by $\int ds/v$. Pivotal quotation: Euler, *Methodus inveniendi* (E101), Bl. 28 (S. 29), `sent_idx 3` (1744) — *Ut si quaerenda sit curva brachystochrona, debebit tempus descensus, quod per $\int ds/\sqrt{x}$ exprimitur, esse minimum.* Sense I.2: an infinitesimal increment (`elementum temporis`), collocated with `elementum` or `differenziale`, the differential dt .
- **Branch II — temporal reference and historical period.** Sense II.3: a point or epoch, in adverbial genitive or ablative (`tum temporis`, `hoc tempore`). Sense II.4: a historical era or generation, with demonstratives (`illius temporis`) to locate the state of a science or a scholar's lifetime. Pivotal quotation: E101, Bl. 140, `sent_idx 9` — *Methodum autem, qua summus illius temporis Geometra est usus...*

Retrieval that treated `tempus` as one vector would mix the brachistochrone with a historian's "in those days". The entry keeps them apart by rule.

vis (lemma `vis`, Latin noun). Corpus frequency 8,933 tokens across 298 documents — 6,418 Latin sentences in 284 of them; five senses in two branches.

- **Branch I — physical and mechanical force.** Sense I.1: an active influence (gravity, central attraction) causing motion. Rule: often a genitive of origin (`vis gravitatis`, `vis centripetae`); direction by participle (`versus centrum tendente`). Quotation: E1, E1, Bl. 0, `sent_idx 0` — the ordinary cycloid as tautochrone, *vi gravitatis uniformiter versus centrum infinite distans tendente*. Sense I.2: a reactive or component force of curvilinear motion (`vis normalis`, `vis centrifuga`, `vis tangentialis`), almost always paired with an orthogonal counterpart. Sense I.3: a general mechanical load, often in the plural *vires*, modelled as functions of position or mass and subjected to integration.
- **Branch II — mathematical and logical efficacy.** Sense II.1: the analytical power of a method (`vis Analyseos`, `vis methodi`), object of *superare* or *perspiciatur*. Quotation: E101, Bl. 84 (S. 85), `sent_idx 2` — *ut methodi meae vis et praestantia...* Sense II.2: the logical force of an equation (`vi relationis`, `vi aequationis`), "by virtue of".

The same lemma, in the same author, already splits into a mechanical game and a methodological one. Section 5 will add Kepler and Newton to the mechanical branch and show that *that* branch is not one game either.

CONCEPT `usage_rule`

A usage rule states how a sense is played: typical constructions, contrasts, and governing verbs. It is the OED's implicit grammar made explicit. A sense without a usage rule is a label; a sense without a quotation is a conjecture.

National dictionaries occupy the same role as the OED, for other languages and other offices: Duden for contemporary German orthography and usage; Le Robert and the Dictionnaire de l'Académie for French; Treccani for Italian; Ma'agarim for historically attested Hebrew; Grimm for historical German. None of their

copyrighted articles is stored in the reference library. What is stored is their authority as metadata: a verdict against a documented sense cites the office, not a pirated entry. The authority remains movable in the sense of ÜG §§96–97.

4 From page to grammar

A scientific page does not arrive as an OED entry. It arrives as an image. The pipeline that turns the image into a citable rule is the PDAI processing chain, specialised at the last stage for LWgrammar.

OPERATION From page to usage rule		
Q How does a printed page become a sense with a usage rule?		
1	ocr	transcribes the page image into text, with the page remaining the unit of evidence
2	stanza	tokenises, lemmatises, dependency-parses, and tags named entities into <code>stz_sentences</code> , <code>stz_tokens</code> , <code>stz_deps</code> , <code>stz_entities</code>
3	gazetteer	marks terms that already belong to a domain inventory and therefore earn an entry
4	lw-slips	copies every sentence using the term onto a slip, with neighbouring words and dependency context
5	lw-compile	groups slips into branches and senses, writes definition and <code>usage_rule</code> , stores the entry with quotations still addressed to (<code>doc_id</code> , <code>pid_x</code> , <code>sent_idx</code>)
→ An LWgrammar entry in the collection database, revisable when new slips arrive.		

The linguistic layer is necessary and not sufficient. Stanza [15] supplies lemma, part of speech, and dependency context as slip features. Named-entity recognition, however, is not a grammar. In kepler-prim the surface *Anima* is tagged PER on kgw-06 Bl. 265 (S. 276): the tagger has treated a cosmological soul as a person. The mistake is instructive. A pipeline that stopped at NER would file the passage under biography. The rule of use — *anima* as motive soul of the world in 1596, still present in the *Epitome* after *virtus motrix* has entered — lives in the entry, not in the tag.

Qualification for an entry follows two paths, unioned. Gazetteer membership is sufficient: a term already inventoried as astronomical, geometrical, mechanical, or philosophical earns an entry because its rule of use is what scholarly search needs. Independently, a both-gate admits a lemma that has been used often enough in the corpus (`corpus_freq` ≥ 10) and asked after often enough (`search_count` ≥ 10). The thresholds are editorial policy, not metaphysics. They keep the card folder from filling with hapax noise while still letting repeated scholarly interest promote a term.

Compilation copies the OED's office practice. Readers (here: the corpus) copy telling sentences onto slips (`lw_slip`). An editor (here: a language model under an editorial brief) groups the slips, drafts branches and senses, and writes a usage rule for each. When new pages enter, only the new slips are handed back with the existing entry; the revision is logged (`lw_revision`). The entry is a living folder, not a pronouncement. Disagreement has an address: every quotation points at a page and a sentence.

Beside the per-collection databases sits a workspace-wide reference library, `LWgrammar/`, whose DuckDB store [16] holds the Wittgenstein sections used above at paragraph-accurate anchors (PU 693 sections, *Über Gewißheit* 676, *Tractatus* 526 propositions, *Zettel* 717), together with the measurement-standard and dictionary-authority rules of Section 2. Full-text search over those sections is BM25 via DuckDB's FTS extension. Connections are short-lived and read-only: open, read, close. The digital texts are the working copies of the Wittgenstein Project; the citation editions remain Suhrkamp [7], [13]. Part II of the *Investigations* is absent from the digital file used here — the file ends at §693 — and is not silently treated as present.

5 Microscopic semantic history

Who used *anima*, *vis*, *lex naturae*, and when, with which meaning? The answers below are not impressions. They are sentences retrieved from Stanza databases of the PDAI collections kepler-prim, enestroem, and natural-philosophy, cited as `doc_id pidx / sent_idx`. The editions themselves are Kepler's *Gesammelte Werke* [17], Euler via the Eneström numbers [18], and Newton's *Principia* of 1687 [14].

5.1 Case A — Kepler: from soul to force, without substitution

In 1596 Kepler's *Mysterium cosmographicum* (KGW I) still speaks the language of motive souls. Two adjacent sentences on kgw-01 Bl. 83 (S. 70) (`sent_idx 9–10`) offer either several motive souls or one at the centre of all orbs, in the Sun:

...aut Motrices animas... aut, vnam esse motricem animam in orbium omnium centro, scilicet in Sole...
...ita nunc vita, motus et anima mundi in eundem Solem recidit...

By 1609 the *Astronomia nova* (KGW III) has a different leading term. On kgw-03 Bl. 43 `sent_idx 24` the motive power is *virtus motrix*, located in the body of the Sun, with physical arguments attached:

...virtutem motricem seu morarum dispensatricem esse in corpore Solis: Accedunt argumenta Physica.

On Bl. 44 `sent_idx 0–1` that power is said to receive quantities plainly as light does:

Virtutem motricem plane ut Lucem, recipere quantitates...

The new game is analogical: force on the model of radiation.

The later *Epitome astronomiae Copernicanae* (KGW VII, 1618–21) does not complete a replacement. On kgw-07 Bl. 295 (S. 299) `sent_idx 8` the Sun still *animam in se habeat*. On Bl. 302 `sent_idx 3` Kepler can list, in one breath, intelligences, motive souls, and solid orbs:

...de intelligentijs, animabus motricibus, et orbibus solidis?

The point is not a synonym swap. A new language-game (*virtus motrix* as light-like power) is played *beside* the old one (*anima* as motive soul). Family resemblance and rule-change are datable on the page. Reading the 1596 sentences with Newton's grammar of *vis* violates the temporally scoped rule R-KEPLER-ANIMA-VIS: Kepler in 1596 does not yet play that game, and *anima* has not left the later work.

