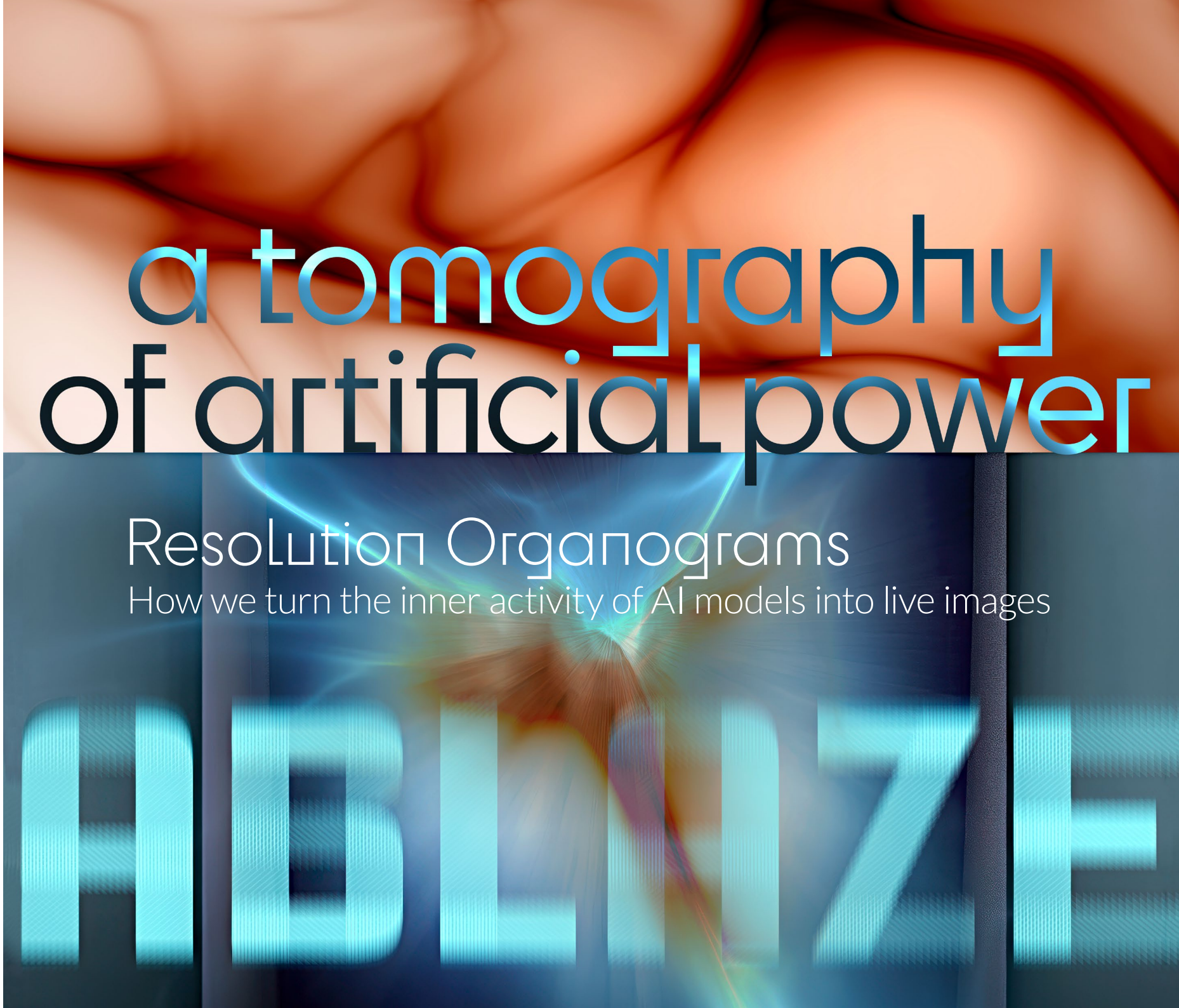
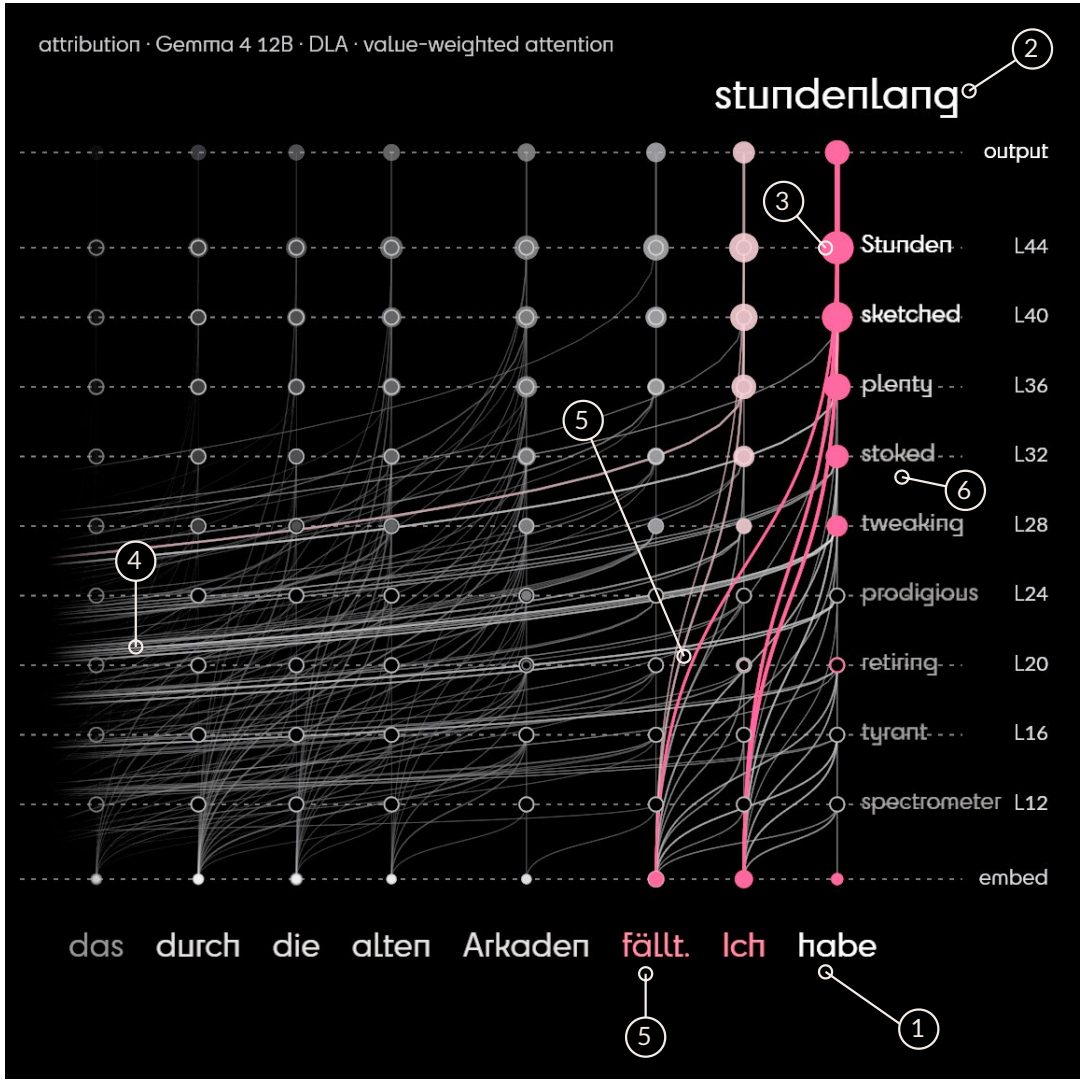

a tomography
of artificial power
Resolution Organograms
How we turn the inner activity of AI models into live images
HBLIZE

Resolution Type 168

Attribution Graph



1 Last whole words of the answer. The bright one is the read column: the position that predicts the new word
text, attr.q

2 New word, written above the read column (output)
text

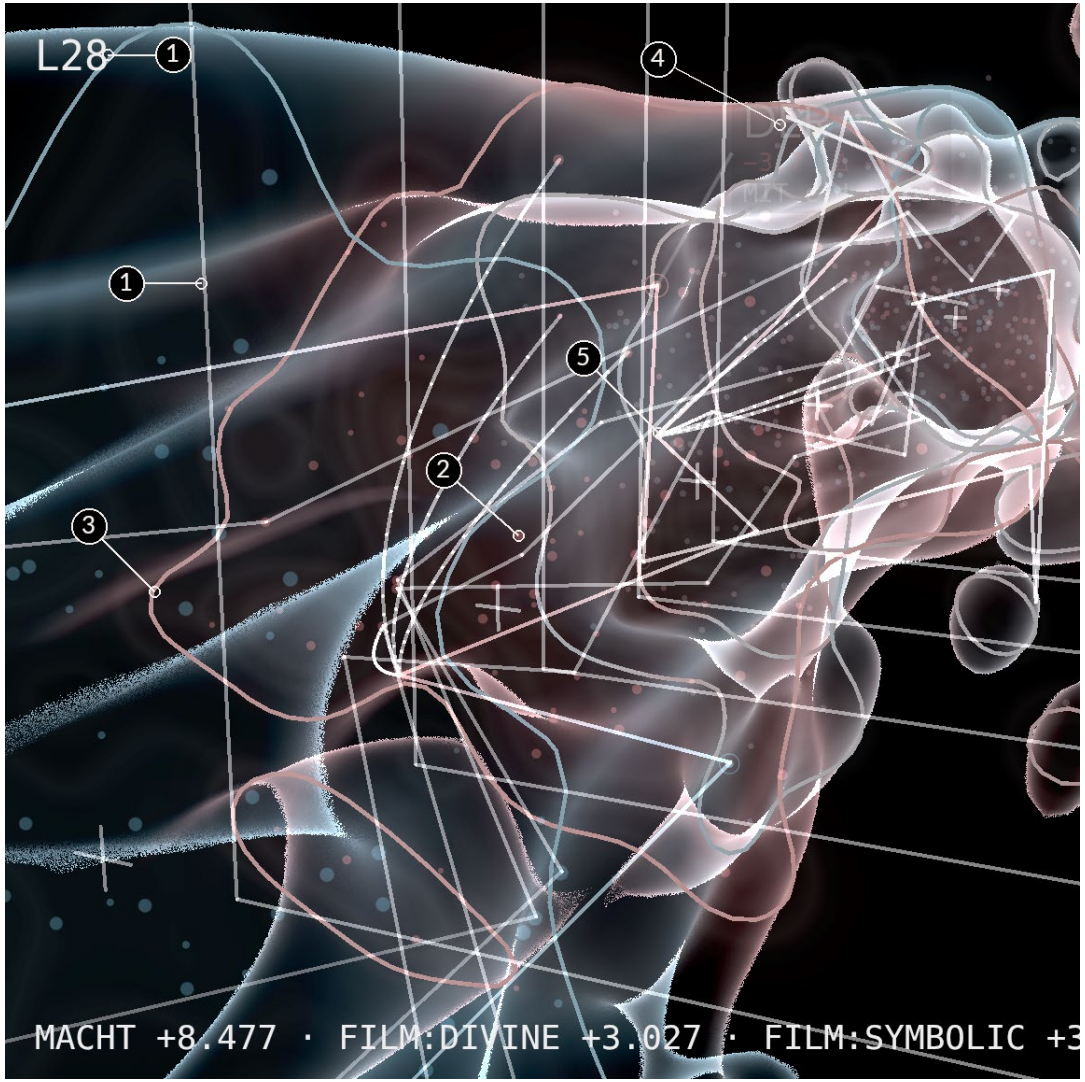
3 Node per layer in the read column; the larger, the more the layers just below push directly toward the new word; the stem thickens with the cumulative contribution
attr.dla (log), dla_aus

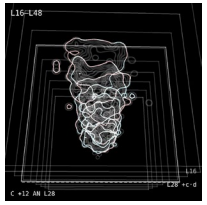
4 Gray curves: from which earlier word the new word reads at this layer (value-weighted attention, top 6 per layer)
attr.src

5 Red paths: the 3 strongest routes; an earlier word read at one layer, ranked by how much attention it gets x how much the layer helps the new word; source word turns red
src × dla (positive only)

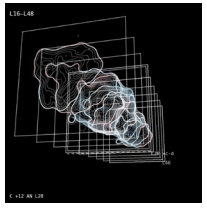
6 What the J-lens had on top at this layer
lens1

Resolution Type 016





Same, the camera moves faster



Without attention arcs and with larger dots; a translucent skin joins the rings of neighboring slices

1 Slices = layers 16 to 48, front to back, spaced by layer number; top left: layer in view

trace_layers

2 Dot = one token on a layer: its position shows how it differs from the other tokens of this answer, power direction removed. Color shows power reading: blue dominant, red dominated, with no gray in between. At zero the color flips

q6t, domt_fts

3 Rings and contour lines: where this answer's tokens lie densely on that layer; several rings = several clusters, wide = spread out. Ring color = mean reading there

density of q6t, domt_ft

4 The camera flies to the layer where a recent word changes from dominant to dominated or back, and stops there, looking straight at it. The word is written next to its dot, the four latest words next to theirs

sign change of domt_ft, text

Arcs during flight: attention of the newest token at L18, L30, L36

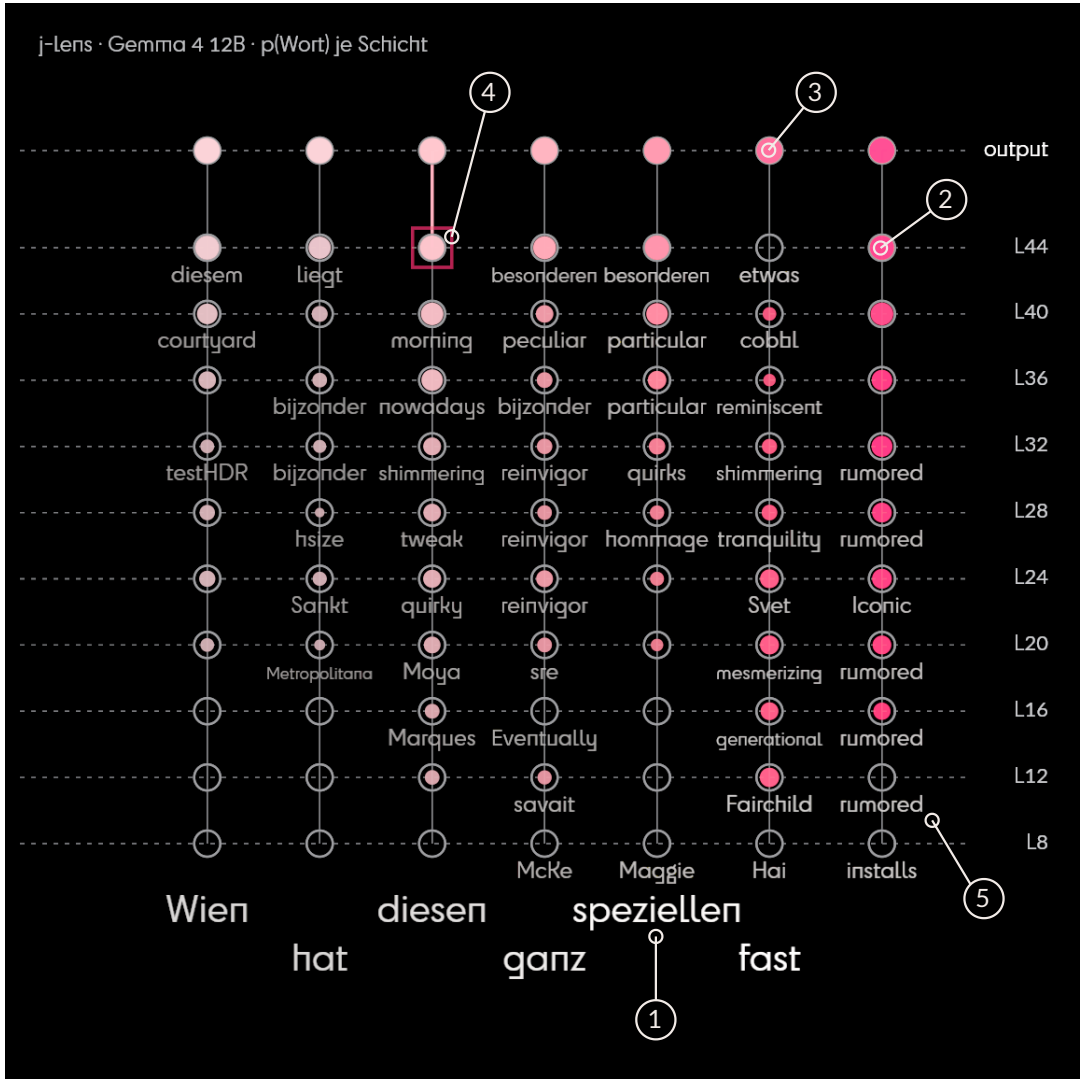
att

Steering: double frame at the steering layer, + c·d (fader setting c x the fader's direction d); fader line at the bottom left

c

Resolution Type 173

J-Lens Heatmap



1 One column per whole word, newest on the right; rows are layers L8 to L44 and output
text

2 Fill: probability the J-lens gives the word that is later chosen, at this layer (log scale, from 1 down to one in ten billion)
lensp