5.2 Case B — *lex naturae* as an emergent phrase

A collocation is not a co-occurrence, and a co-occurrence is not a rule. In kepler-prim, the lemmata *lex* and *natura* share a Latin sentence **45 times**, across 16 documents. A proximity search narrows this to **23 sentences** within five tokens and **15** within one (10 and 9 documents respectively). The compiled OED entry for *lex natura* in kepler-prim spans two branches and rests on 16 proximity-confirmed quotations from across the *Gesammelte Werke*.

Every one of those figures depends on decisions that the figure itself does not display, and we state them because two of them move the number by more than a factor of six.

A sentence is addressed by the triple (document, page, sentence). The sentence index restarts on each page, so a key built from document and sentence alone merges sentences that merely share an index on different

pages: the same-sentence count then reads **313** instead of 45. That is not a stricter or looser criterion, it is a wrong one, and it inflates precisely the quantity a philological claim would rest on.

Second, what counts as an instance of *natura*. Admitting the adjective *naturalis* as a variant of the noun yields **54** same-sentence hits, 27 within five tokens and 16 within one. Those are the figures an earlier draft of this paper reported, and they were not miscalculated: they follow exactly from the variant list that draft named. The move is not obviously illegitimate either, since it is how one folds *naturae* or the OCR artefact *naturâa* into the lemma. But *naturalis* is a different word with a different grammar, and admitting it counts sentences in which Kepler speaks of natural things among sentences about nature's law. A defensible-looking definition, stated openly, still produced a fifth more evidence than the claim can carry.

Third, language. kepler-prim is a bilingual edition — 325,702 sentences tagged Latin against 137,435 tagged German — so one expects a language filter to matter. Here it does not: all 45 co-occurrences fall in Latin sentences. The expectation was reasonable and the data declined it, which is the ordinary fate of plausible methodological assumptions and the reason for recording the language rather than assuming it.

The lemma census therefore stores the sentence language and the counting mode alongside every figure, including the rejected one, so that a reader can reproduce the 54 as readily as the 45. A figure that cannot state its own definition is not a finding.

But the compiled entry reveals something the raw co-occurrence count does not. Kepler's *lege Naturae* is an **ablative construction**. In *De Stella Nova* (1606; KGW I, kgw-01 Bl. 346 (S. 336) sent_idx 3) he faults those who ascribed every celestial event to the free choice of the gods: had they weighed from the start *ista omnia rata lege Naturae circumire* — “that all these things come round by a fixed law of Nature” — they would never have fallen into such credulity. On the next page (Bl. 347, S. 337, sent_idx 1) a spirit is *quadam naturae lege connexus*, “bound by a certain law of nature”, to men and to the world. The ablative places *lex* as the manner or bond of a process, and its opponent is *arbitrium liberum*, not falsehood. Nature's law is that by which motion runs orderly; Kepler does not name it and set it on the table to be conserved.

Euler does exactly that. Eneström E88, *Nova theoria lucis et colorum* (1744), Bl. 27 (S. 28) sent_idx 5, rejects a rival account of refraction because *amplissima quaedam et constantissima lex naturae non adimpletur* — “a most comprehensive and most constant law of nature is not fulfilled”. Four sentences later his own theory is commended because *lex naturae gravissima conservatur*. Here the phrase stands in the nominative as the subject of fulfilment and conservation. E57, *Inquisitio physica in causam fluxus ac refluxus maris* (1739), Bl. 1 (S. 20) sent_idx 3, shows the same standing in another case: *cum ex legibus naturae satis superque constet* — the laws are the premisses from which a conclusion follows. One may argue *from* a law only if the law is a proposition. Whether nominative or ablative of premiss, Euler's *lex naturae* is something that can be true, violated, upheld, and reasoned from.

This is the rule difference, and it is grammatical before it is doctrinal: Kepler's *lege Naturae* states the manner in which things happen; Euler's *lex naturae* names a proposition that bodies obey or violate. Two usage rules, not one concept. A system that answers “Kepler on the law of nature” by retrieving the 45 co-occurrence sentences — or by finding embedding neighbours of *lex* and *natura* — returns evidence that fits neither rule. A system that compiles the collocation, distinguishes ablative from nominative constructions, and reads Kepler's grammar against Euler's, understands the difference. The compiled OED entry is the understanding (R-LEX-NATURAE-ATTEST).

■ 5.3 Case C — *vis motrix* across three authors

The mechanical branch of *vis* is still not one rule.

- **Kepler**, kgw-03 Bl. 44 sent_idx 0: *virtus motrix* on the model of light — analogical, radiating from the solar body.
- **Newton**, newton-principia-1687 Bl. 11 sent_idx 6: *vis motrix ex vi acceleratrice ducta in quantitatem ejusdem materiæ* — *vis motrix* taken from the accelerating force multiplied into the quantity of matter. The rule is definitional and relational: one quantity is the product of two others.

- **Euler**, *Briefwechsel mit J. A. von Segner* (Opera Omnia IVA 8, doc 978-3-319-75942), Bl. 564 (S. 551) sent_idx 24: *Vis motrix est caussa, quae per se semper producet motum* — a causal definition. Force is the cause that of itself always produces motion.

Same lemma, three rules. But the Segner letter does something more interesting than supply a third definition: Euler states the rule of use himself, and states what it excludes. Three sentences on (sent_idx 27) he writes that the character of motive force

...wird von der Bewegung her genommen, und kan, so weit ich sehe, von nichts andern hergenommen werden, so daß derjenige, der nicht wüste was eine Bewegung ist, unmöglich wissen könnte, was man in der Mechanic eine Kraft nennet; er würde die selbe von einer der Kräfte seiner Seele nicht unterscheiden.

Whoever does not know what motion is cannot know what is called a force in mechanics — *he would not distinguish it from one of the powers of his soul*. This is a usage rule given in the source: the mechanical game is fixed by its tie to motion, and its boundary runs exactly where Case A found Kepler’s difficulty. What LWgrammar formalises as `usage_rule`, Euler here writes out in German while defining in Latin.

The partition is not by author, though. In the *Correspondance* of Opera Omnia IVA 7 (doc 978-3-7643-8743-3, Bl. 382 sent_idx 8) Euler writes *vis motrix est ut productum ex massa in celeritatem* — Newton’s relational rule, not the causal one. Euler defines causally where he settles what mechanics is about, and relationally where he computes an impact. So a terminology-compare that reports “all three use *vis motrix*” has understood none of them, and one that reports “Euler defines it causally” has understood half of Euler. The LWgrammar entry for Euler’s *vis* (Section 3) already splits mechanical from methodological senses; the present case splits the mechanical sense again — by rule, not by author.

A further warning sits in the lemma layer itself. *visus* (seeing) is a different lemma. In kepler-prim it occurs in 264 Latin sentences. A string search for `vis` will harvest it. Lemma normalisation is required, and even then it does not disambiguate Kepler’s analogical *virtus* from Newton’s product from Euler’s cause. Only the rule does.

METHOD How a case cites

Every historical claim above names a collection document, a zero-based page index, and a sentence index. Disagreement is an invitation to open that page, not a dispute about an impression.

6 The ai-dictionary: scoped senses and submitted rules

The cases of the previous section make a structural demand. If Kepler’s *lege Naturae* and Euler’s *lex naturae* are two rules and not one concept, then the dictionary that records them cannot have one article per lemma. It must have one article per lemma *and scope*: the same headword carries different senses for different authors, works, and periods, and each sense carries its own rule of use. The workspace reference library described in Section 4 has been extended into exactly that object. We call it the **ai-dictionary**: a dictionary category whose entry is not “what the word means” but “what the word means *for this author, in this work, in this period, in this context of use*” — the OED’s historical principle, sharpened into a queryable scope.

6.1 The entry form

Four tables carry the category. `term_entry` holds the headword with lemma and language. `term_sense` holds the scoped senses: each row names its author, work, period, and context of use, gives the definition, and writes the rule of use out as text. `sense_locus` holds the attestations, and here the PDAI reference policy is not optional decoration but the schema itself: a locus is the quadruple (collection, document, page index, sentence index). The page index is the *physical* scan (`page_<pid>`, the leaf); the entry’s readable `ref` names the number actually printed on that leaf and declares the physical index beside it — S. 277 [Bl. 288] —

falling back to the declared sheet alone (Bl. 302) only where the running head carries no legible number. A citation never lets the physical index stand in for the printed page without saying so. `usage_rule_entry` holds the rules that are not tied to one headword — the grammatical and stylistic rules against which formulations are judged — each with a tier (blocking or advisory), a statement, positive and negative examples, and a status.

The *lex naturae* case of Section 5 lives in this form. The Kepler sense (`lex-natura#kepler#1`, scoped to 1600–1630) states the ablative rule — nature’s law as the manner in which processes run, opposed to *arbitrium liberum* — and carries six harvested loci from the *Gesammelte Werke*: two in KGW I (kgw-01, Bl. 190 → S. 180; Bl. 288 → S. 277), two in the *Epitome* volume (kgw-07, Bl. 77 → S. 103; Bl. 406 → Bl. 406), and two in the correspondence volume kgw-21-1. A second, deliberately distinct sense records the juridical-moral use of the natural-law tradition, so that a passage of that register is not forced under Kepler’s physical rule. The scoping answers the retrieval failure this paper began with: a query for “Kepler on the law of nature” resolves against the Kepler sense and its loci, not against a lemma-wide average.

Two properties of the store follow the riverbed dynamic of Section 2 rather than convenience. First, senses and rules carry a status (active, revised, revoked) and a source; a correction revises, it never deletes, and a revoked rule remains visible as revoked. Second, compilation is layered: the seed layer of entries is compiled from declared source files and can be recompiled at any time, while submitted rules and harvested loci survive recompilation untouched. The dictionary is a living folder in the OED’s sense, with the revision log as part of the object.

6.2 The submission path

A dictionary that admits new rules only at compile time is an archive. The ai-dictionary has a live path: a rule observed in practice — typically a correction by the author — is submitted with a statement, examples, scope, and tier, receives an identifier, and is active from that moment for every consumer of the store.

The founding submission is the truth-bearer rule:

RULE R-GG-001 (SUBMITTED)

Sachverhalte und Tatsachen — states of affairs and facts — are not true or false. They are parts of the world: a state of affairs obtains or does not obtain. Only Aussagen — propositions, what a declarative sentence expresses — are true or false. “This fact is true” is therefore not an emphatic assertion but a grammatical error. Anchor: TLP 2; PU §§134–137.

The rule is Wittgensteinian in content and in kind: it is a grammatical sentence in the sense of PU §251, defended by the negation test (“this fact is false” is not a weaker claim but a confusion of categories), and it is *submitted*, dated, and scoped like every other rule — the author’s authority is recorded as a source, not presumed as an oracle. At the time of writing the store holds 21 entries, 23 scoped senses, and 21 usage rules (16 blocking, 5 advisory), together with seven review gates that stage the rules into a delivery protocol.

6.3 From annotation to rule

The submission path becomes productive when it is wired into the editing loop. The ai-lektorat annotation viewer (port 5114) operates over any Quarto paper and provides an inline track-changes interface: the annotated paragraph is replaced by a revision block with changed spans rendered as `<del>/<ins>` hunks, each decidable independently. A three-pane picker — lector family, lector, scripted instruction or free instruction — selects the reviewing stance (Figure 1); a segmented mode control switches between *diagnose*, *propose*, and *frei*. The model is always the current member of the configured model family (at the time of writing: `gemini-3.8-flash` via the `gemini` family entry in `standards/model-routing.yaml`; the string `gemini-3.8-flash` no longer appears anywhere in the codebase).

Wittgensteinian Grammar and OED Dictionaries for AI

A foundation of semantic understanding of scientific texts

AUTHOR

Gerd Graßhoff  

AFFILIATIONS

Humboldt-Universität zu Berlin

Deutsches Museum

ABSTRACT

Semantic understanding of scientific texts is conformity to a rule of use, not similarity of tokens. What an expression means is shown by the grammar of its employment; whether a formulation hits that meaning is decided by whether it follows a rule or contradicts it — and a violation can be stated only by naming the specific rule whose authority is breached. This paper develops that claim from Wittgenstein's later philosophy (rule-following, the standard metre, perspicuous representation) and from the entry form of the Oxford English Dictionary. It then shows how, after OCR and linguistic analysis, those rules can be compiled as LWgrammar entries from a declared corpus and consulted as a reference library. The argument is carried by cases: length without a measurement standard; Kepler's *anima* beside *virtus motrix*; the phrase *lex naturae* attested in Euler as a nominative proposition and in Kepler as an ablative mode; three grammars of *vis motrix* in Kepler, Newton, and Euler. A formulation score operationalises the thesis: no verdict of violation without a cited rule and an anchor. The instrument is built and measured: an ai-dictionary whose senses are scoped by author, work, and period, with every attestation addressed to a page and a sentence; a rule-submission path with a revision lifecycle; a conformity check that assigns each rule the verdict engine its content demands (dependency syntax, lexis, or natural-language inference), calibrated to 16/16 on a seeded golden test; and a logical-relation scan that finds, among the 663,972 sentences of an archival collection, those that imply or contradict a given sentence. Meaning becomes a citable inventory of rules, scoped by time, person, and discipline.

KEYWORDS

Wittgenstein, rule-following, Oxford English Dictionary, LWgrammar, historical scientific Latin, semantic conformity, Kepler, Euler

The problem: token similarity is not understanding

LEKTORAT ANNOTATION

Multi

Wide

Collapse

No segment selected (click a ¶ icon or select text)

FAMILY	LECTOR	INSTRUCTION
Composition	Stil	Register break — marketing or tutorial voice (DEF-P04)
Language	Terminologie	Nominalisation load (DEF-P03)
Apparatus	Copyedit	Uniform hedging (DEF-P02)
Presentation		Buried topic sentence
Chief		
Terminology		

Mode

Diagnose

Propose

Frei

stil - gemini-3.8-flash

INSTRUCTION (TYPE OR DICTATE)

What should the lector check or revise?

SOURCE TEXT (TRIMABLE)

Selected passage or full segment...

Scope

segment

Engine

gemini-3.8-flash

☒ Also register as LWgrammar rule after acceptance

▶ Execute revision

Figure 1 The annotation panel with three-pane picker (Lector family | Lector | Instruction), mode control, and “Also register as LWgrammar rule” toggle pre-set by the classify step.

Before the revision runs, the server fetches all active blocking rules from the LWgrammar MCP (`lw_usage_rules scope=all tier=blocking`) and prepends them as a mandatory context section ahead of the lector skill. The lectors that own rules in the store (`bibliothekar`, `terminologie`) receive additionally filtered pulls for their own registered rules. The outcome line in the sidebar reports `✦ LWgrammar rules consulted` or, when the MCP is unreachable, `⚠ LWgrammar not consulted`.

The accepted hunks compose a merged text client-side (Figure 2); “Create version” posts the result as a new version with provenance (`POST /api/versions`, `origin accepted`).

what does it mean to *follow* a meaning-rule or to *contradict* one — and how, after a page has been read by OCR and parsed linguistically, can one build an instrument that makes that judgement checkable on scientific texts?

Five contributions follow from that question:

1. **Rule-following as a measure of conformity.** A formulation is semantically in order when it follows an identifiable rule of use; a violation is reportable only with the rule named and cited.
2. **The OED entry as perspicuous representation.** The dictionary that begins from dated quotations and lets the definition summarise the evidence is the lexicographical form of Wittgenstein's *übersichtliche Darstellung*.
3. **LWgrammar.** The same entry type, compiled from one declared corpus after OCR and Stanza analysis, with the rule of use written out as *usage_rule*.
4. **A casebook of microscopic semantic history.** Who used *anima, vis, lex naturae* when, and under which rule — with every quotation anchored at `(doc_id, pidx, sent_idx)`.
5. **Context-specific judgement.** Time, person, and discipline scope the same engine; publication-format profiles (the analogue of a LaTeX document class) configure style dimensions without replacing the rule inventory.

The register of what follows is immanent work on the *Philosophische Untersuchungen* and *Über Gewißheit*, paired with a built instrument and with corpus evidence that can be opened at a page. Secondary literature on Wittgenstein is not rehearsed. Companion papers carry what this one assumes: the PDAI page format (Graßhoff 2026b), the linguistic index (Graßhoff 2026e), the dictionary as a scholar's operation (Graßhoff 2026a), the Partitur as the score behind an agent's answer (Graßhoff 2026f), and the golden-query testbed that holds those operations to a declared oracle (Graßhoff 2026d). Computational epistemology treats the validity of *arguments* (Graßhoff 2026c); this paper treats the grammar of *words*. The two tasks are neighbours, not substitutes.