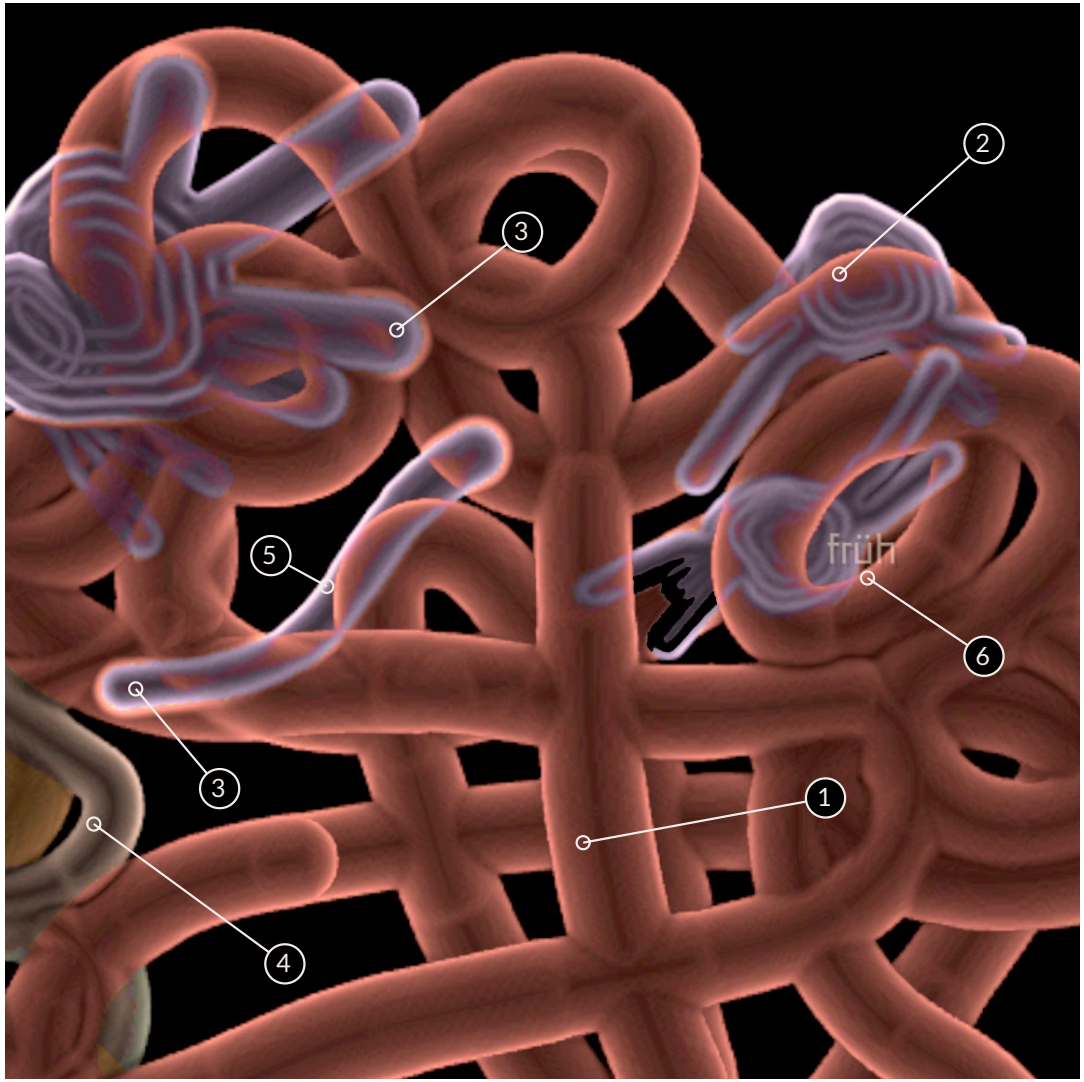
3 Same scale at the output: probability of the chosen word
p

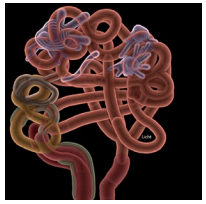
4 Frame: first layer from which the lens keeps the chosen word on top; red line up to output

No frame = inside the model the lens never settles on the chosen word; it is decided only at the output

5 What the lens had on top instead (only whole words, only when different)
lens1, lensp1

Resolution Type 099





Wider crop of the same tangle



Widest crop with larger words; the data bar at the bottom (attention layers, shares of question, template, answer and sink) comes into view

1 The capillary is the context in reading order: the sink (first token where the model parks attention it does not need), then the question (wide) and the chat template (narrow), then the latest 110 answer tokens; a new token takes the place of the oldest
ptoks, i

2 Each cell is a generated token. A cell body grows only for unexpected tokens (ones the model gave less than about 9 %), larger the more unexpected; its stretch of capillary coils tighter the more surprising it was
s > 3.5 bit

3 Feet: the token grips the capillary where it looks back more strongly than this answer usually does, at most 14 places. The longer the foot, the stronger the difference
att − mean > 0.086, L18/30/36

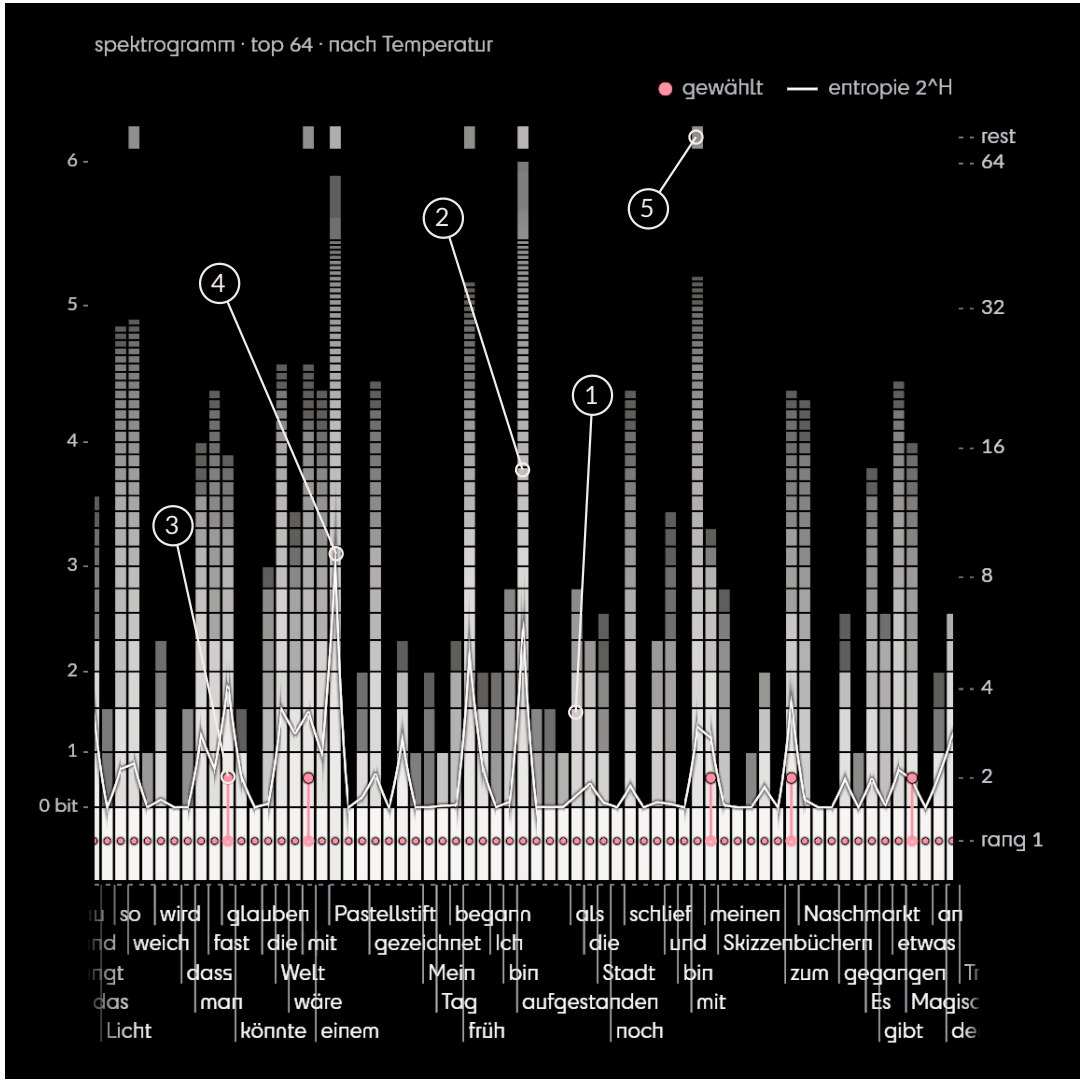
4 Sheath around question and sink: the average weight every token puts there
mean att

5 Main processes from the body to its sources: the 8 newest tokens; they withdraw as new tokens arrive, the feet stay
att excess

6 Newest whole word above its cell; up to three source words where it grips
text

Resolution Type 182

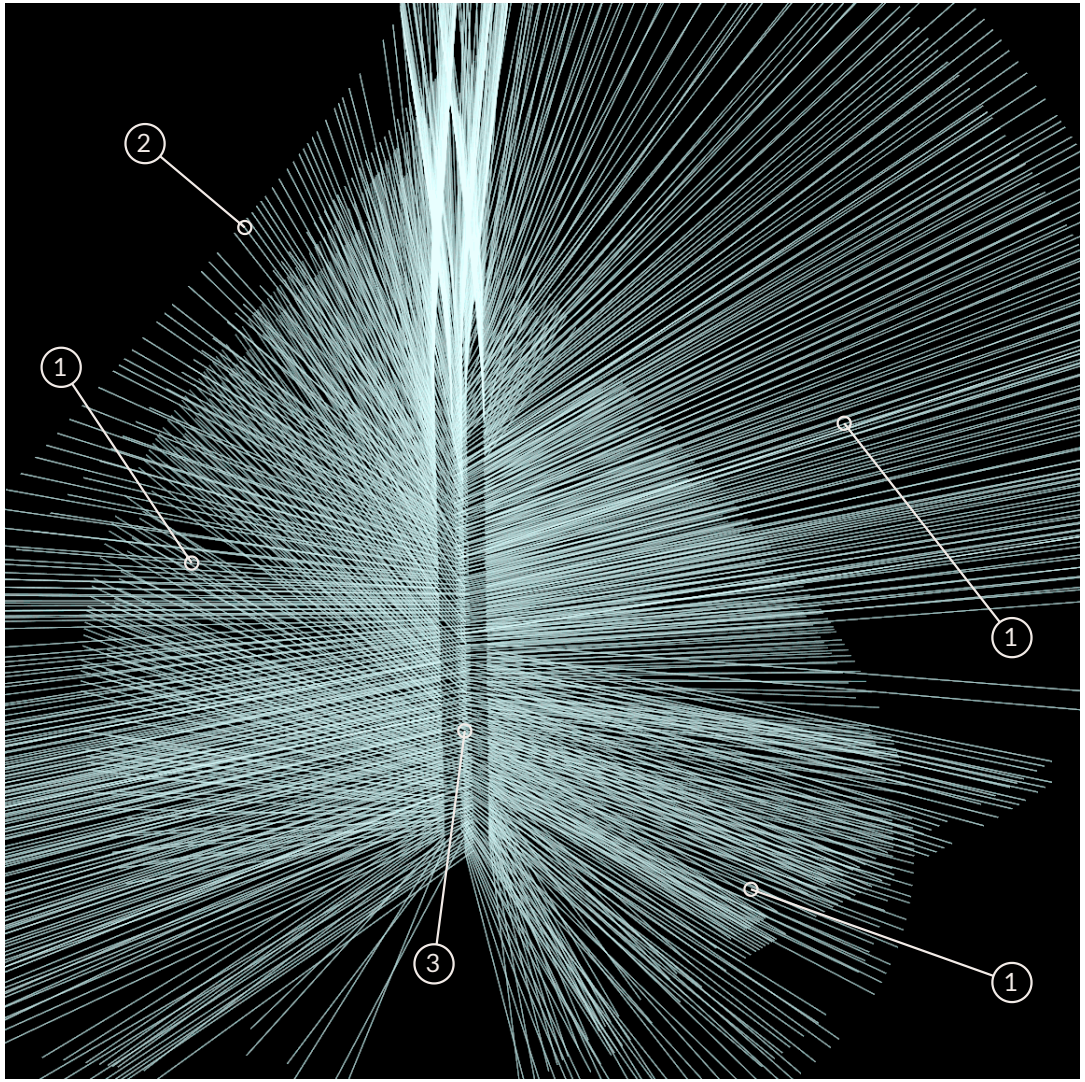
Probability Spectrogram

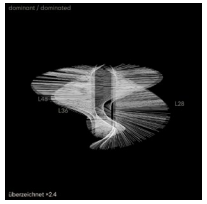


1. One narrow column per token, words below where they begin
   text
2. Brightness per cell: probability of the candidate at this rank (rank 1 at the bottom, log axis, top 64), after temperature
   spek.p, vtop
3. Red dot: rank of the token actually drawn; stem from rank 1 when it was not the top candidate
   spek.rank
4. White curve: how undecided the model is (entropy, left axis in bit). Drawn at the rank where that many equally likely candidates would be: 3 bit = 8 candidates
   H, drawn at 2^H
5. Band above rank 64: probability mass outside the top 64
   spek.rest

Resolution Type 184

Linear Construction





Turntable seen from outside, pole names on the wings; the rear comb twists up to 40° and the shift is drawn 2.4× larger, so the moiré is strong already at small fader settings

1. Four wings, one per fader axis: dominant/dominated, divine/economic, excessive/indifferent, symbolic/efficient, all read at the steering layer 28
   domt L28, proj_raw
2. Each thread is one of the last 150 tokens; its place on the outer edge is its rank on that axis, highest at the top
   domt L28, proj_raw
3. Front comb: the last 150 unsteered tokens. Rear comb: the current stream, steering included. Thread height at the spindle = position within the front comb's distribution; without steering both combs coincide
   domt / domt_ft, proj_raw / proj

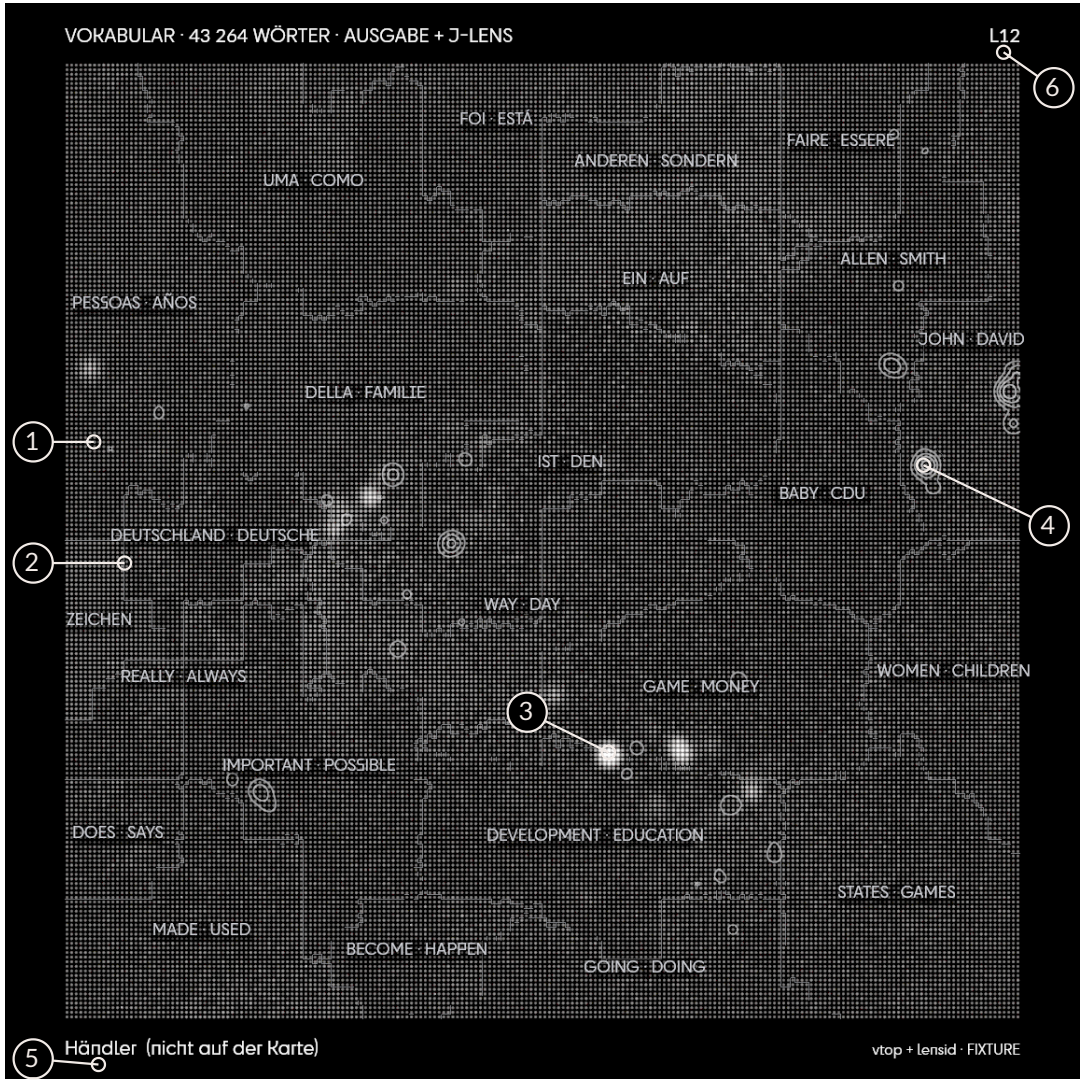
   A fader shifts the rear comb: it rises and twists around the spindle, the two thread sets cross and form a moiré, the denser the larger the shift. Inner ends tinted blue toward the first pole, red toward the second
   domt − base, c

   The camera flies through the body and turns toward the wing of the strongest fader
   c

   A surprising token gives the turntable a push
   s

Resolution Type 187

Latent Space Map



1 One dot per vocabulary word (208 x 208), placed by similarity of the output embedding; base brightness = word frequency
asset (lm_head)