Revision V3

✓ Accept all

✗ Reject all

Create version

Apply to .qmd

The register of what follows This is paper immanent works work immanently on the Philosophische Untersuchungen and Über Gewißheit, paired pairing them with a built instrument and with corpus evidence that can the bereader opened can open at a page. Secondary literature does not Wittgenstein rehearse is secondary not literature rehearsed on Wittgenstein. Companion papers carry what this one assumes: the PDAI page format (Graßhoff 2026b), the linguistic index (Graßhoff 2026e), the dictionary as a scholar's operation (Graßhoff 2026a), the Partitur as the score behind an agent's answer (Graßhoff 2026f), and the golden-query testbed that holds those operations to a declared oracle (Graßhoff 2026d). Computational epistemology treats the validity of arguments (Graßhoff 2026c); this paper treats the grammar of words. The two tasks are neighbours, not substitutes.

LEKTORAT ANNOTATION

Multi

Wide

Collapse

DEF-P08

MINOR

Dead abstraction

The phrases 'corpus evidence that can be opened at a page' and 'Secondary literature on Wittgenstein is not rehearsed' use agentless passive constructions.

→ Attribute the actions to explicit agents: let the paper refrain from rehearsing secondary literature, and let the reader open the corpus evidence.

Revision ready — accept/reject hunks inline, then Create version. + LWgrammar rules consulted.

VERSIONS

V0

V1

V2

V3

Diff vs.

previous

LLM stil 2 findings candidate 2026-09-07 12:06

What follows pairs immanent work This paper works immanently on the Philosophische Untersuchungen and Über Gewißheit, pairing them with a built instrument and with corpus evidence that the reader can be opened open at a page. Secondary It does not rehearse secondary literature on Wittgenstein is not rehearsed: Wittgenstein. Companion papers carry what this one assumes: the PDAI page format (Graßhoff 2026b), the linguistic index (Graßhoff 2026e), the dictionary as a scholar's operation (Graßhoff 2026a), the Partitur as the score behind an agent's answer (Graßhoff 2026f), and the golden-query testbed that holds those operations to a declared oracle (Graßhoff 2026d). Computational epistemology treats the validity of arguments (Graßhoff 2026c); this paper treats the grammar of words. The two tasks are neighbours, not substitutes.

VERSION TEXT (EDIT TO REFINE)

This paper works immanently on the Philosophische Untersuchungen and Über Gewißheit, pairing them with a built instrument and with corpus evidence that the reader can open at a page. It does not rehearse secondary literature on Wittgenstein. Companion papers carry what this one assumes: the PDAI page

Save as V+1

✓ Apply to .qmd

Reject

Figure 2 Inline revision block with word-level hunks. Each hunk carries a ✓/✗ decision; “Accept all” composes the merged text; “Create version” stores it and, when the generalise toggle is on, triggers the rule-proposal step.

When the user accepts a revision and the “Also register as LWgrammar rule” toggle is on, the server runs the *rule-propose* role (a small Gemini call) to extract a generalised rule statement, positive and negative examples, a tier, a scope, and a detector specification from the accepted change. A compact rule card appears in the sidebar (Figure 3). One click submits the rule via `lw_submit_rule` with provenance evidence (paper, segment identifier, annotation identifier, quoted passage); the returned rule code is stored on the annotation record.

what does it mean to *follow* a meaning-rule or to *contradict* one — and how, after a page has been read by OCR and parsed linguistically, can one build an instrument that makes that judgement checkable on scientific texts?

Five contributions follow from that question:

- 1. Rule-following as a measure of conformity.** A formulation is semantically in order when it follows an identifiable rule of use; a violation is reportable only with the rule named and cited.
- 2. The OED entry as perspicuous representation.** The dictionary that begins from dated quotations and lets the definition summarise the evidence is the lexicographical form of Wittgenstein's *übersichtliche Darstellung*.
- 3. LWgrammar.** The same entry type, compiled from one declared corpus after OCR and Stanza analysis, with the rule of use written out as `usage_rule`.
- 4. A casebook of microscopic semantic history.** Who used *anima, vis, lex naturae* when, and under which rule — with every quotation anchored at `(doc_id, pidx, sent_idx)`.
- 5. Context-specific judgement.** Time, person, and discipline scope the same engine; publication-format profiles (the analogue of a LaTeX document class) configure style dimensions without replacing the rule inventory.

The register of what follows is immanent work on the *Philosophische Untersuchungen* and *Über Gewißheit*, paired with a built instrument and with corpus evidence that can be opened at a page. Secondary literature on Wittgenstein is not rehearsed. Companion papers carry what this one assumes: the PDAI page format (Graßhoff 2026b), the linguistic index (Graßhoff 2026e), the dictionary as a scholar's operation (Graßhoff 2026a), the Partitur as the score behind an agent's answer (Graßhoff 2026f), and the golden-query testbed that holds those operations to a declared oracle (Graßhoff 2026d). Computational epistemology treats the validity of *arguments* (Graßhoff 2026c); this paper treats the grammar of *words*. The two tasks are neighbours, not substitutes.

V4 (accepted) ✓ Accept all ✗ Reject all Create version Apply to .qmd

This paper works immanently on the *Philosophische Untersuchungen* and *Über Gewißheit*, pairing them with a built instrument and with corpus evidence that the reader can open at a page. It does not rehearse secondary literature on Wittgenstein. Companion papers carry what this one assumes: the PDAI page format (Graßhoff 2026b), the linguistic index (Graßhoff 2026e), the dictionary as a scholar's operation (Graßhoff 2026a), the Partitur as the score behind an agent's answer (Graßhoff 2026f), and the golden-query testbed that holds those operations to a declared oracle (Graßhoff 2026d). Computational epistemology treats the validity of arguments (Graßhoff 2026c); this paper treats the grammar of words. The two tasks are neighbours, not substitutes.

LEKTORAT ANNOTATION Multi Wide Collapse

Scope: segment Engine: gemini-3.8-flash

☒ Also register as LWgrammar rule after acceptance

rule candidate

Execute revision

Annotation saved: 27af320110f680d9
Run started (gemini-3.8-flash)
✓ Stage: segment-review (still)
✓ Stage: segment-propose
✓ Run complete (ok)

LWGRAMMAR RULE PROPOSAL R-GG-010

Statement: State scholarly actions in the active voice with an explicit agent instead of using impersonal passive constructions.

✓ This paper does not rehearse secondary literature on the topic.
✗ Secondary literature on the topic is not rehearsed.

Confirm & Submit Dismiss

DEF-P03 MINOR Nominalisation load

The opening sentence relies on a copular predicate and nominalisation ('The register of what follows is immanent work on...') instead of expressing the action directly through an active verb.

→ Recast with an active verb governed by the paper: 'This paper works immanently on the *Philosophische Untersuchungen* and *Über Gewißheit*, pairing them with...'

DEF-P08 MINOR Dead abstraction

The phrases 'corpus evidence that can be opened at a page' and 'Secondary literature on Wittgenstein is not rehearsed' use agentless passive constructions.

→ Attribute the actions to explicit agents: let the paper

Figure 3 Rule card after acceptance of an active-voice revision: the card shows the proposed rule statement, a positive and a negative example, and the assigned code R-GG-010. The “Confirm & Submit” button records the rule in the LWgrammar store via the MCP write path.

The worked case is R-GG-002 — Wittgenstein citation grammar. The failing form, which citeproc produced from `[@wittgensteinPU2022, PU §201]`, was (Wittgenstein 2022, sec. 201) when written as `[@wittgensteinPU2022, §201]`: the CSL locator label “sec.” rendered where the siglum § was expected. The annotation identified this as a *bibliothekar* issue. On acceptance the server proposed the rule:

Wittgenstein’s works are cited by work siglum and section number (PU §201, ÜG §96), never by author-year and locator label (“sec. 201”).