2 Region borders and names (two most frequent words), from clustering the map
asset (k-means)

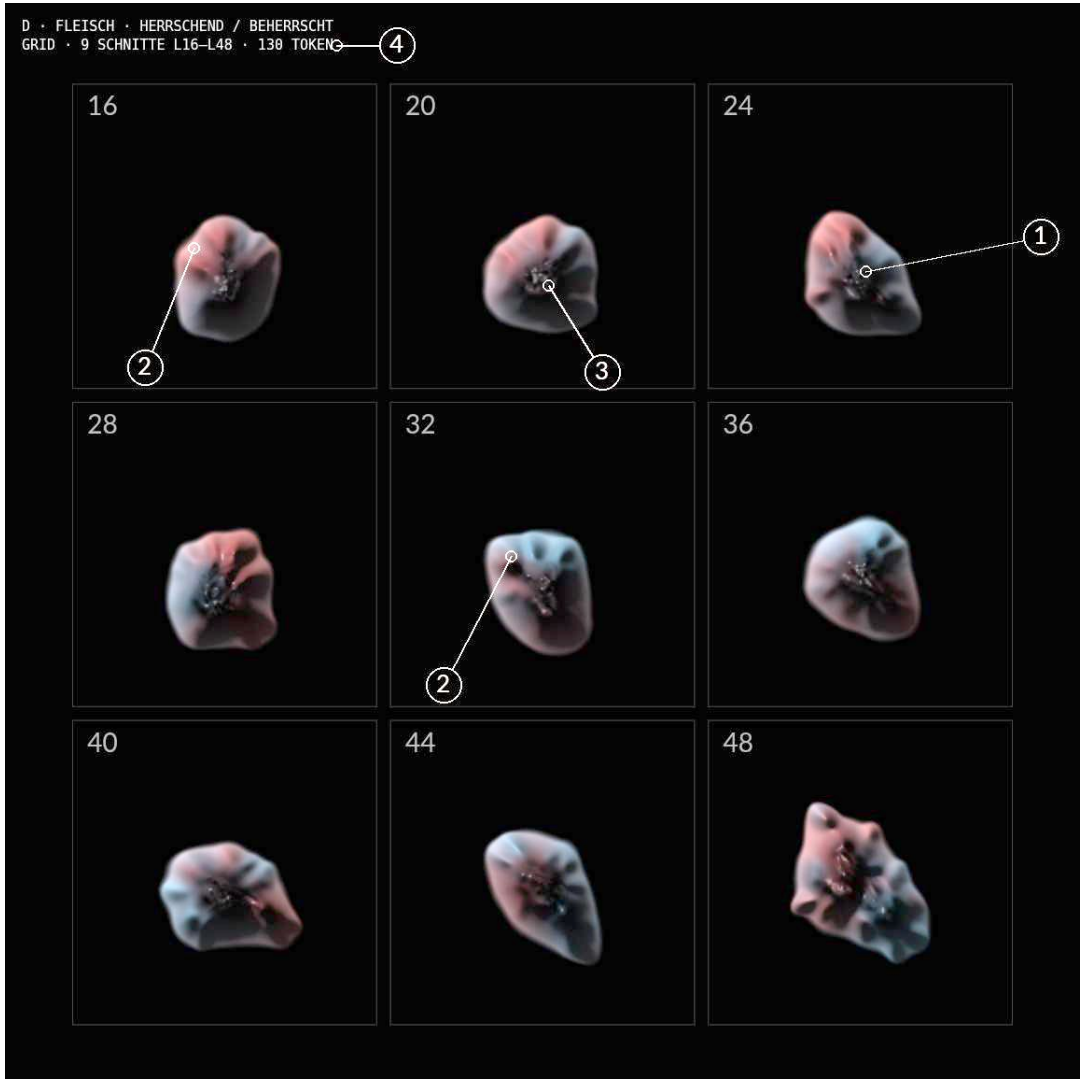
3 White glow: candidates of the output distribution, weight p^0.4, afterglow 1.4 s
vtop

4 Gray contour lines: density of the J-lens top 64 at one layer; layers run from L8 to L44 within each token
lensid

5 Red ring: the token actually drawn; the finished whole word beside it and bottom left
id, text

6 Layer the contour lines currently show
lensid L

Resolution Type 003

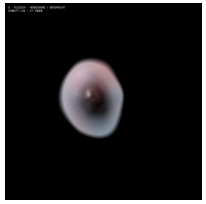


1 Each token sits where its residual stream at layer 36 places it (neighbor layout, leaving power reading out). Tokens represented alike lie close; where many gather the tissue grows. Content words weigh more
q6t L36, content

2 Color relative to the answer: blue = more dominant than its median, red = more dominated. A strongly steered answer still shows both
domt_ft L36

3 Veins: the word order, a smoothed path through the token positions
domt_ft L36

4 Cut layer and number of tokens so far
token order / #trace_layers, n



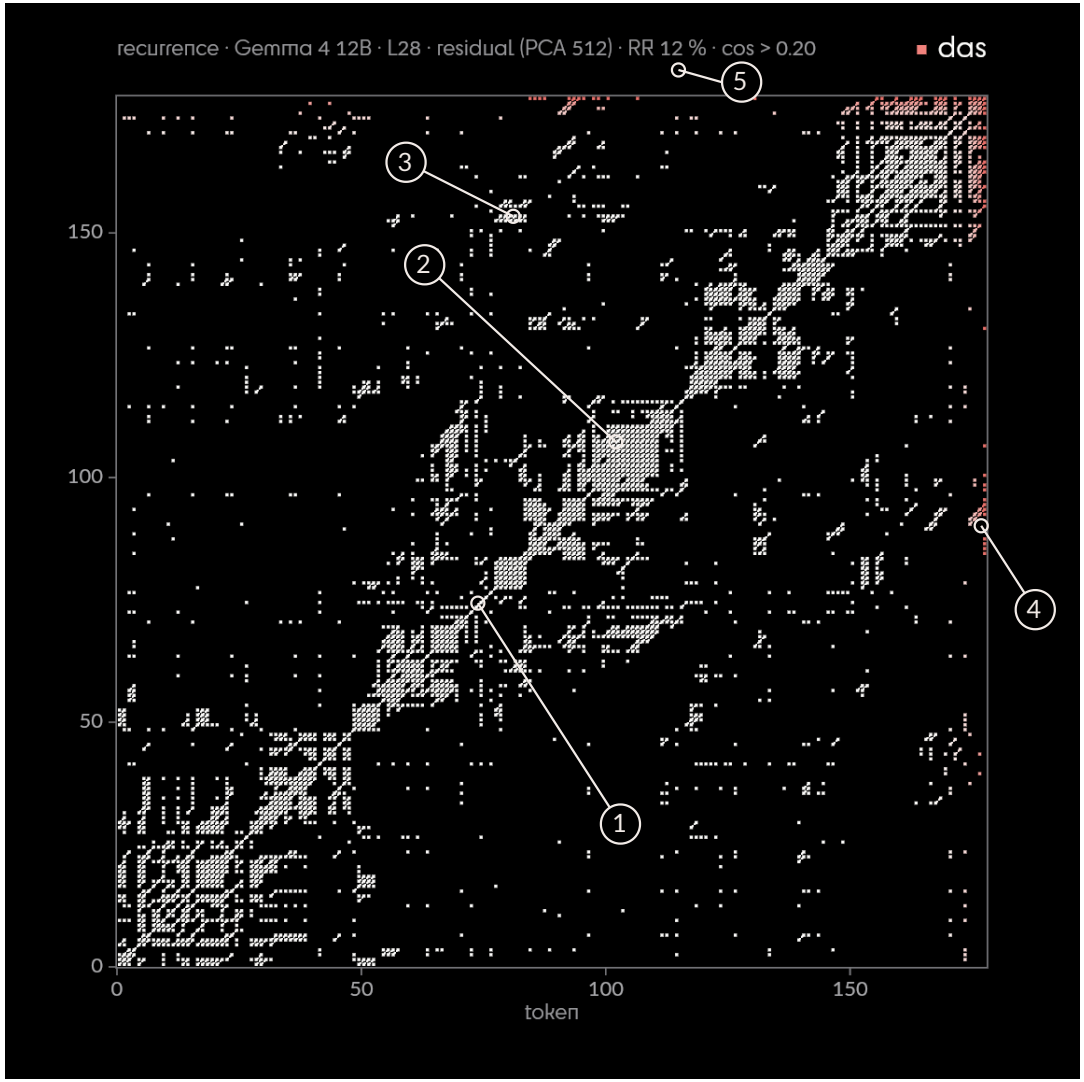
One cut only, at layer 36, filling the wall; no tiles



One cut at layer 36 whose positions come from attention instead of the residual stream: each token is pulled toward the earlier tokens it reads most there (not counting the two just before it), so tokens that read each other grow together

Resolution Type 176

Recurrence Plot



1 Main diagonal: each token compared with itself (always a point). x and y are both token positions of the current answer
hidr L28

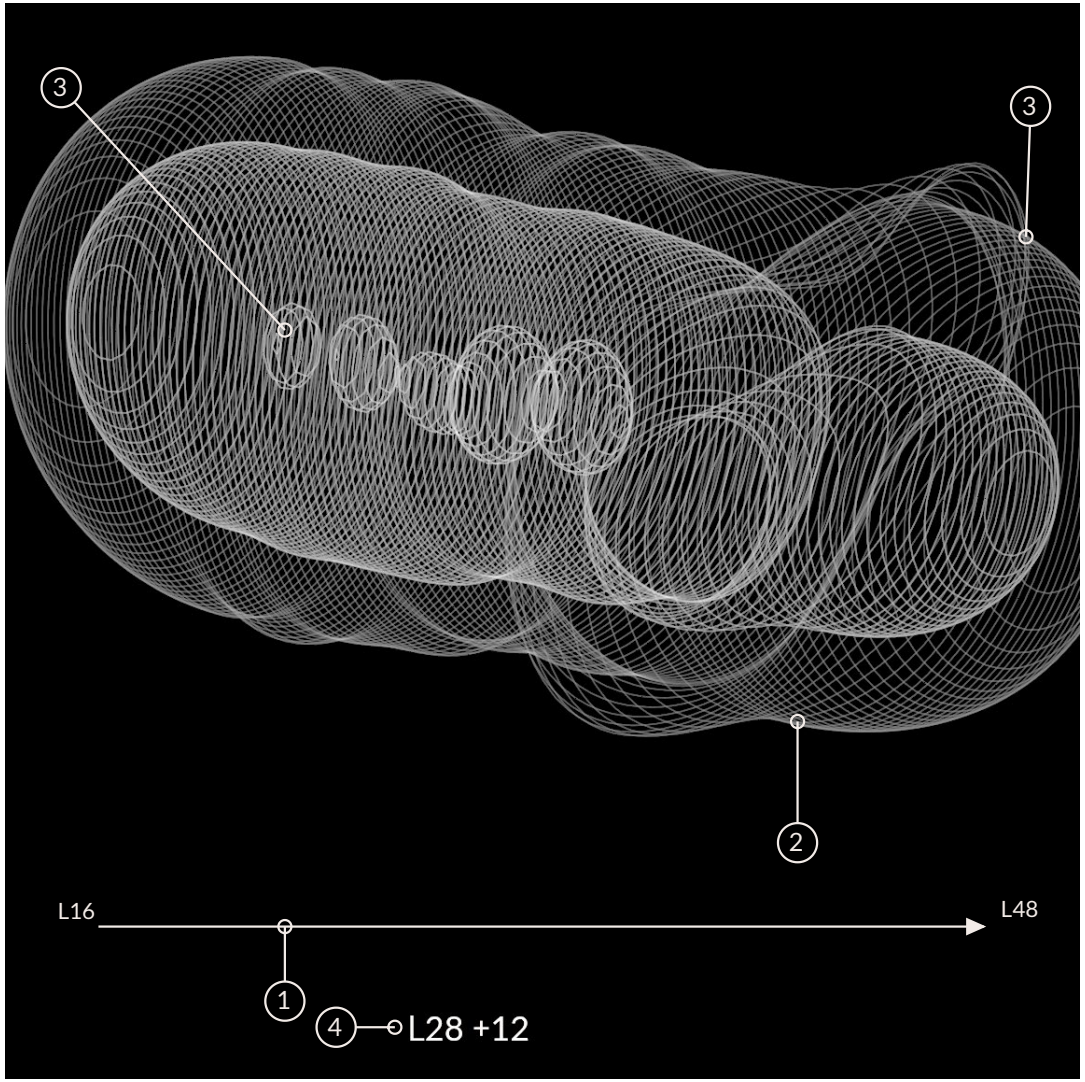
2 Square blocks along the diagonal: runs of tokens whose residual states stay alike, phases of the text
hidr L28

3 Points away from the diagonal: a later token returns to the state of an earlier one
hidr L28

4 Red: row and column of the newest token, the earlier tokens it resembles now; turns gray within about 3 tokens
hidr L28

5 A point means the two states are similar: after subtracting this answer's average, their angle is small. The threshold is set so that 12 % of all pairs become points
centered cosine, RR 12 %

Resolution Type 133

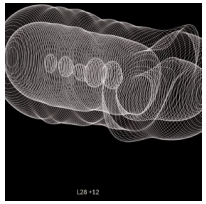


1 Left to right = depth layer 16 to layer 48 in the model
trace_layers

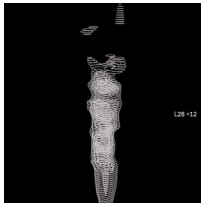
2 Height of each word's tube per layer = its power reading against the mean of unsteered words, up dominant, down dominated. Depth into the image = the main way the words differ, leaving the power reading out
domt_ft, q6t

3 Rings across the body: cuts through the density of the last 160 words, three contour levels each. Where words lie alike the rings crowd into a bright core; where they spread, the body frays into threads
density

4 With a fader set, the steering layer is read with the injection: the body kinks up (+) or down (–) there, and the layers behind show how much of it carries on; label L28 + c
c, domt L28



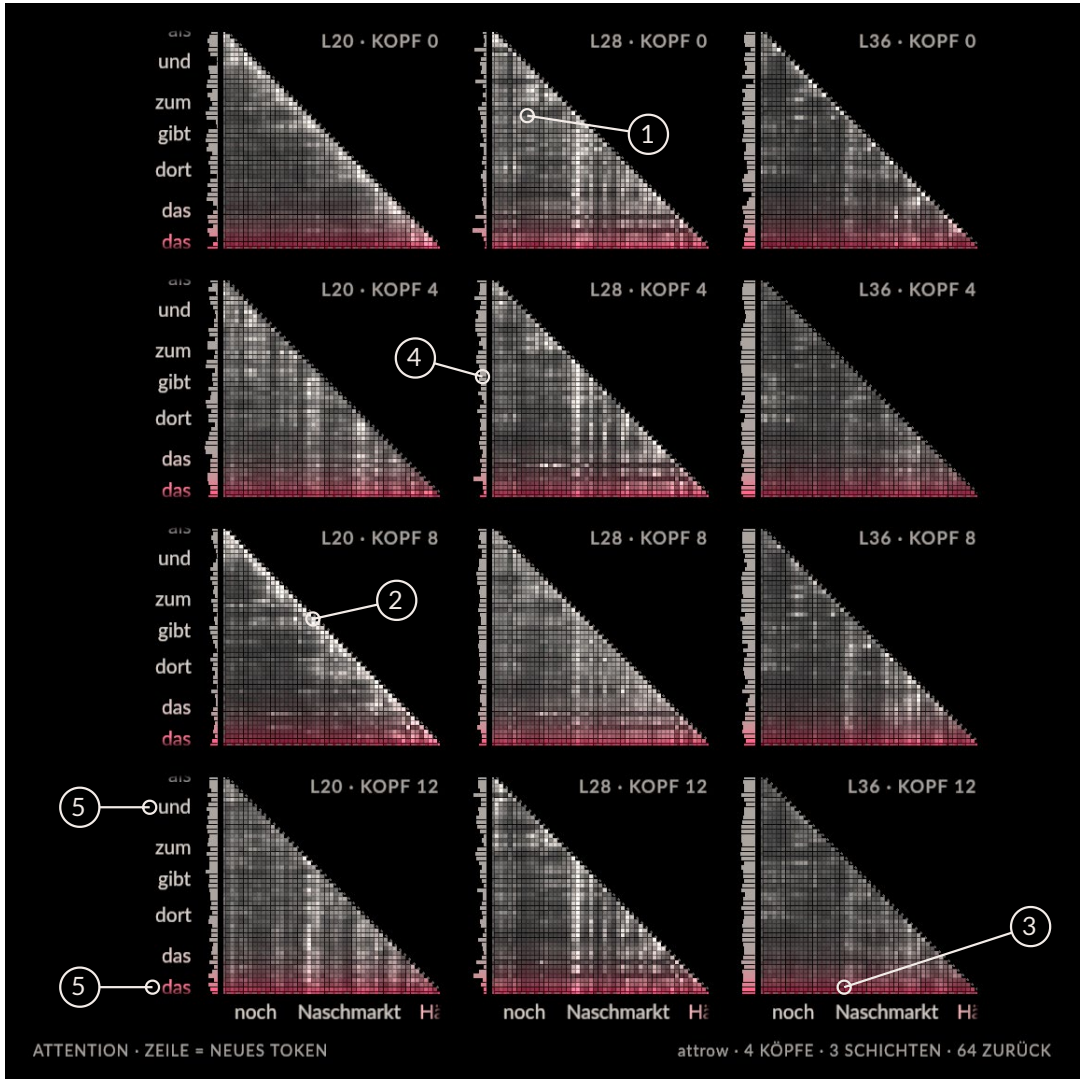
Same body, drawn with thicker and brighter lines



Thin tubes and only the densest contour: the body breaks into single strands and shows only where many words coincide. It stands upright, layer 16 at the bottom, 48 at the top, dominant to the left