The detector `{"kind": "lexical", "markers": ["sec. \\d", "PU", "ÜG", "Wittgenstein"]}` was registered alongside the rule. From that point forward, any revision involving Wittgenstein passages has the rule injected into the prompt context, and `lw_rule_conformity` can flag the violation in new text without a full revision run (Figure 4). The source fix — `[@wittgensteinPU2022, PU §201]` — renders cleanly as “(Wittgenstein 2022, PU §201)” in both HTML and PDF output.

game. The example is “Every rod has a length”:

Beispiel: »Jeder Stab hat eine Länge«. Das heißt etwa: wir nennen etwas (oder dies) »die Länge eines Stabes« — aber nichts »die Länge einer Kugel«.

(Wittgenstein 2022, sec. 251)

The negation “Some rods have no length” is not a false physical hypothesis. It is nonsense in this game: it withdraws the very technique by which “length of a rod” is said. That is the test this paper will use on scientific prose. A formulation that drops the standard by which a quantity is defined does not make a weaker empirical claim. It leaves the game in which that quantity can be said.

The standard metre: a role, not a property

PU §50 is the load-bearing example, and the wording matters. Wittgenstein writes *aussagen*, not *sagen*:

Man kann von einem Ding nicht aussagen, es sei 1 m lang, noch, es sei nicht 1 m lang, und das ist das Urmeter in Paris. — Damit haben wir aber diesem natürlich nicht irgend eine merkwürdige Eigenschaft zugeschrieben, sondern nur seine eigenartige Rolle im Spiel des Messens mit dem Metermaß gekennzeichnet.

(Wittgenstein 2022, sec. 50)

One cannot predicate of the Paris prototype that it is, or is not, one metre long. The sentence would look like a measurement. It is a stipulation of the game. The prototype has no occult property; it has a role. All other length-talk in that game is a quantitative multiple of that role, on the basis of a measuring procedure.

Revision V1

✓ Accept all ✗ Reject all Create version Apply to .qmd

One cannot predicate of the Paris prototype that it is, or is not, one metre long. ~~The~~ ~~That~~ ~~is~~ sentence ~~would appear~~ ~~to~~ ~~look~~ ~~to~~ ~~be~~ ~~reported~~ ~~as~~ a measurement. ~~It~~, but ~~it~~ ~~is~~ stipulates ~~a~~ ~~stipulation~~ ~~rule~~ of the game. The prototype has no occult property; it ~~has~~ ~~plays~~ a role. All other ~~length~~ ~~talk of length~~ in that game ~~is~~ ~~a~~ ~~quantitative~~ ~~multiple~~ ~~quantitatively~~ ~~of~~ ~~multiplies~~ that role, ~~on the basis of~~ ~~a~~ ~~when~~ ~~measuring~~ ~~one~~ ~~procedure~~ ~~measures~~.

LEKTORAT ANNOTATION Multi Wide Collapse

The clauses ‘would look like a measurement’, ‘is a stipulation’, and ‘is a quantitative multiple of that role, on the basis of a measuring procedure’ convert active processes into abstract nouns paired with copular or linking verbs.

→ Recast nominalisations (‘measurement’, ‘stipulation’, ‘measuring procedure’) into active verbal constructions to improve clarity and force.

DEF-P08 MINOR Dead abstraction

The phrase ‘on the basis of a measuring procedure’ invokes a measurement process without an agent executing the procedure.

→ Specify the agent or render the clause actively (for instance, ‘when one measures’).

Revision ready — accept/reject hunks inline, then Create version. ✦ LWgrammar rules consulted.

VERSIONS V0 V1 Diff vs. previous

LLM stil 2 findings candidate 2026-09-07 12:09

One cannot predicate of the Paris prototype that it is, or is not, one metre long. ~~The~~ ~~That~~ sentence ~~would look like~~ appears to report a measurement. ~~It is~~ measurement, but it stipulates a stipulation rule of the game. The prototype has no occult property; it ~~has~~ plays a role. All other ~~length-talk~~ talk of length in that game ~~is~~ a quantitative multiple of quantitatively multiplies that role, on the basis of a measuring procedure: role when one measures.

VERSION TEXT (EDIT TO REFINE)

One cannot predicate of the Paris prototype that it is, or is not, one metre long. That sentence appears to report a measurement, but it stipulates a rule of the game. The prototype has no occult property; it plays a role. All other talk of length in that game quantitatively multiplies that role when one measures.

Save as V+1 ✓ Apply to .qmd Reject

Figure 4 A later revision of a passage about the Paris prototype runs with R-GG-010 injected. The sidebar reports “ LWgrammar rules consulted” and the lector flags DEF-P08 Dead abstraction — passive nominalisations — as a direct consequence of the active-voice rule in the store.

The loop from correction to enforceable rule involves four components that are now wired together: the annotation viewer, the revision engine, the MCP write path, and the conformity checker. The store is the stable centre; the viewer and the checker are consumers on opposite sides of the edit cycle.

6.4 Services

The ai-dictionary and the annotation-rule loop are exposed as two services, both accessible tailnet-wide (Tailscale):

LWgrammar MCP (<http://127.0.0.1:5116/mcp>, streamable HTTP, launchd `com.ai-platform.lwgrammar-mcp`): 16 tools in four families — dictionary (`lw_term`, `lw_usage_rules`, `lw_submit_rule`, `lw_revoke_rule`), modes (`lw_modes`, `lw_mode`, `lw_sprachformen`, `lw_situate`, `lw_gate`, `lw_application_examples`), consult/lector (`lw_consult_understanding`, `lw_lector`, `lw_rule_conformity`), and corpus reference (`lw_search`, `lw_cite`). Reads are served from any Tailscale node; writes (`lw_submit_rule`, `lw_revoke_rule`) execute only on mac-studio.

ai-lektorat viewer (<http://127.0.0.1:5114>, launchd `com.ai-platform.lektorat`): annotation, revision, version, classify, rule-propose, curate-publish, render, and apply endpoints. The viewer is the primary write source for new rule submissions initiated by the author during revision.

6.5 Modes: the same store under different contexts

The context axes of the next section (time, person, discipline, profile) need a practical switch. The store therefore compiles **modes**: named recipes that select, for one kind of task, the registers, rules, and gates that govern it, and emit them as a compact context for a language model. Writing an English conceptual paper, a German Gutachten, and a reading of seventeenth-century Latin are three modes over one store — the inventory does not change, the selection does. A mode is the analogue, on the consultation side, of what the publication profile is on the judging side.

The implemented registry holds five modes: `conceptual-paper-en` (English scholarly prose, covering analytic-core and system-register senses, the full blocking-rule set, and ai-erasmus register passages), `gutachten-de` (the German evaluative register with Gutachten exemplars and verdict-grounding rules), `documentation` (the lucid prose rules with terminology lock for the project scope), `orchestrator-book` (the system-register, action-and-goal, and analytic-core terminology frames together with the chapter-gate protocol), and `historical-reading` (term senses filtered by author, work, and period — the mode under which a query for *lex naturae* in Kepler’s *Astronomia Nova* returns Kepler’s senses and their loci, never the house sense projected backward). Each mode emits a compact deterministic pack — the relevant rules, senses, and exemplary passages, a few kilobytes — suitable for direct injection into a language-model context.

6.6 The consultation hub: three application forms

A mode pack is not a consultation. Three forms of consultation correspond to three distinct practical relationships a language user has with the store.

Writing. The task is to formulate text in the author’s registers. Before drafting, the agent calls the store with the mode name and the list of terms the text will use; it receives back the blocking rules (ordered by tier), the house senses of each planned term with their usage rules and forbidden synonymies, and two or three exemplary passages from the registered register corpus. During drafting, per-term spot checks confirm that each term is introduced in the right scope. After drafting, the self-check gate runs. The pack binds: a blocking rule is a gate, not a suggestion; an advisory rule is deviated from only with a stated reason. This is the institutional form of PU §201: to write correctly under the standard is to know what it means to go against *this* rule, and the pack makes that explicit.

Lectoring. The task is to audit a text — one’s own or another’s — against the standard, with rule-grounded corrective suggestions. The call names the file and optionally a scope; the pipeline returns findings keyed by rule code, each with the offending passage, the violated rule, the corrective sense as a positive example, and a concrete reformulation suggestion. The suggestion is generated by the rule, not by the model’s stylistic preference: what licenses the correction is the cited rule, not an aesthetic judgement. Output follows the ai-lektorat finding contract, so the same findings can drive an annotation viewer or a review-stage gate.