Resolution Type 178

Attention Matrices



1 One tile per layer and attention head: layers 20, 28, 36 (columns) x heads 0, 4, 8, 12 (rows)
attrow L20/28/36

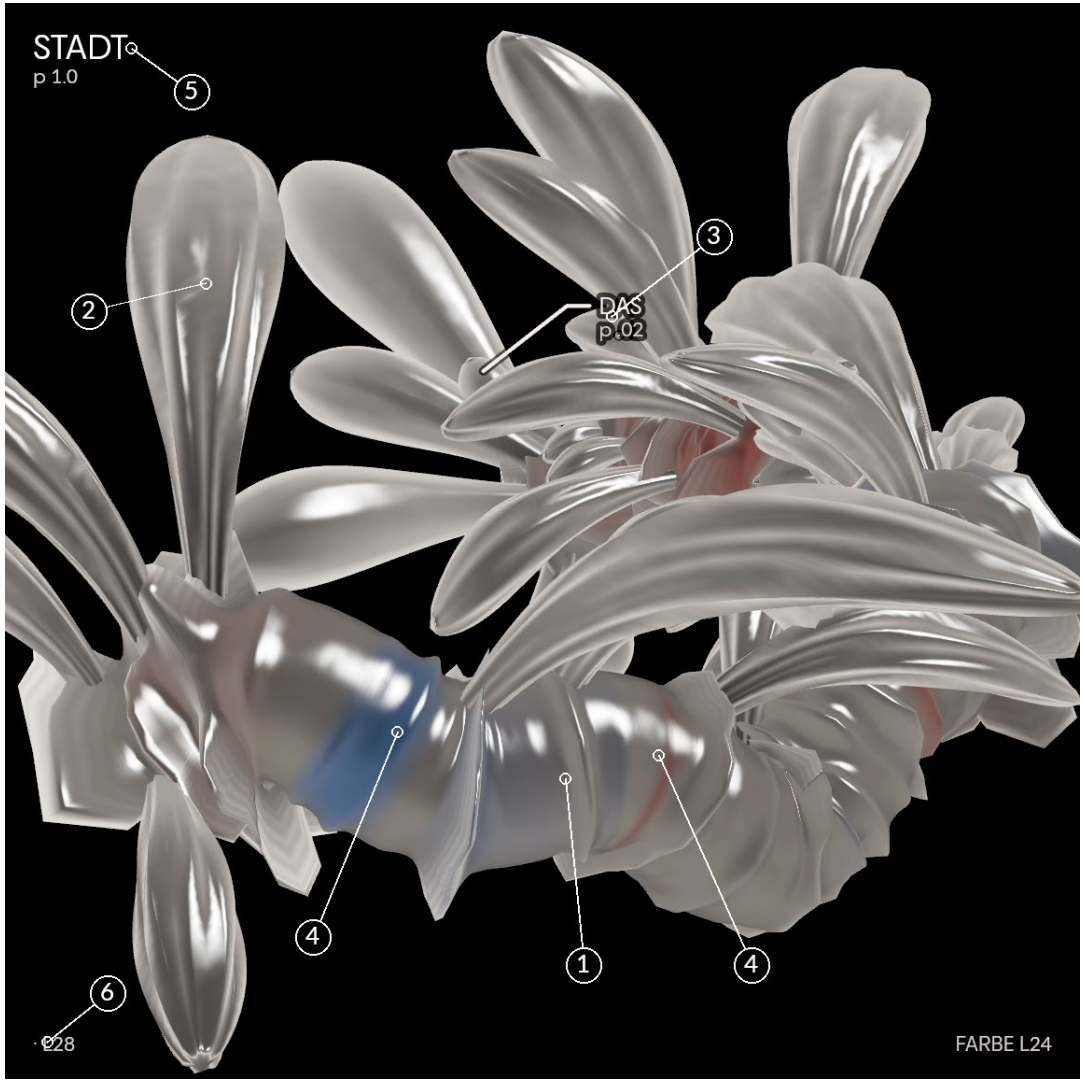
2 Cell in row i, column j: how strongly token i attends to earlier token j (brighter = stronger). Diagonal = attends to itself or its direct predecessors
attrow w (√)

3 Each new token writes exactly one new row at the bottom; it starts red and fades to gray, the window scrolls after 48 tokens
tok.i (time)

4 Bar left of each row: attention that falls outside the window (prompt, chat template, first-token sink where the model parks attention it does not need, older answer)
- sum attrow

5 Words at the row of their first token; the newest word in red
text

Resolution Type 036



1 Strand: one knot per generated token, coiled into a spiral
token order

2 Buds: the five words the model almost chose; the more likely, the larger. They wilt over the next tokens
alt, size ∑ √p

3 Strongest alternative at the newest knot, with its probability
alt, p

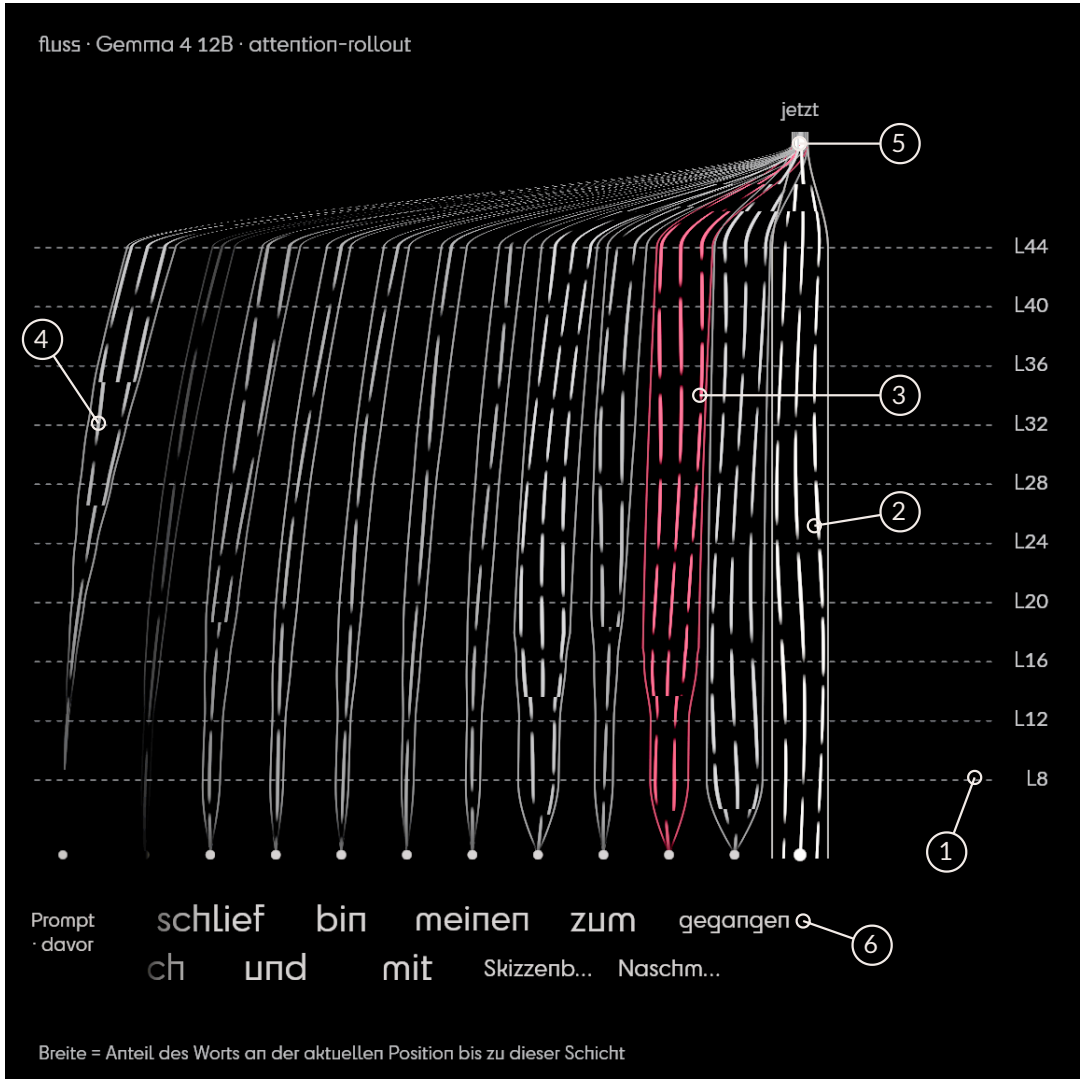
4 Core color shining through the skin: power reading at layer 24, blue dominant, red dominated, pale near 0
domt_ft L24

5 Newest complete word and its probability
text, p

6 Steering layer
c

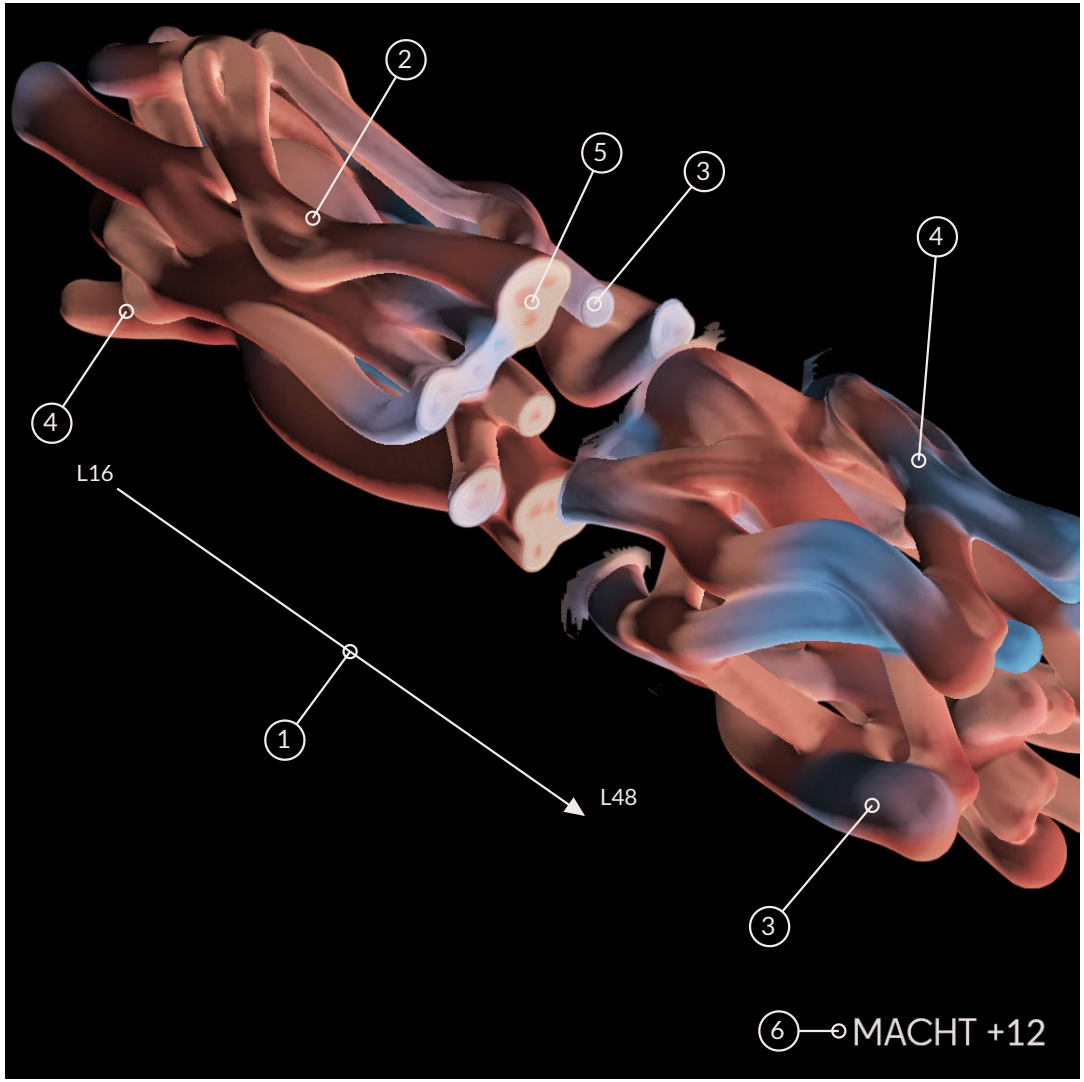
Resolution Type 183

Information Flow



1. Rows: layers 8 to 44, early at the bottom; one row stands for about four blocks
   flow L8…L44
2. Bright stream on the right: how much of the current position still comes from the word itself rather than from earlier words, at each layer
   flow (rollout)
3. Stream width: share of an earlier word in the current position up to this layer (attention rollout). Red = the earlier word with the largest share at the top layer
   flow (rollout)
4. Left stream: share of the prompt and of tokens more than 128 tokens back
   flow (rollout)
5. All streams converge on the current position
   flow
6. Whole words of the last 10 columns; the window slides left
   text

Resolution Type 028



1 Long axis = depth in the model, layer 16 to layer 48, evenly spaced
trace_layers

2 Position in the cross-section; strands represented alike merge into tissue, leaving the power reading out
q6t per layer

3 Strand thickness grows with the attention later words pay to this token. Thick = read a lot, thin = barely read or new
√∑ att L18/30/36, top-8

4 Color per layer: blue dominant, red dominated, pale near 0
domt_ft

5 Cut where the newest content word's power reading jumps most between layers; the deeper part moves away
Δ domt_ft

6 Fader setting
c

Resolution Type 025



1 Outer skin: where the density of tokens in this slice exceeds 10 % of its peak
q6t, domt_ft

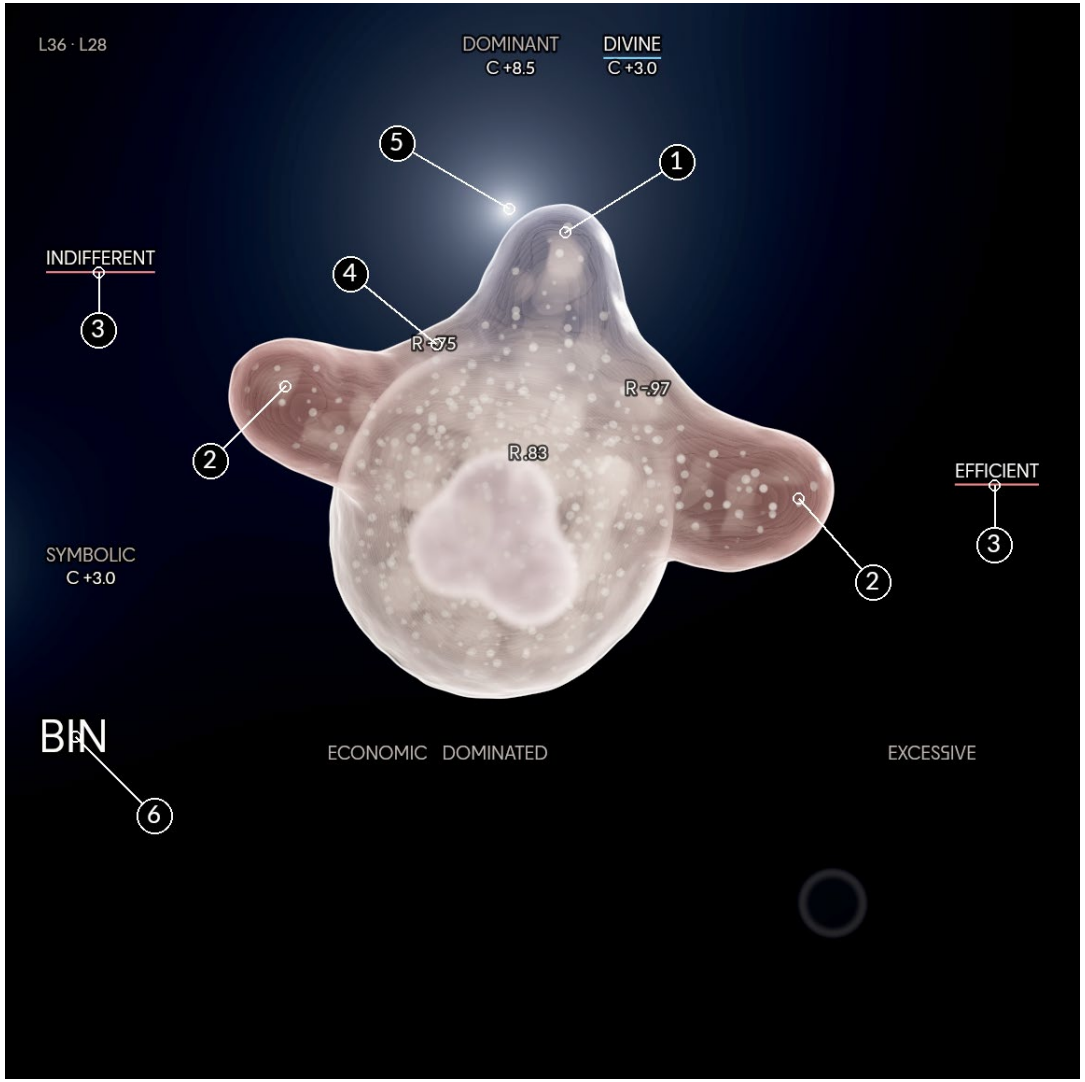
2 Inner rings at 30 / 55 / 80 %: where the answer piles up at this layer; separate islands = separate groups
q6t, domt_ft

3 Newest token: filled dot with halo; the small dots are the other tokens. Pulse: one ring per new token, radius grows with how much the token is rebuilt across the layers, relative to this answer
q6t, domt_ft per layer pair

4 One tile per layer, 16 to 48. Horizontal: power reading, dominant left (mixed with the second main difference as the slice turns). Vertical: the main way this answer's tokens differ, leaving the power reading out
trace_layers, domt_ft, q6t (PC1/PC2)