Understanding. The task is to resolve what a term meant in a specific historical or foreign text — not to project the house sense backward. The call names the term, the author, and the work; it receives the sense lattice filtered by that frame, each sense-instance with its attestation loci (`collection`, `doc_id`, `pidx`, `sent_idx`) and real quotations, plus contrast senses showing what the term meant to other authors or in other periods. A missing sense is reported as an explicit gap — the institution offers a provenance-governed catalogue of actual usage, not a closed claim about meaning.

The three forms are not interfaces to three separate systems. The same store, the same rules, and the same loci serve all three; what changes is the direction of the query and the nature of the result. Writing extracts a prescription; lectoring extracts a finding; understanding extracts an attestation.

6.7 Lectoring for maximal linguistic precision

The lectoring form requires a pipeline that is lemma-based, not surface-string-based, for the obvious reason that inflection hides terms. The pipeline runs in five steps.

Segmentation and lemmatization. The text is split into prose paragraphs (Markdown headers, code fences, and YAML front matter are excluded from auditing). Each paragraph is parsed by a stanza pipeline

— language-detected en/de/la models — into sentences with per-token lemma, POS tag, and dependency relation. Term keys are lemmata; *Sachverhalt* and its case forms resolve to the same key.

Term identification. Two candidate channels are measured rather than assumed. The first: noun-phrase head lemmas and lemma n-grams (up to trigrams) over POS patterns, matching against the headword table. This channel is deterministic, free, and complete for terms whose form the store knows. The second: GLiNER2.5 (`fastino/gliner2.5-multi-v1`, the workspace’s query-comprehension authority) with terminology-harvest labels — technical term, philosophical concept, system component, Latin phrase. GLiNER adds recall for terms whose surface form the store does not yet know, and for multi-word technical phrases that stanza n-grams miss. In the benchmark on a 285-line chapter (`action.qmd`), the stanza channel alone found 8 of 9 registered terms; GLiNER found all 9, at the cost of 7 additional false candidates — a union of the two channels recovers full recall at acceptable precision. Term identification is therefore a union by default, with GLiNER candidates promoted to findings only when they also appear in the dictionary.

Batch sense resolution. The deduplicated lemma inventory joins against the dictionary tables in one DuckDB query — term entries, scoped senses, usage rules — returning the applicable rules for each term under the current scope context. This step takes milliseconds regardless of inventory size.

Finding generation. Four detector classes cover the finding space. Rule violations (a term used against a blocking usage rule) are detected by the conformity service described in the next section. Forbidden synonymy fires when two lemmata that are locked to one concept appear in the same text — the *cluster/package/route* class. Sense drift fires when one term is used in senses from two incompatible context scopes across the draft. Gaps are terms without any entry in the store, reported as curation candidates when they appear repeatedly as technical-register nouns (capitalized, NOUN/PROPEN tag, frequency ≥ 2).

Suggestion generation. For each finding, the applicable rule rows — statement, positive example, negative example — are assembled into a concrete, rule-cited suggestion: what to replace, with what, and why. The rule of application, not the model’s taste, licenses the suggestion; the model’s role is limited to phrasing it as a sentence in the text’s register.

The lector pipeline was run against the three orchestrator book chapters (`index.qmd`, `action.qmd`, `justification.qmd`, 940 lines of prose in total) whose motivating terminological errors appear in this paper’s introduction. The conformity engine returned no conflicts on the sentence/proposition class (R-LG-008) or the truth-bearer class (R-GG-001) in the final drafts — the passages that previously contained those constructions had been rewritten. The verification ban (R-LG-007) fired on legacy schema field name occurrences but with the accompanying prose making the exemption explicit; those findings are advisory and the authors chose to retain the field name with the explicit gloss “acceptance check (stored as `verification` for compatibility)”. Gap candidates included `Harness`, `enfolding`, and `frontier`, which are present in the system-register terminology guide and could be promoted to the dictionary in a subsequent curation round.

7 Measuring semantic conformity

7.1 Four verdicts

Given a formulation and a context (time, person, discipline, collection, publication profile), the instrument returns one of four verdicts for each load-bearing expression:

Verdict	Condition
conform	A rule applies, and the formulation follows it (a length <i>with</i> a standard; a term in its documented sense).
violation	A structural miss (length without standard) or a fired violation-group (Kepler + 1596 + Newton’s vis). Always with <code>rule_id</code> and anchor.

Verdict	Condition
innovation	The use departs and is marked as a coinage — a rule in the making.
obsolete	The use follows a daggered rule.

Innovation and obsolete are not assigned automatically in the present deterministic score; they remain editorial. Conform and violation are assigned by matchers that live in the rule table as data (`scope.terms`, `scope.violation_when`), not as buried constants.

7.2 The formulation score

The Partitur `lwgrammar-formulation` is the thesis as infrastructure [4]. Five stages, no language model on the hot path:

1. **formulation-parse** — isolate measures (value + unit), whether a standard or procedure is named, and the observation terms.
2. **rule-locate** — collect applicable rows from `ref_rule` (grammar principles, measurement standards, dictionary conventions, corpus-usage rules, style rules).
3. **reference-cite** — resolve each rule’s anchors to section text (PU §50, ÜG §341, a CGPM resolution, a corpus `source_ref`).
4. **conformity-judge** — emit verdicts. A violation without `rule_id` and an anchor is a *stage failure*, not a soft warning.
5. **formulation-assemble** — write the answer as sentences that quote the rule. If no rule applies, that fact is stated as a status. An empty hit-list is not returned as “nothing found”.

OPERATION Iron Rule	
Q When is a semantic verdict allowed to say “violation”?	
1	formulation-parse finds the quantity 2 m in “Der Stab ist 2 m lang.” and records that no SI, BIPM, or procedure cue is present
2	rule-locate activates R-SI-LENGTH-REF because a measure is present
3	reference-cite resolves the rule’s citation to PU §50
4	conformity-judge emits violation, with <code>rule_id</code> and the anchor; refusing to emit the verdict if either is missing
→ Violation of R-SI-LENGTH-REF (grammar of measuring), anchor PU §50.	

The same score, given “Der Stab ist 2 m lang, gemessen nach der SI-Meterdefinition (BIPM 2019)”, returns **conform** under the same rule. The rule has not changed; the formulation now makes the move the game requires. Kepler 1596 read as Newton’s *vis*, and Kepler credited with *lex naturae*, fire corpus-usage rules whose violation-groups are stored in the library. In each case the answer names the rule. That is PU §201 operationalised: going against a rule is going against *this* rule.

7.3 Context axes and style profiles

The same engine judges differently when the context changes, because `rule-locate` resolves rules under a scope. The scope is not a parameter the caller sets from scratch each time: it is the scope of the active **mode**, which names the combination of context axes appropriate to the task. The axes are:

- **time** — LWgrammar scopes such as `1e1748`; metre definitions 1799 / 1960 / 1983;
- **person** — an author’s compiled entry, not a field-wide average;
- **discipline** — gazetteer domains (astronomy, mechanics, philosophy);
- **collection** — kepler-prim is not enestroem;

- **profile** — a publication-format profile, declared on the Partitur as `stil_profil`, the analogue of a document class.

A mode compiles these axes into a single scope identifier and supplies the rules, senses, and style profile relevant to that scope. The `historical-reading` mode, for instance, sets the person and collection axes to the author and collection in question and suppresses all house-rule blocking that does not apply to a reading context: one does not audit Kepler’s Latin for violations of the English conceptual-paper register.

Style is not a second system. The ai-lektorat defect catalogue (excellence rules and DEF-codes for argument, structure, prose, terminology, bibliography, integrity, figures, examples) is ingested into the same `ref_rule` table, with a dimension path (`stil/komposition/argument`, `stil/formulierung/terminologie`, ...). A finding counts only when the code resolves to a registered rule. The profile `arxivai` does not replace that catalogue. It *weights* it for this venue: British scholarly English, examples after claims, thesis locatable in the abstract and the first section, new terms motivated and explained in plain language before their first load-bearing use (in the `lettreai Typst` theme, as short info boxes), citations complete and not placeholders, figures referenced in the running text, lists set as lists.

The limits of the instrument are the limits of a practice. Entries are relative to a declared corpus. Compilation is editorial and revisable. Scopes are historical. The library is not a meaning-oracle and does not pretend to be. What it offers is stricter and rarer: a place where a disagreement about meaning has an address.