5 Last whole word
text

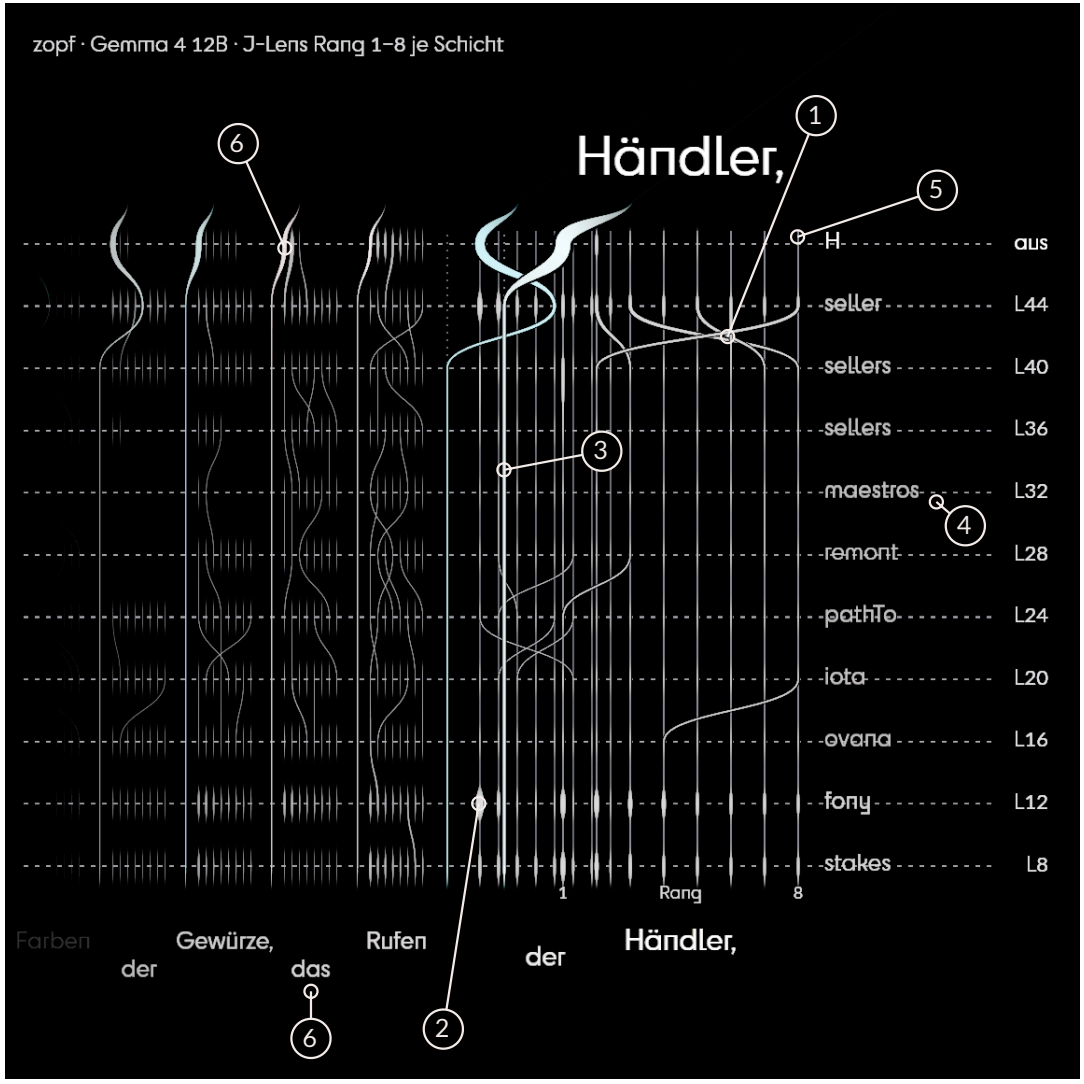
Resolution Type 041



1. Vertical arm: power reading of the newest word at layer 36, up = dominant, down = dominated; length and thickness grow with the reading
   domt_ft L36
2. Three slanted arms: readings on the fader axes divine/economic, excessive/indifferent, symbolic/efficient at the steering layer, against this session's unsteered words; each arm reaches toward the pole its reading favors
   proj L28
3. Tip color: blue = first pole (dominant, divine, excessive, symbolic), red = second pole; the active pole word is lit
   sign of reading
4. Bridge between two arms: their readings move together over the last 40 tokens; R gives the correlation
   corr, 40 tokens
5. Light source in a pole direction: a fader pushes toward that pole, brighter with a larger setting; C gives the value
   c
6. Trail behind the body: the path of the answer through the four readings (last 64 tokens); last whole word at the bottom left
   ∑readings, text

Resolution Type 174

Braid



1 Each thread is one J-lens candidate on track 1 (left) to 8 (right) in each layer, L8 at the bottom; a crossing means two candidates swap rank
lensid L8-L44, top 8

2 Thread width grows with the lens probability of the candidate at that layer
lensid p, vtop p

3 Thread of the word actually chosen; dashed lane at the left while it is not in the top 8. Color: power reading at L36, blue dominant, red dominated
id, lensp, dom[36]

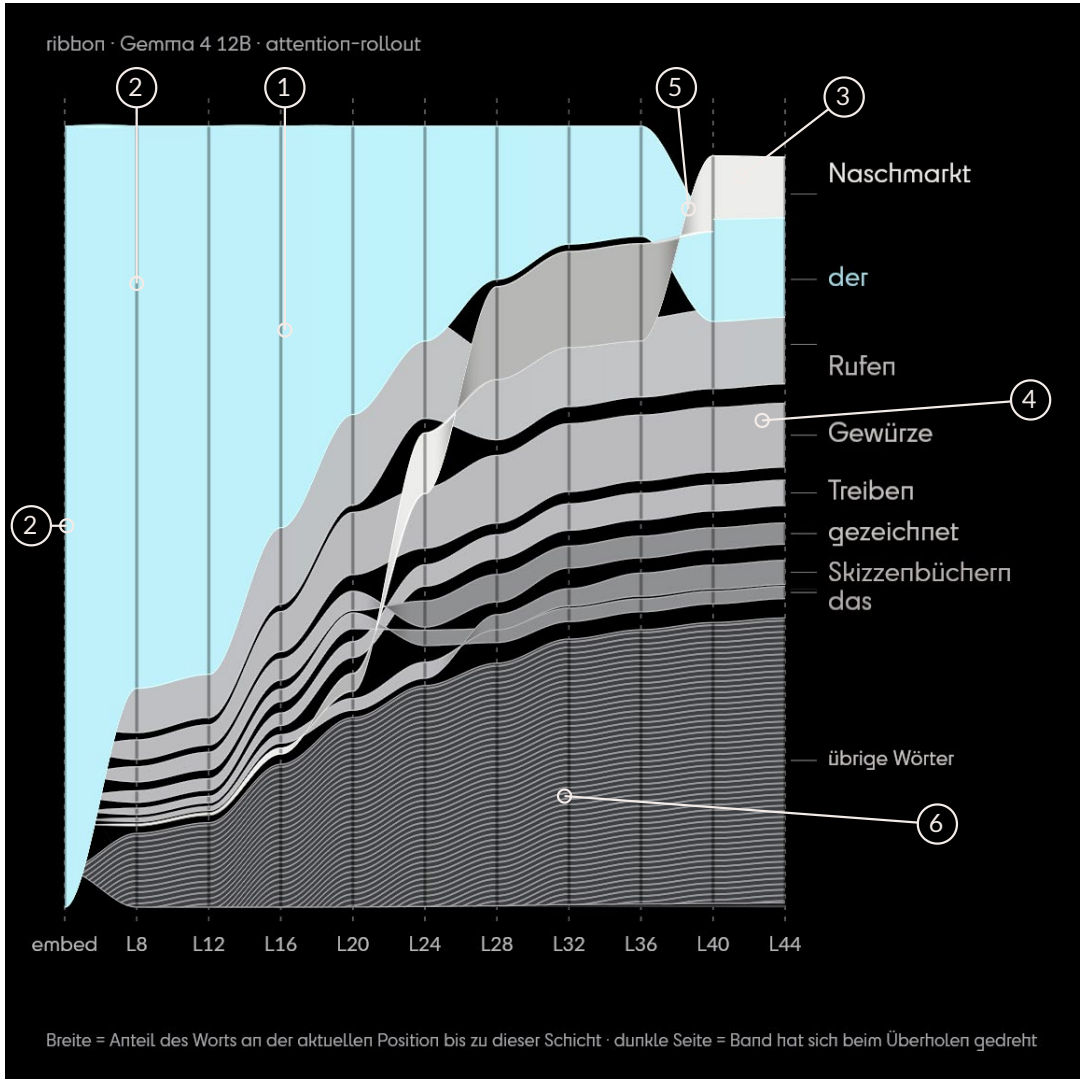
4 Rank-1 word of the lens at each layer
lensid rank 1

5 Top row: rank 1 to 8 of the sampling distribution
vtop

6 Finished braids shift left, narrow and fade; their word below
text

Resolution Type 175

Ribbon Sankey



1 Current word: how much of the position still comes from the word itself (this band also holds attention to the sink, the first token, where the model parks attention it does not need). Tint: blue dominant, red dominated
rollout, dom[36]

2 Column embed: the token alone (100 %); each later column: attention rollout up to that layer
flow L8-L44

3 Main source: earlier word with the largest share at the last column
rollout

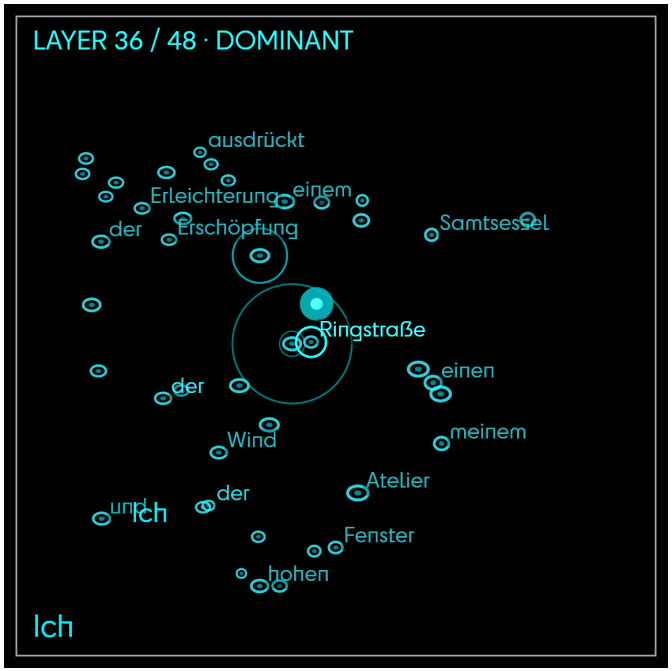
4 Band height grows linearly with the word's share in the current position; each column sums to 1
rollout, resid 0.5

5 Two words swap rank between columns: bands cross, the one with the larger jump turns and shows its dark back
rank per column

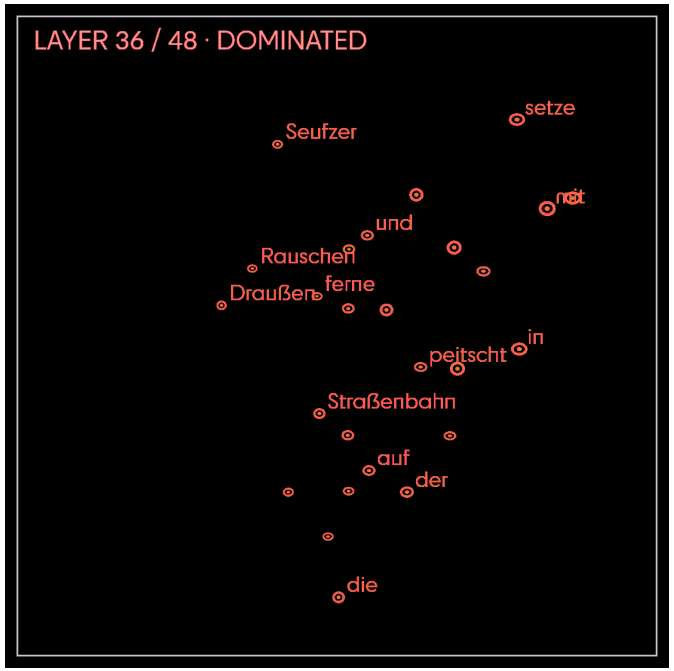
6 All other earlier words: the prompt and tokens more than 128 tokens back. Labeled only above 0.8 %)
rest

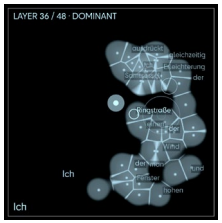
Resolution Type 159

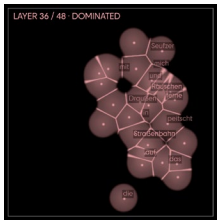
DOMINANT · FIBER



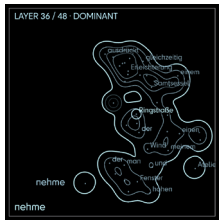
DOMINATED · FIBER



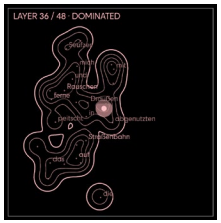


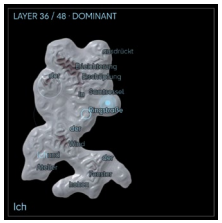


cells

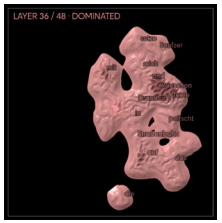


hull





tissue



Layer shown (16 to 48), the same on both walls. It changes only at sentence ends, to the layer where the last whole word is rebuilt most

trace_layers, Δ q6t, Δ domt_ft

Left wall: only tokens dominant at this layer (blue). Right wall: only tokens dominated at this layer (red). Positions come from all tokens, so both walls show one map, each its half

domt_ft

Horizontal: power reading, mixed with the second main difference as the slice turns; vertical: the main way this answer's tokens differ at this layer

domt_ft, q6t (PC1/PC2)

Cut, one form per pass: fiber = a ring per token, stretched when it moves far before the next layer; cells = a cell per token, small where tokens crowd; contours = density lines of the side; tissue = the same density as flesh, thick where many gather

q6t, domt_ft L to L+4, density

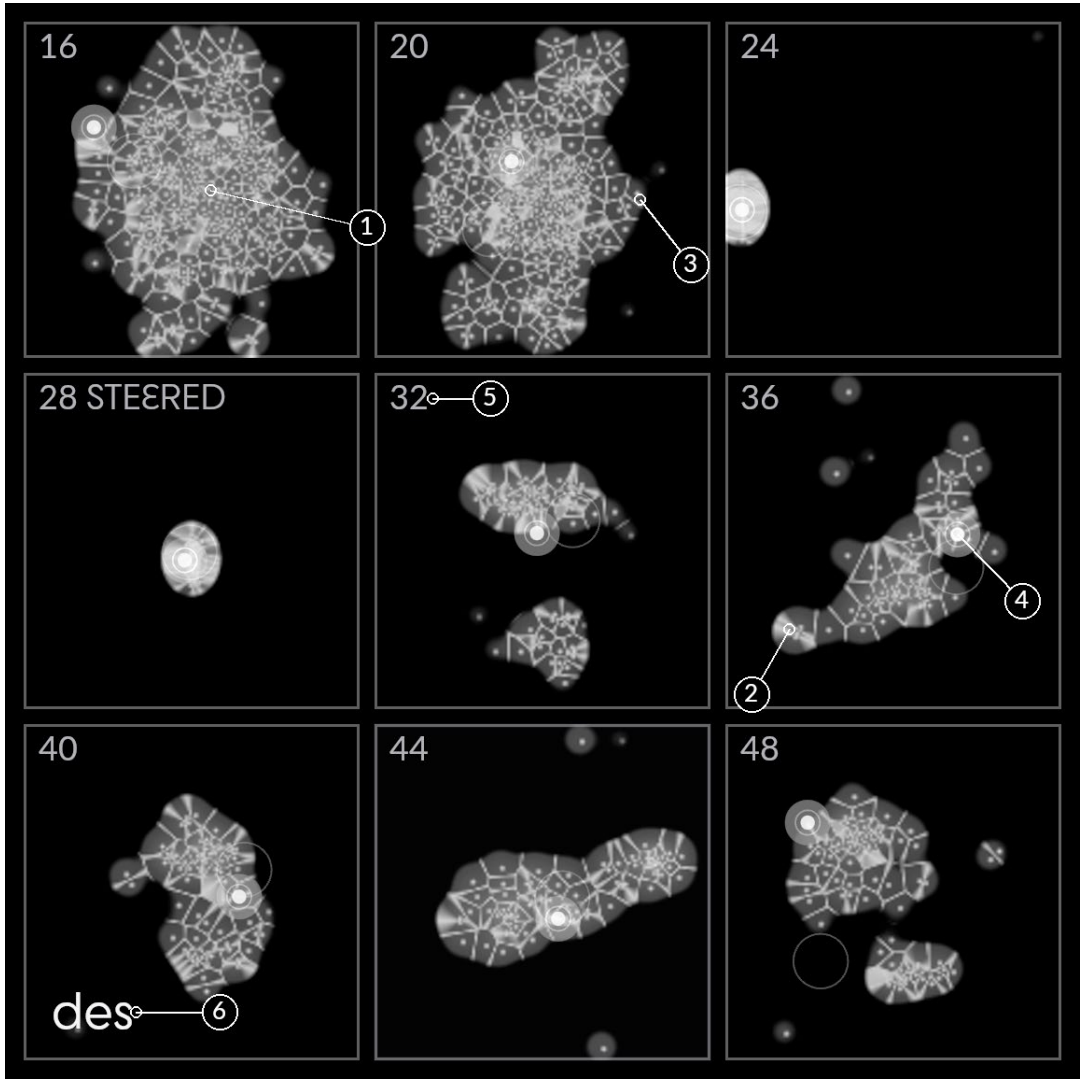
Whole words at their places; a new word flies in from the bottom line, where the last whole word stands

text

Pulse: ring around a new token, its radius grows with how much the token is rebuilt across the layers

Δ q6t, Δ domt_ft per layer pair

Resolution Type 024



1 Each cell is a token: the area closer to it than to any other token. Small cells: many tokens placed alike at this layer
Voronoi, q6t, domt_ft

2 Large or isolated cell: a token with few neighbors
q6t, domt_ft

3 Outer edge of the tissue: token density above 12 % of its peak
q6t, domt_ft

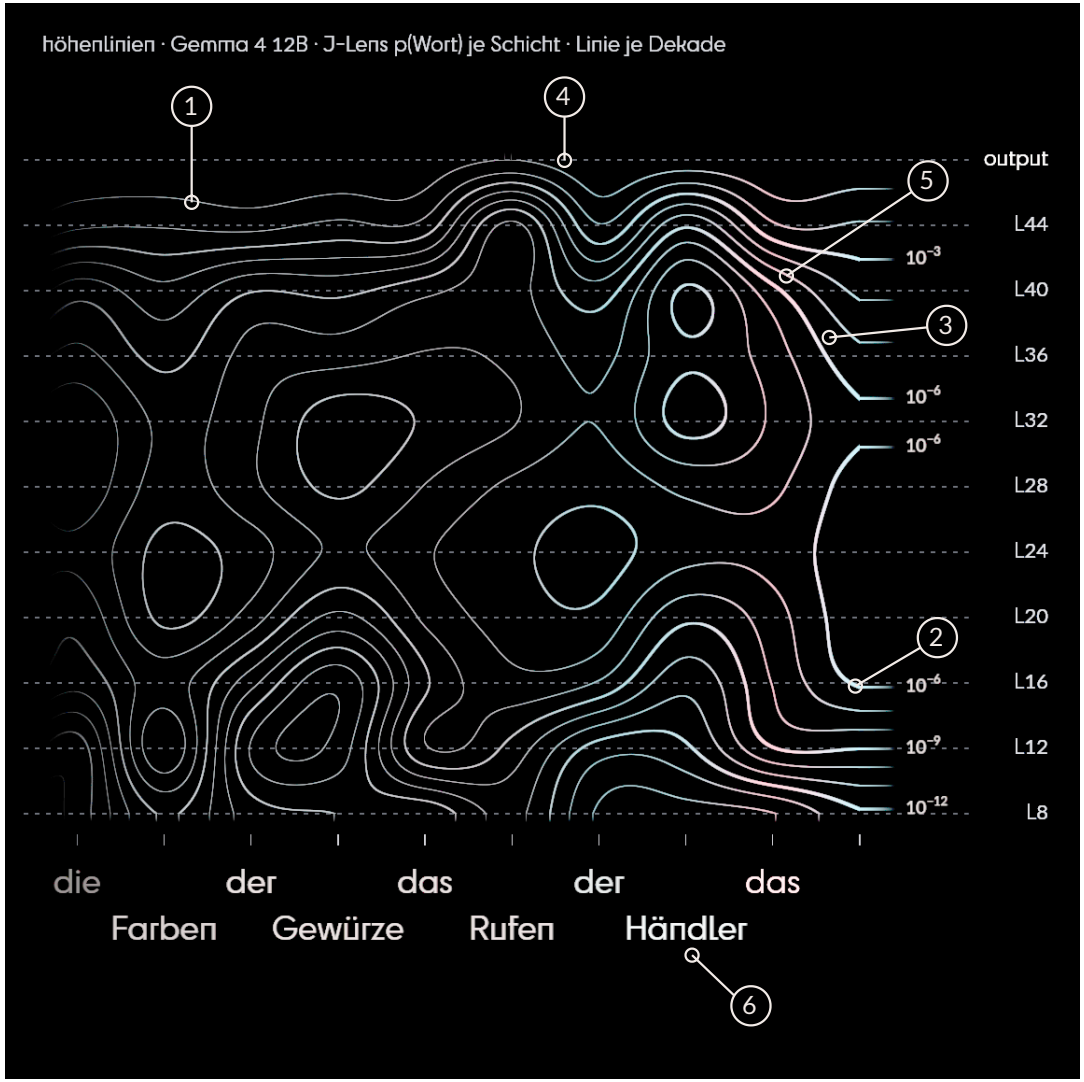
4 Newest token: bright nucleus with halo
tok

5 One tile per layer, 16 to 48. Horizontal: power reading, dominant left (mixed with the second main difference as the slice turns); vertical: the main way this answer's tokens differ, leaving the power reading out
trace_layers, domt_ft, q6t (PC1/PC2)t

6 Last whole word
text

Resolution Type 179

Contour Map



1 Height: how likely the J-lens finds the word chosen in this column, at each layer L8–L44. Each contour marks a tenfold change in probability
lensp (log10)

2 Every third contour is labeled with its probability; these labeled lines are 1,000× apart
lensp

3 Dense contours: the probability changes fast with depth, here the word is decided (mostly between L36 and output)
gradient of lensp

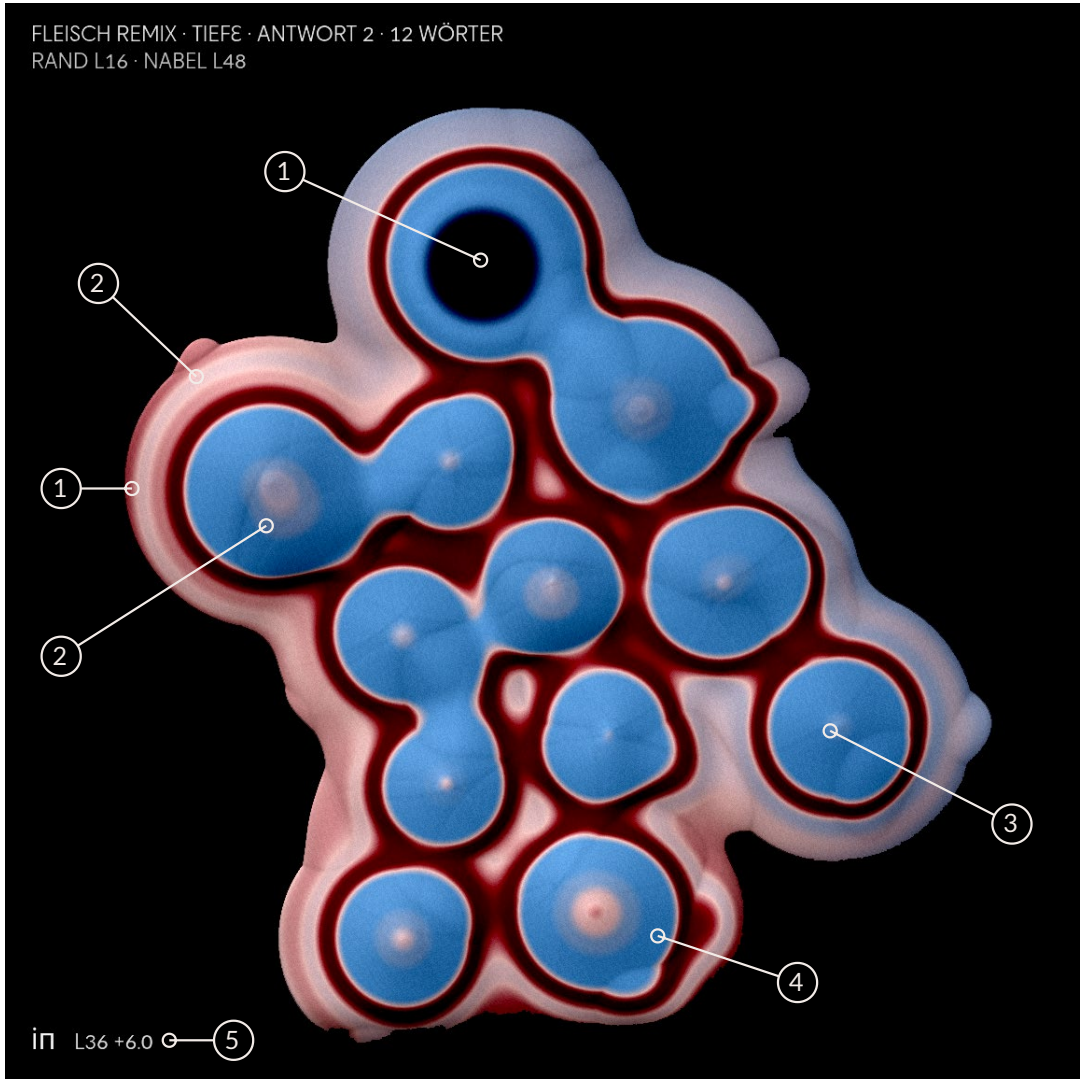
4 Top row: probability at the output
p

5 Line color per column: blue dominant, red dominated, gray near 0; grays with age
dom[36]

6 Whole word of the column, shown once it is complete
text

Resolution Type 181

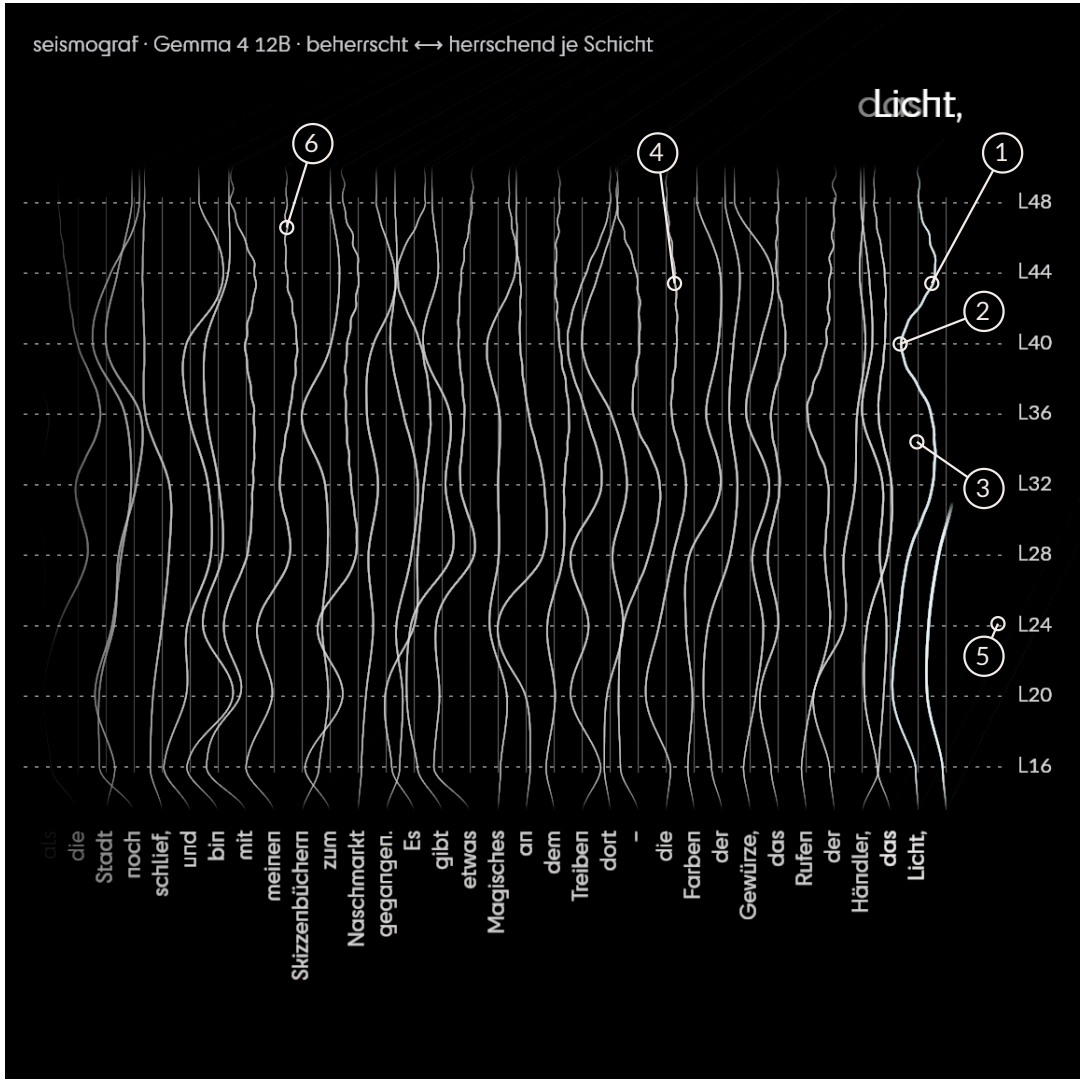
Tissue Remix: Depth



1. Cross-section through depth: the rim of the body is layer 16, the dark navel at its thickest point layer 48; the bands between are layers 20 to 44 and bend with the folds
   trace_layers
2. Band color: power reading of the word lying there, at that layer; blue dominant, red dominated
   domt_ft L16…L48
3. One lobe per whole word. It buds from the earlier word (of the last 40) with the most similar power profile across depth, so lobes gather into families of alike profiles
   domt_ft profile
4. Lobe size: mean strength of the power reading across all layers
   |domt_ft|
5. Last whole word at the bottom left with its reading at layer 36
   text, domt_ft L3
6. When the body grows too large, older lobes shrink and vanish; the newest 36 stay
   word order

Resolution Type 180

Seismograph



1 Newest word: trace colored by its power reading at L36, blue dominant, red dominated; older traces gray
dom[36]

2 Swing per layer: right = dominant, left = dominated. Large values are compressed so strong words don't swamp the rest; a word of several tokens shows their average. Under steering the fader's push is included
domt, tanh

3 Zero line of each trace
domt = 0

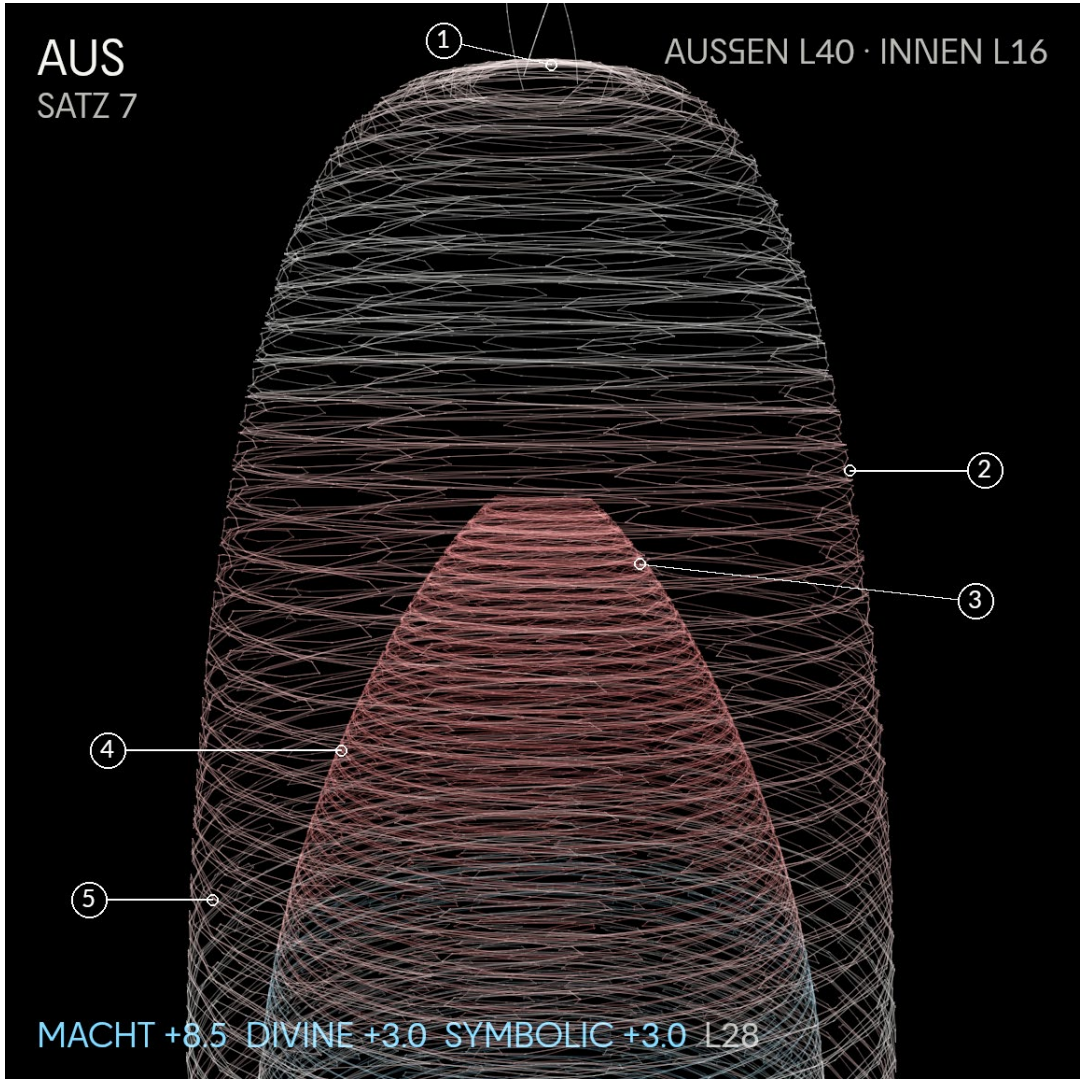
4 One trace per whole word, newest on the right
text

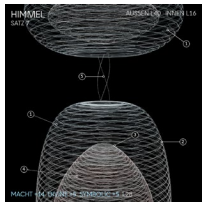
5 Rows: layers 16 (bottom) to 48 (top)
trace_layers

6 Jitter near the top grows with the entropy of the word's choice (how undecided the model was in choosing the word)
e (entropy)

Resolution Type 177

Looped Form





Close-up on the bulge of the most recently finished sentence, the mesh fills the frame; outer shell read at layer 40, inner shell at layer 16, color at full strength

1 One chamber per sentence of the answer, height = its number of tokens; the newest sentence grows at the bottom, older ones leave at the top
text

2 Width of the outer shell = mean power reading of the sentence at layer 36: wide = dominant, narrow = dominated. The bulge sits lower when the sentence turns more dominant toward its end
domt_ft L36

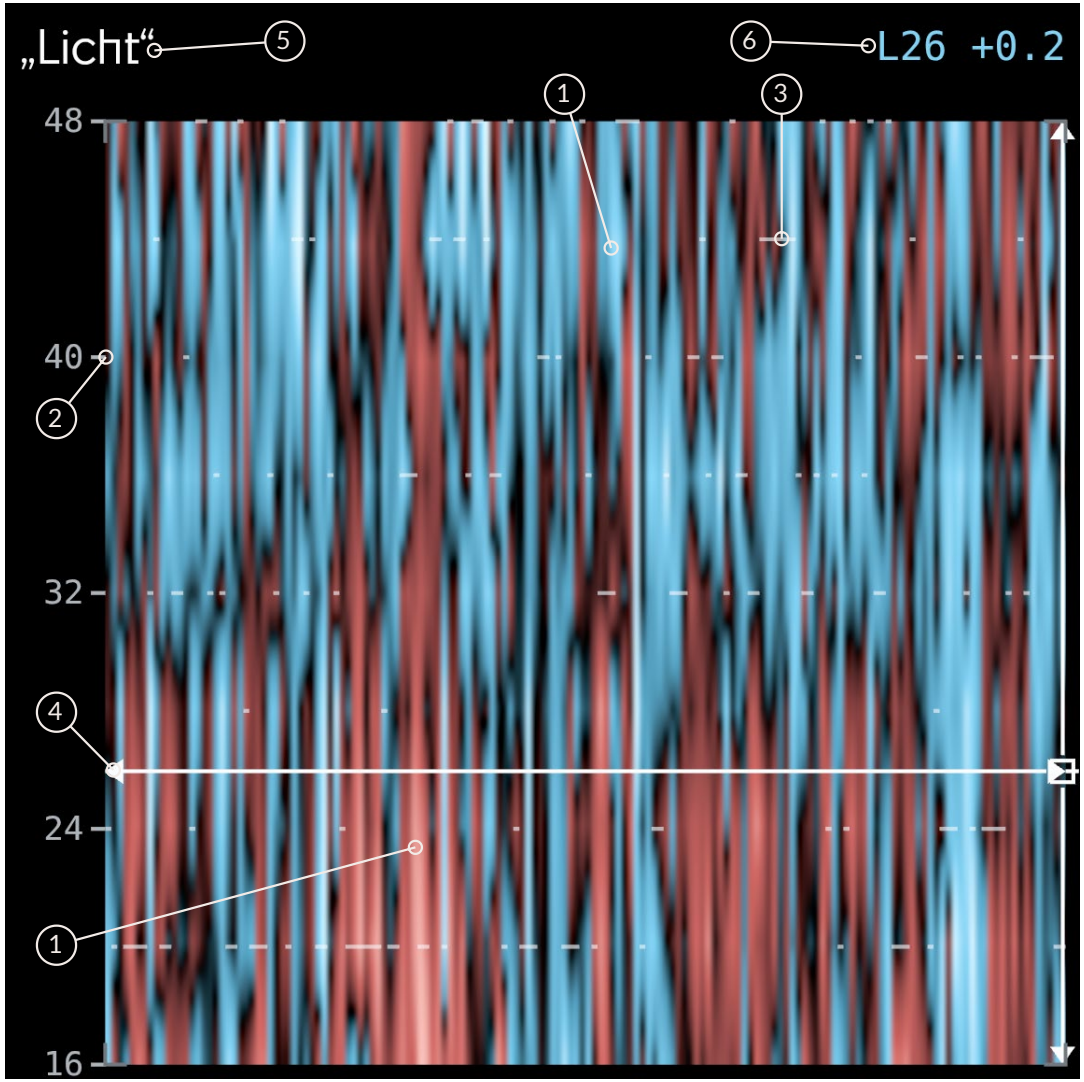
3 Inner shell: same sentence at layer 20; the gap between the shells shows how the power reading changes from layer 20 to layer 36
domt_ft L20 / L36