8 Two services, measured

The formulation score of the previous section is deterministic: its matchers live in the rule table as data. Two further services extend the judgement to cases a matcher cannot reach, and both are built and measured rather than proposed. Their shared discipline is the golden test: a set of passages, or sentence pairs, with verdicts fixed *in advance* of any run, against which every engine choice and every threshold is calibrated. An engine claim without such a measurement is treated the way this paper treats a violation verdict without a rule — as not yet a judgement.

8.1 The conformity check

Given a passage and the rule inventory, the conformity check returns, per rule, one of three verdicts: **conforms**, **conflicts**, or **not applicable**. A topic gate keeps the third verdict honest: a rule whose subject matter is absent from the passage produces no finding at all, rather than a vacuous “conforms”.

The instructive result is *which* engine may judge *which* rule. A single zero-shot inference engine — a natural-language inference model asked whether the passage entails a violation hypothesis — was measured first and scored 4 of 16 on the golden set: subtle grammatical distinctions such as “a fact cannot be true” are exactly what a similarity-trained judge misses, which is this paper’s opening claim returned as an experimental result. The engine was therefore demoted, not discarded, and each rule is now judged by the engine its own content demands:

- **Syntax rules** are judged on the dependency parse. The truth-bearer rule R-GG-001 fires when a parse configuration obtains — the lemma *fact* or *Sachverhalt* as subject under the predicate *true* or *false* — not when suspicious vocabulary co-occurs. The rules that nothing follows from a *sentence* (only from the proposition it expresses) and that a *goal* is not a truth-bearer are detected the same way.
- **Lexical rules** are judged by the words they themselves ban or demand: the verification ban, the marketing register, the hedging budget. For these rules the word list is not a heuristic for the rule; it *is* the rule’s subject matter.
- **Semantic rules** — in the present inventory, the metaphor ban — keep the inference engine, but contrastively: the passage is scored against a violation hypothesis *and* a conform hypothesis stored on the rule, and only a clear margin between the two yields a verdict. These findings are advisory by tier.

Under this division the golden set is met in full: 16 of 16 verdicts, 11 of 11 conflicts and 5 of 5 conforming counterparts, with no false positive on the not-applicable passages. Every finding names its rule, its engine,

and its evidence — a parse configuration, a matched expression, or an inference margin — so a contested verdict can be taken apart at the level where it was made.

■ 8.2 Logical relations across a collection

The second service turns the judgement outward, from a passage against rules to a sentence against a corpus. Given a test sentence A and a collection, it returns every sentence B for which B **implies** A, A implies B, the two are **equivalent**, or B **contradicts** A — with each finding addressed, as always, to (document, page, sentence).

Pairwise inference over a whole collection is not feasible directly, and the query philosophy of the PDAI platform forbids it anyway: the most restrictive fast filter runs first. The scan is therefore a funnel. A multilingual embedding of the collection’s sentences [19] supplies semantic recall: the top candidates by similarity to A, retrieved from the collection’s vector store in roughly a second. Only this candidate slice reaches the judge, a multilingual inference cross-encoder [20] that scores each pair in both directions. The direction matters twice over. Entailment is asymmetric, so B-implies-A and A-implies-B are distinguished and their conjunction reported as equivalence; contradiction, by contrast, is symmetric, and the scan exploits that as a guard — a contradiction is reported only when it holds in *both* directions, which was measured to suppress the spurious one-directional contradictions that negated phrasings of A otherwise provoke.

The golden test seeds the measurement. Ten sentences were drawn deterministically from the karman-nachlass collection — the technical reports and correspondence of Theodore von Kármán’s Nachlass, 663,972 parsed sentences — and one of them fixed as B₁: a hydraulics report sentence stating that Colorado River water must be lifted about 1,700 feet over a mountain range. Two test sentences were written down with their verdicts in advance: A₁, a paraphrase that B₁ implies, and A₂, its gravity-flow negation that B₁ contradicts; the other nine seeds are neutral to both. The cross-encoder reproduces all 20 designed verdicts. The full funnel then runs both test sentences against the entire collection: in both runs the seeded B₁ is retrieved with the designed relation — as an implication of A₁ among 21 non-neutral findings, as a contradiction of A₂ among 32 — with recall in 1–13 seconds and the verdict stage near 50 seconds for 200 candidate pairs scored in both directions. The scan also returns what no seed predicted but the design hoped for: the same report’s neighbouring formulations of the lift, found independently on other pages, each with its address.

METHOD One discipline, two directions

The conformity check reads a passage against the rule inventory; the relation scan reads a corpus against a sentence. Both refuse the shortcut this paper argues against: neither reports a verdict from similarity alone. Similarity only nominates candidates; the verdict is an inference with a direction, and it carries its evidence and its address.

Both services, together with the dictionary lookup, the mode compiler, and the rule-submission path, are served from one process that keeps the linguistic models warm and attaches the collection databases read-only — one language institution consulted over the local network, rather than a copy of its rules pasted into every prompt.

■ 9 Guarding what a model writes

The services so far judge a text that exists. A language model that revises or writes text for the workspace needs the same rules before its output reaches a reader. The case that made this necessary was a lectoring run on a German paper about the Famulus platform: 136 paragraphs were revised under one instruction, which asked for “brilliant, eloquent, vividly readable German”. The revised paragraphs praised what the author had written (“remarkable”, “masters three heterogeneous holdings”), added intensifiers and remarks about the text, and lost propositions of the source. None of this broke a rule of grammar in the school sense; every case broke a rule of use of the practice the paper belongs to — an author states his claims, and a lector does not grade them.

9.1 Task forms, context, and measurements

LWgrammar is therefore the only source of linguistic instruction for every model call that writes or rewrites text, whatever the model. A client asks for a *task form* (revise a paragraph, reformulate it, shorten it, answer from evidence, write notebook text) and receives it as the system instruction. It passes a context: language, author and style family, scope, the whole document, the fixed terms of the document, the citable bibliography keys, and, for text without a source, the evidence the text is drawn from. After the model has written, the result is measured against its source. Every measurement is blocking: a revision passes only if none of them is worse than in the source, and every measurement names the rule that it enforces.

Measurement	Rule
propositions kept, none added, none reversed; logical particles unchanged	R-LG-024
no evaluative predicate about the text, its author or its subject	R-LG-021
no remarks about the text or its author	R-LG-022
no rule identifiers or instruction echoes in running prose	R-LG-023
one concept, one word	R-LG-001
no conflict with any registered usage rule	R-GG-001 and following
no inflation or register change against the author's style family	R-LG-026
no invented or lost citations; numbers with provenance	R-LG-014, R-LG-006
no aesthetic objective in the instruction itself	R-LG-025

The last row checks the instruction, not the text: “brilliant” and “eloquent” are not admissible objectives, because they license the evaluation and inflation the other rows forbid. An owner's content directive — strike, add, complete, cite — changes content on purpose; the measurements it necessarily touches are reported but do not block. No directive licenses an evaluation, a meta remark or an intensifier. Text without a source, such as a synthesised answer, is checked for numbers and citations only where it goes beyond the evidence it was given; where no evidence was passed, those two measurements are reported as not applicable rather than as passed. The detectors of R-LG-021 to R-LG-026 are derived from the evaluation itself, and each change of a detector is written to a history table.

9.2 Measured on the damaged revisions

Against the 136 damaged paragraphs, each measured against the paragraph as it stood before the lectoring run, the guard rejects 133. The identity control (each of the 141 source paragraphs measured against itself) fails in no case. The three revisions that pass are the limit of the present measurements. A positive control is still missing. Measured against the author's later rewrite of the whole paper, only 8 of 77 aligned paragraph pairs pass; that rewrite changes and extends content by the author's instruction, which was not passed to the guard as a directive, so the figure measures the difference between a reformulation and a rewrite, not a false-positive rate. A clean positive control needs revisions that the author accepted as faithful reformulations, and those are what the decision record below collects.

The guard was also run on the other seven prose paragraphs of this section, as text without a source. A word in quotation marks counts as mentioned, not used, so the quoted “remarkable” and “brilliant” do not count as evaluations. Four of the seven paragraphs still fail: three because they speak of the author, one because

it names rule identifiers. Here the author and the rules are the subject of the text, not remarks about it, and the guard does not yet tell the two apart.