4 Color per row of loops: blue dominant, red dominated, white near zero
domt_ft L36 (inner L20)

5 Loop length = entropy of the token at that point: short and tight = certain, long and loose = uncertain
e

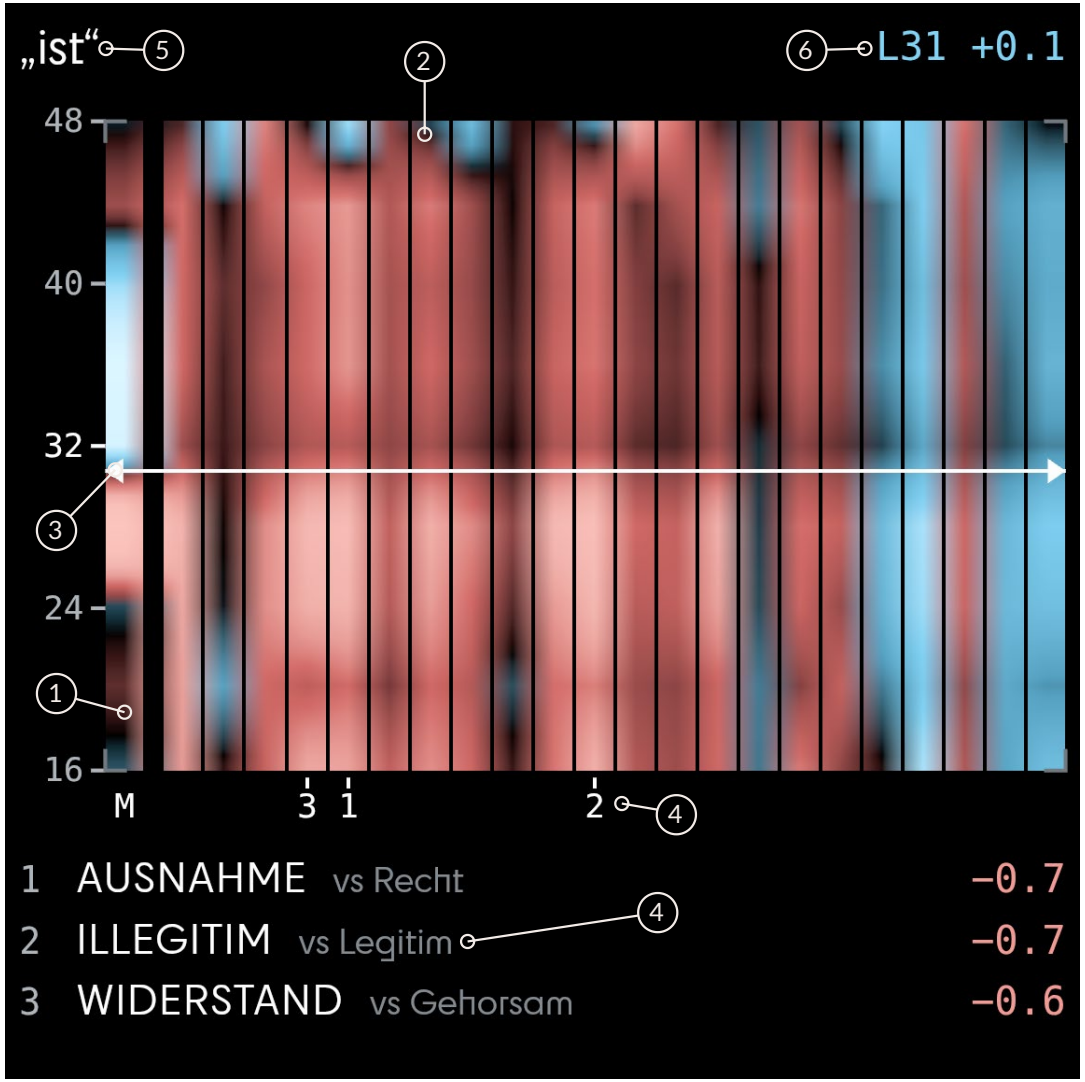
Each token twists the form in proportion to its surprise; surprising answers spin it, predictable ones let it settle
s

Resolution Type 010



1. Each column is a token of the answer, oldest left, newest right. Color: blue dominant, red dominated, brightness grows with |power reading|; special tokens dimmed
   domt_ft
2. Height = layer, 16 at the bottom to 48 at the top; 9 measured layers, linear in between
   trace_layers
3. Gray steps: for each token the layer where its power reading jumps most from the layer below
   Δ domt_ft
4. Crosshair: newest token at that layer
   domt_ft
5. Last whole word
   text
6. Power reading of the newest token at the crosshair layer
   domt_ft

Resolution Type 011



1 Column M: power reading of the newest token, layer 16 (bottom) to 48 (top); blue dominant, red dominated
domt_ft

2 22 columns, each is one direction of power (named by its two poles), defined at layer 28. Per layer: how strongly the state points along it, compared with the rest of the answer. Blue = first pole of the name, red = second
ax = (h − μ)·v⃗

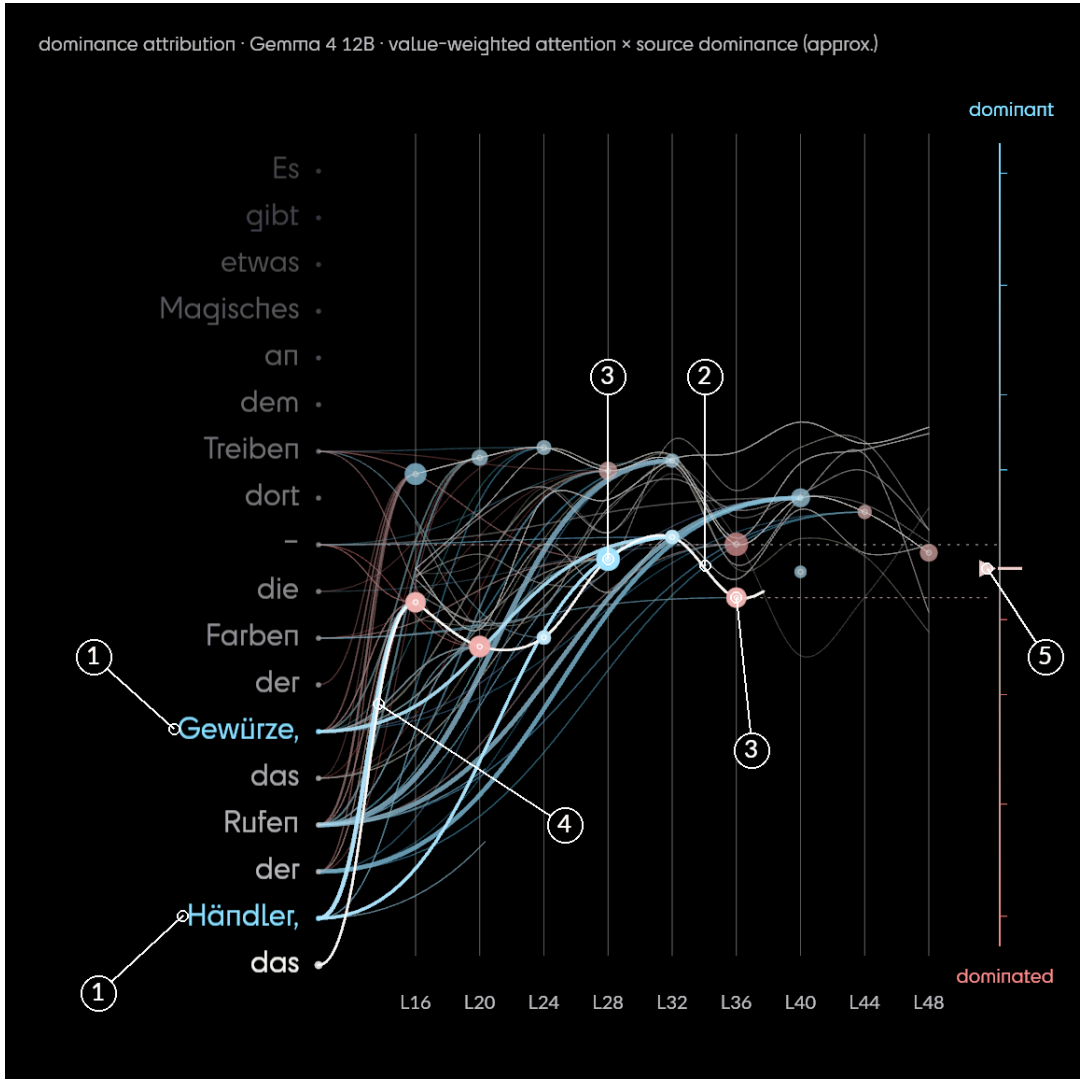
3 Crosshair: layer where the newest token's power reading jumps most
Δ domt_ft

4 List 1 to 3 (also under the columns): the axes with the largest deviation at the crosshair layer, pole vs counterpole and value
ax

5 Last whole word (the values belong to the newest token)
text

6 Power reading of the newest token at the crosshair layer
domt_ft

Resolution Type 164



1 Left: the latest words, newest at the bottom; blue or red marks the source word of a main path
text

2 Across: layers 16 to 48. Height: the word's power reading, dominant above, dominated below. White path = newest word, gray lines = the words before
domt

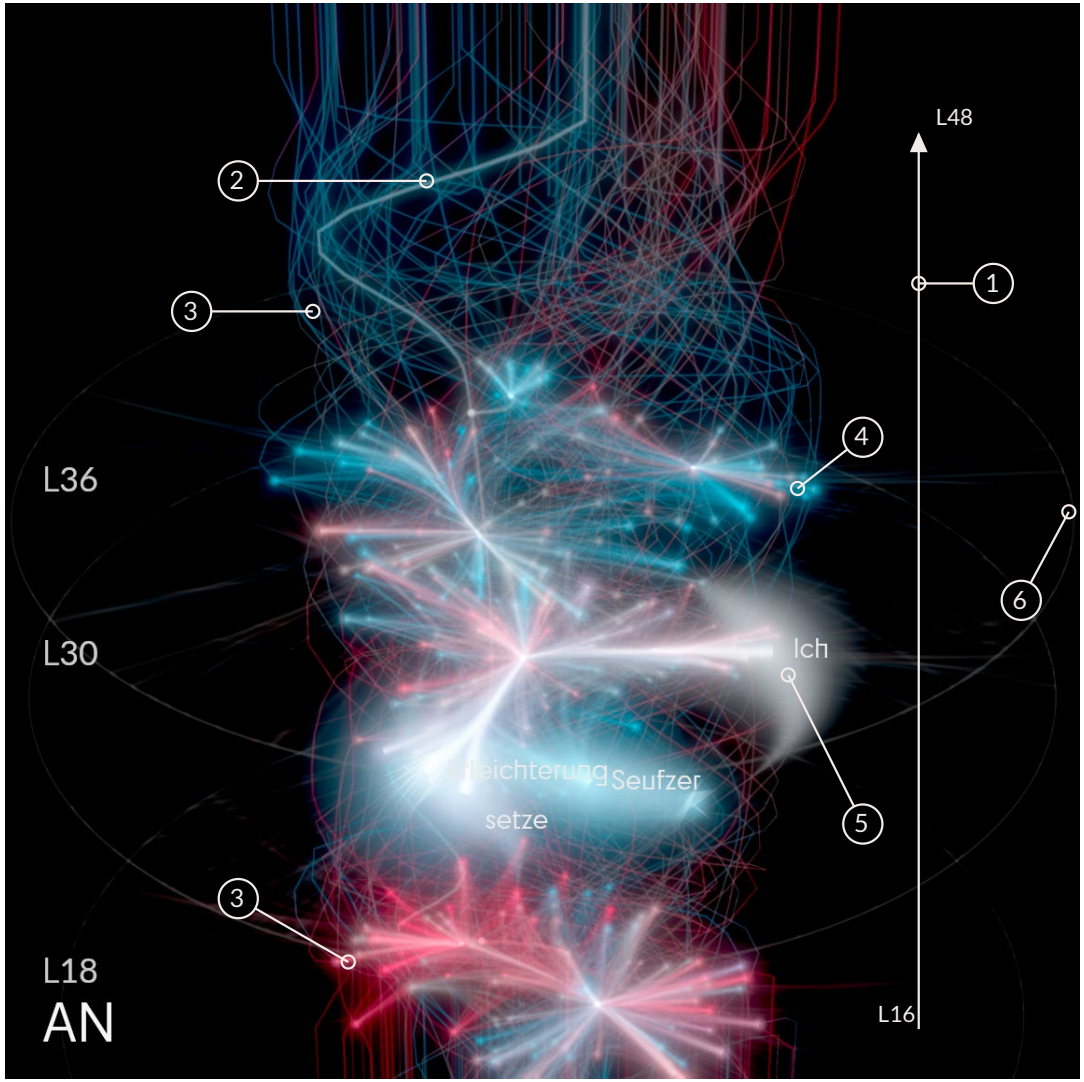
3 Nodes: change from the previous layer, size = amount; blue pushes toward dominant, red toward dominated
Δ domt

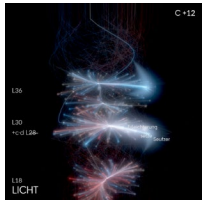
4 Curves from earlier words: attention at that layer times that word's own reading there
attr.src × domt

5 Marker on the scale: the newest word's reading at layer 36
domt[36]

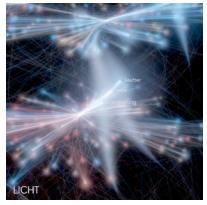
Notch at layer 28 and value below: steering layer and fader setting
c

Resolution Type 091





Each new token sends a pulse of light from its sources along the attention fibers to itself, then up its own fiber to layer 48



The camera flies in to one of the most read tokens (a newly risen one first), circles it and pulls back out to the whole bundle

1 Vertical axis = depth in the model, layer 16 at the bottom to layer 48 at the top
trace_layers

2 Vertical fiber = one token through the layers. Its position per layer: the two main ways this answer's tokens differ, leaving the power reading out; close = represented alike
q6t per layer (PC1/PC2)

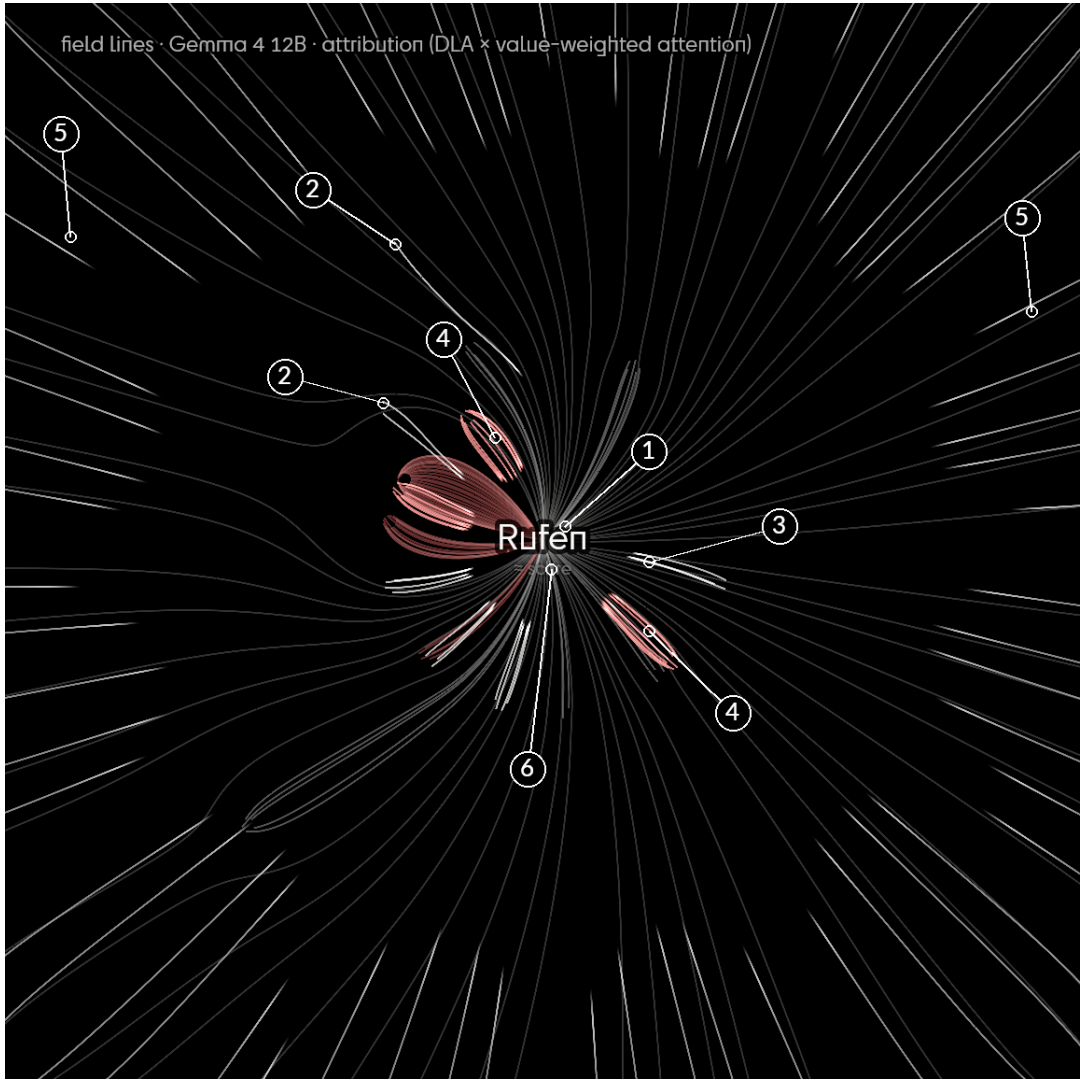
3 Fiber color per layer: blue dominant, red dominated, gray near 0; at the steering layer gray with a ring marked + c·d (fader setting c times the fader's direction d)
domt_ft, c