9.3 Decisions become rule evidence

The author decides on every revision the lector proposes. Each accept and each reject is recorded with the guard's verdict on that revision, and three cases carry information. A reject of a revision the guard had failed confirms the failed rules. An accept of a revision the guard had failed contests them: the rule receives counter-evidence and a candidate for its revision opens. A reject of a revision the guard had passed shows that something the author rejects is covered by no rule, and a candidate for a new rule opens. The author's decision outranks the machine's measurement, but no rule changes by itself: candidates stay open until the author turns one into a rule, revises or revokes the contested rule, or dismisses the candidate, which is kept.

9.4 Senses in time: the store behind Case A

The rules a guard applies are scoped, and so are the senses behind them. For Kepler the ai-dictionary holds two dated senses of the cause of planetary motion: `anima#kepler#motrix`, attested in the *Mysterium cosmographicum* of 1596, and `vis#kepler#motrix`, attested from 1609 to 1621. The change between them is stored with Kepler's own words: in a note of the second edition of the *Mysterium* (1621; kgw-08, Bl. 109, sentences 16–22) he writes that one has his principle of celestial physics if one substitutes *vis* for *anima*, and that he had once believed the moving cause to be a soul, until he saw that it weakens with distance as the light of the Sun does. The store records that replacement as a relation between two senses with a locus and a year. It does not turn it into a synonymy: Case A showed *anima* still in use in the *Epitome*. The rule R-KEPLER-ANIMA-VIS is the consequence: a 1596 sentence is not to be read with the later sense.

If the senses are dated, a specimen of text can be situated by them. The situating service measures the distance of a specimen to the dated works of an author by the frequencies of their most frequent word forms (Burrows' Delta), adds the hits of markers of dated senses such as *anima motrix* and *vis motrix*, and returns a time window, a discussion context and the nearest works; below 250 tokens it returns no verdict. On 18 specimens from the Kepler volumes, each situated with its own volume excluded, the mean error of the estimated year is 8.8 years at 500 and at 1,000 tokens, against 9.4 years for a baseline that answers the mean year of the remaining works; 10 of the 18 true years fall inside the returned window. The improvement over the baseline is small, and the discussion context is less reliable than the year: a passage from the *Astronomia nova* (1609), situated with that volume excluded, is dated 1605 with a window of 1596–1619, but assigned to the context "correspondence". Situating is therefore a measured service with declared error, not an attribution.

10 Relation to fine-tuning

Fine-tuning a language model on authored text encodes a practice of use, but it does so implicitly: the practice is captured in the gradient updates rather than written down as rules, and it cannot be queried or corrected without re-training. Two things this paper's approach does that fine-tuning alone cannot: it makes the rules inspectable (the store is read; a decision can be traced to a cited rule) and it makes them revisable (a submitted correction takes effect immediately, for every consumer of the store, without a training run).

The two approaches are complementary. A fine-tuned model learns the statistical regularities of the register; the institution supplies the declared exceptions — the norm that must be followed precisely because it is a norm, not merely the most probable next token. A model that has been trained on Kepler will produce Keplerian prose; it will not, without the dictionary, know that *lex naturae* in an *Astronomia Nova* passage follows the ablative-rule sense and must not be glossed by the juridical-natural-law sense of a century later. For that it needs the locus with its coordinates, the competing sense with its contrast, and the rule that makes the distinction a distinction.

The store is therefore consultable ground truth for any language model, fine-tuned or not. A model that consults the pack before drafting is not merely imitating the register: it is following the rules of the practice, which is what it means to write in the language of that practice in the first place.

11 Conclusion

Semantic understanding of scientific texts is not a better embedding. It is a normative achievement: a formulation either follows a rule of a practice or it contradicts one, and the contradiction is only stateable with the rule in hand. Wittgenstein's later philosophy supplies the grammar of that achievement — use, perspicuity, practice, hinges, a riverbed that can shift for reasons. The OED supplies the form in which a practice of use can be *seen*. LWgrammar writes that form from the pages of a corpus, after OCR and linguistic analysis, and keeps every quotation on a page. The formulation score makes the thesis executable: no violation without a cited rule. The ai-dictionary makes the scope executable: one headword, many rules, each with an author, a period, and an address. And the two measured services make the judgement scale without surrendering it — a conformity check whose engines are assigned by what each rule is about, and a relation scan in which similarity only nominates and inference decides.

What this changes for historical scientific Latin is microscopic and therefore decisive. *vis* is not one thing. *anima* does not become *vis* by translation. *lex naturae* is not Kepler's phrase because *lex* and *natura* co-occur. Length is not a number. Each of those sentences is a rule, with an anchor, and a machine that cannot cite the rule has not understood the sentence.

The companions take neighbouring tasks: AI Erasmus, the dictionary as an operation the scholar performs; the AI Partitur, the score that makes an agent's steps readable; the golden-query testbed, the thirteen questions against which those operations are held [5]; computational epistemology, the validity of reconstructed arguments. Further authors can be added to the reference library the way Wittgenstein was added; further national dictionaries can be added as authorities without storing their text; further publication profiles can be declared the way *arxivai* is declared. The inventory grows. The demand does not: a judgement of meaning names its rule.

Bibliography

- [1] G. Graßhoff, "AI Scholar: From PDF to Portable Document AI Infrastructure," 2026.
- [2] G. Graßhoff, "Querying with AI Linguistic Trees: Stanza-Backed Dependency Parses, Lemma Normalisation, and DuckDB," 2026.
- [3] G. Graßhoff, "AI Erasmus — A Platform for a New Era in Digital Humanities," 2026.
- [4] G. Graßhoff, "The AI Partitur — A Score Language for the Humanities," 2026.
- [5] G. Graßhoff, "Golden Queries in AI Erasmus: A Testbed for Scholarly Inquiry," 2026.
- [6] G. Graßhoff, "Computational Epistemology: Foundations for the Assessment of Propositional Validity in AI Systems," 2026.
- [7] L. Wittgenstein, *Philosophische Untersuchungen*, 11th ed. Frankfurt am Main: Suhrkamp, 2022.
- [8] Bureau International des Poids et Mesures, *The International System of Units (SI)*, 9th ed. Sèvres: BIPM, 2019. [Online]. Available: <https://www.bipm.org/en/publications/si-brochure>
- [9] Conférence Générale des Poids et Mesures, "Resolution 6 of the 11th CGPM (definition of the metre by the krypton-86 wavelength)." 1960.
- [10] Conférence Générale des Poids et Mesures, "Resolution 1 of the 17th CGPM (definition of the metre by the speed of light)." 1983.
- [11] OPERA Collaboration, "Measurement of the neutrino velocity with the OPERA detector in the CNGS beam." 2011.
- [12] T. Adam and others, "Measurement of the neutrino velocity with the OPERA detector in the CNGS beam," *Journal of High Energy Physics*, vol. 2012, no. 10, p. 93, 2012, doi: 10.1007/JHEP10(2012)093.
- [13] L. Wittgenstein, *Über Gewißheit*, 16th ed. Frankfurt am Main: Suhrkamp, 2023.

-
- [14] I. Newton, *Philosophiæ Naturalis Principia Mathematica*. London: Jussu Societatis Regiæ ac Typis Josephi Streater, 1687.
- [15] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. D. Manning, “Stanza: A Python Natural Language Processing Toolkit for Many Human Languages,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Association for Computational Linguistics, 2020, pp. 101–108. doi: 10.18653/v1/2020.acl-demos.14.
- [16] M. Raasveldt and H. Mühleisen, “DuckDB: An Embeddable Analytical Database,” in *Proceedings of the 2019 International Conference on Management of Data (SIGMOD)*, ACM, 2019, pp. 1981–1984. doi: 10.1145/3299869.3320212.
- [17] J. Kepler, *Gesammelte Werke*. München: C. H. Beck, 1938.
- [18] G. Eneström, *Verzeichnis der Schriften Leonhard Eulers*. in Jahresbericht der Deutschen Mathematiker-Vereinigung, Ergänzungsband~IV. Leipzig: Teubner, 1910.
- [19] J. Chen, S. Xiao, P. Zhang, K. Luo, D. Lian, and Z. Liu, “BGE M3-Embedding: Multi-Lingual, Multi-Functionality, Multi-Granularity Text Embeddings Through Self-Knowledge Distillation.” 2024.
- [20] M. Laurer, W. van Atteveltdt, A. Casas, and K. Welbers, “Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI,” *Political Analysis*, vol. 32, no. 1, pp. 84–100, 2024, doi: 10.1017/pan.2023.20.