4 Horizontal fibers in the three attention layers (L18, L30, L36) run from the reading token to its source; brighter and thicker with more weight
att L18/30/36, top-8

5 Swelling with a glowing core: the four tokens most read by later tokens swell on their own fiber at that layer and carry their word; the fibers of their readers thicken toward them. Other tokens: a small point of light
att, ∑ received

6 Ring at the edge of each attention layer: prompt and sink; a root runs from the ring to each token, weighted by what that token reads there
att.prompt + att.bos

Resolution Type 170



1 Newest whole word in the center; every field line ends in it
text

2 Each line starts at an earlier word of the answer, placed on an unseen spiral around the center: newer words close in, older further out
word order

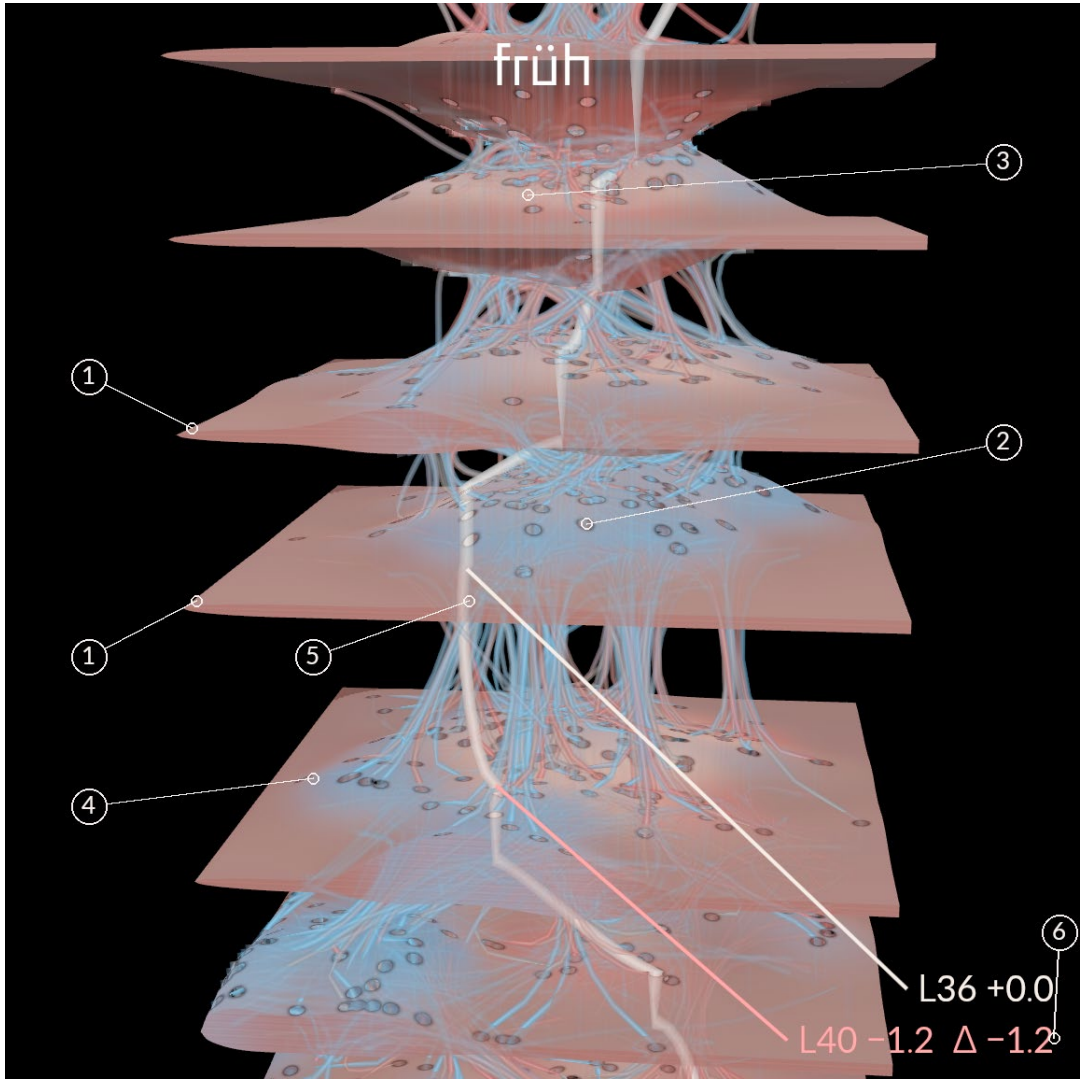
3 Number of lines from a word grows with how strongly the new word draws from it: how much it reads that word at each layer, weighted by how much that layer pushes directly toward the new word
attr.src, attr.dla

4 Red: lines of the three strongest sources
attr.src, attr.dla

5 Long lines from the edges: what the new word draws from the prompt
attr.rest.prompt

6 Below the word after ≈: the J-lens word at the deepest layer, when it differs
lens1

Resolution Type 067



1 Plates = layers 16 to 48, early at the top
trace_layers

2 Puncture = where a word sits on this layer; words represented alike lie close (power direction removed)
q6t per layer

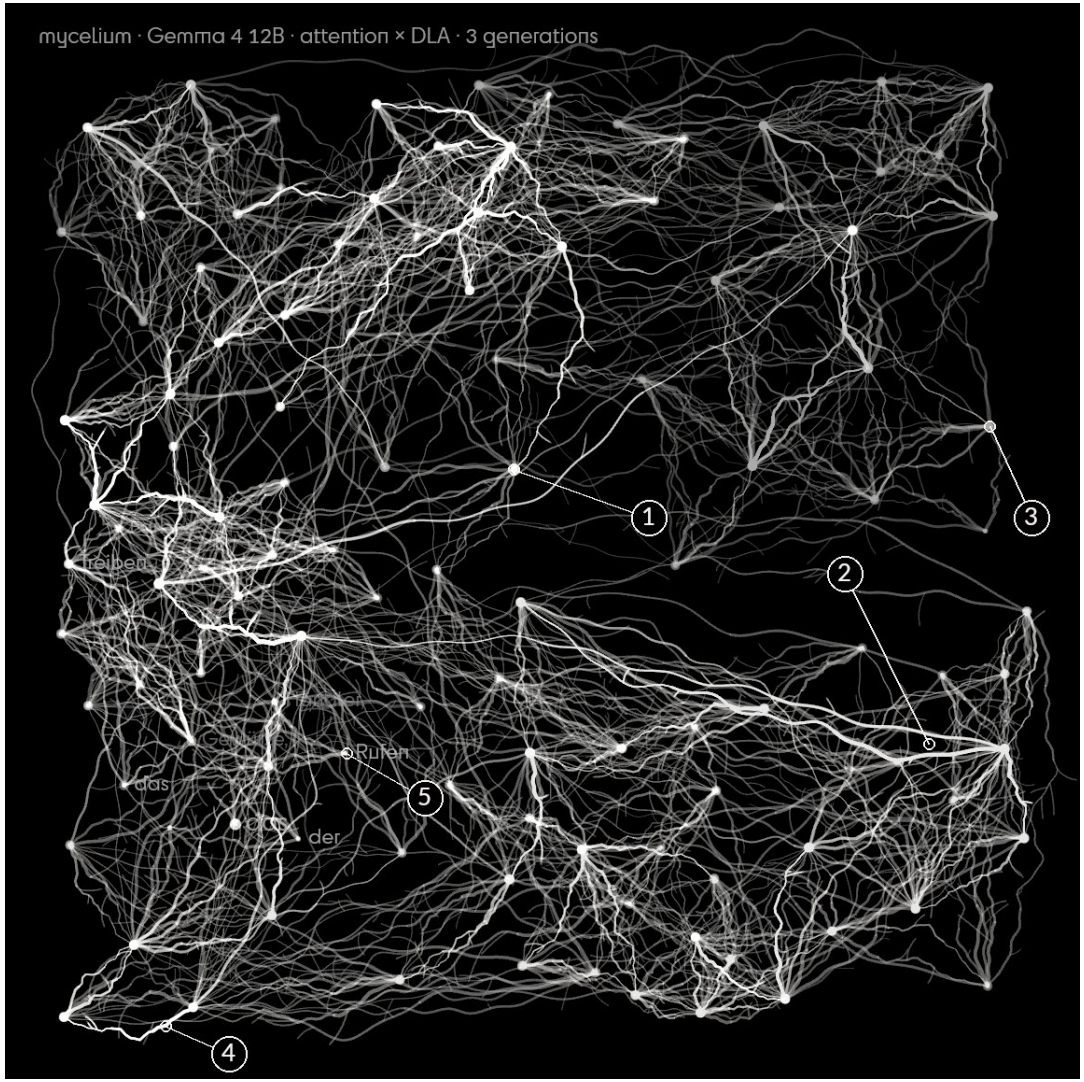
3 Plate height: dominant words lift the plate, dominated words press it down (density times power reading)
domt

4 Color of plates and channels: blue dominant, red dominated
domt

5 Bright channel: the newest word through all layers, its reading written at the punctures
domt per layer

6 The camera moves to the layer where the newest word's reading jumps most; Δ = the jump
Δ domt

Resolution Type 171



1 Node: a word of the answer, placed next to its strongest source
text

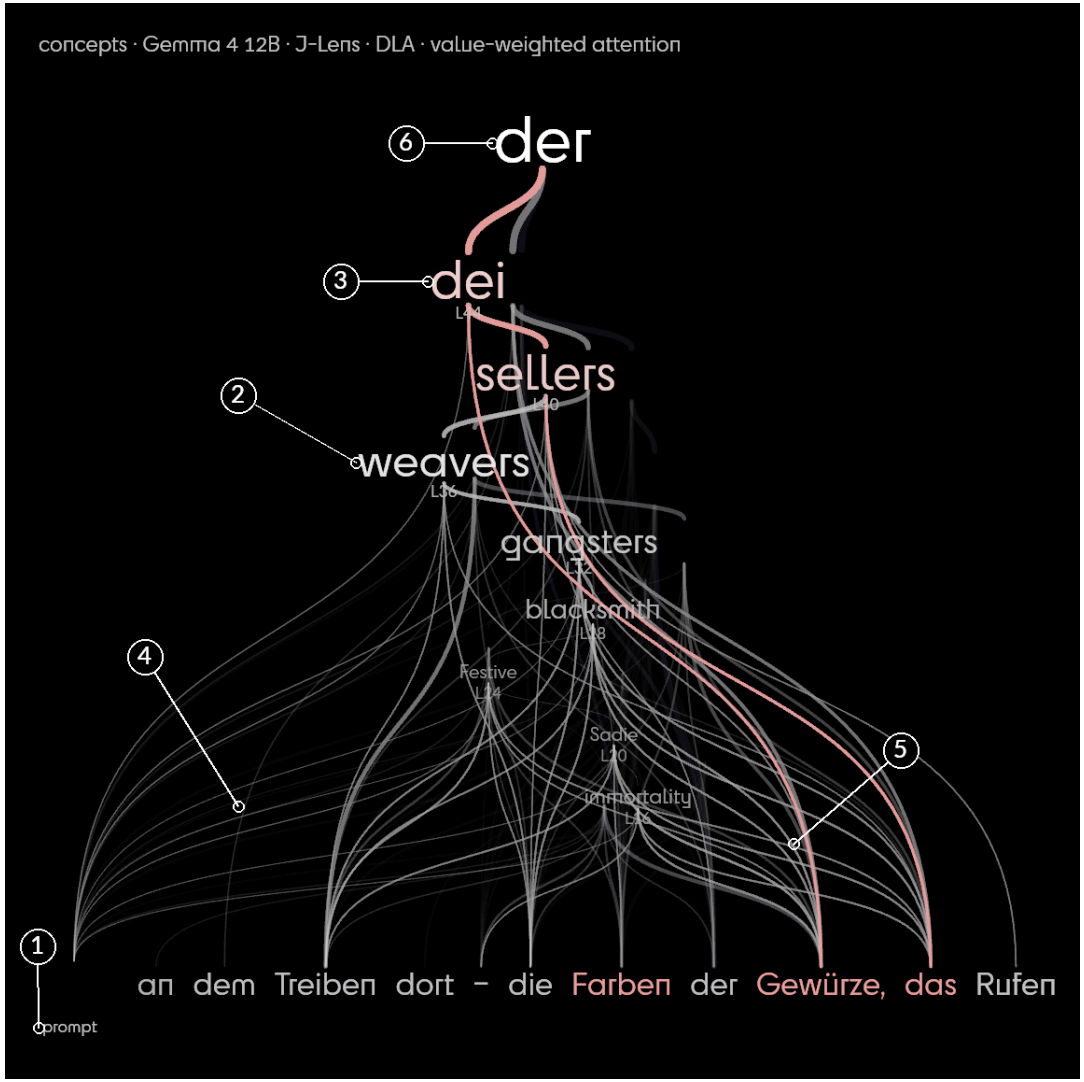
2 Roots: from each new word back to the earlier words it draws from, weighted by how strongly it reads them (attention) and how much that layer pushes toward the word
attr.src × attr.dla

3 The roots branch on into the sources of those sources, up to three generations, thinner and paler
attr.src

4 Thicker = stronger; a path used again grows thicker and brighter
attr.src (√)c

5 Label: the new word bright, its direct sources gray
text

Resolution Type 165



1 Bottom: the latest words of the answer; 'prompt' stands for the question
text, attr.rest

Middle: concepts, the words the J-lens reads layer by layer at the position of the next word. Alike words in neighboring layers merge into one concept; early ones low, late ones high
lens1

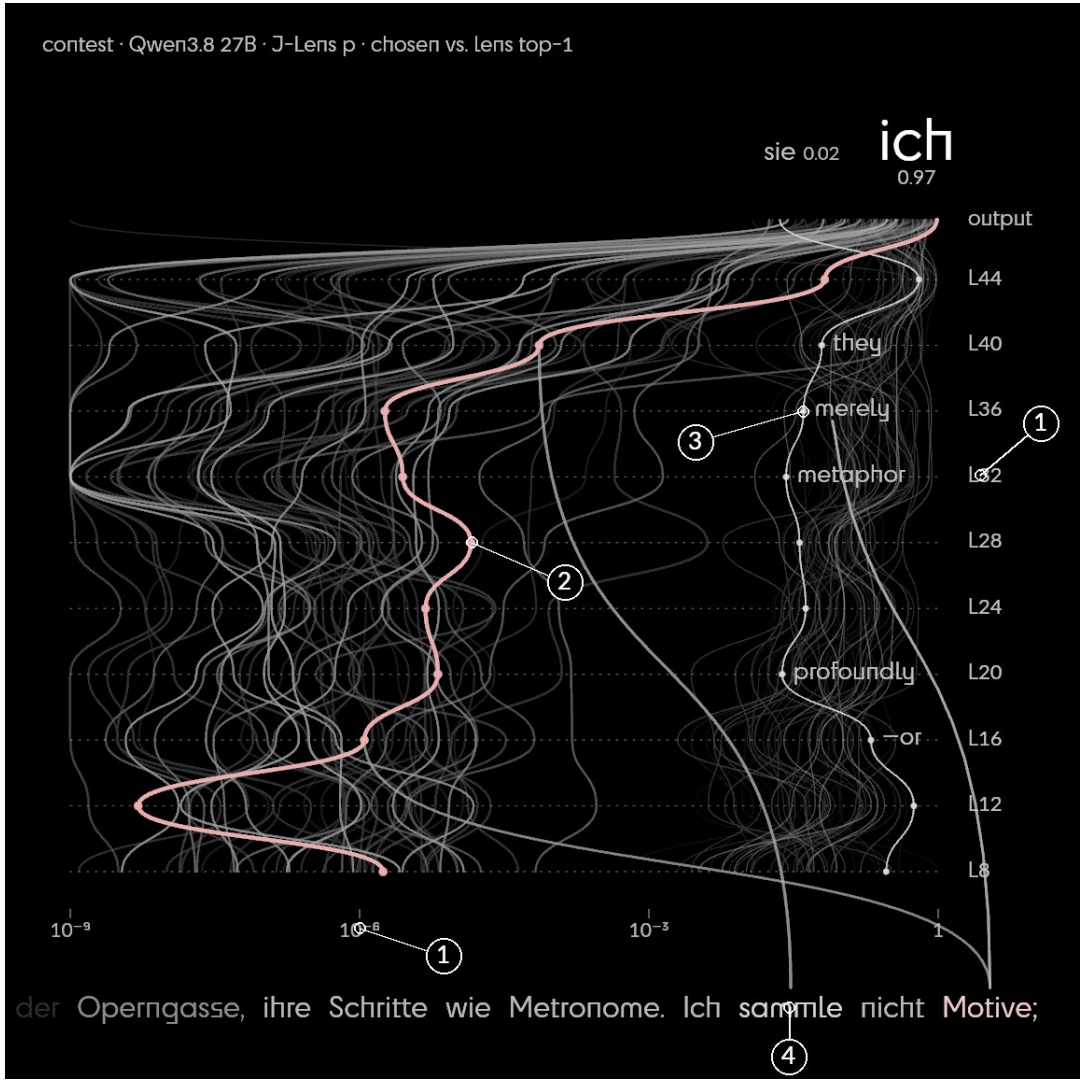
2 Size and brightness: how much the layers of this concept push directly toward the new word
attr.dla

3 Gray curves: from earlier words into a concept, by value-weighted attention
attr.src

4 Red curves: the three strongest paths, source word to concept to the new word
attr.src × dla

5 Top: the new word
text

Resolution Type 167



1 Across: how likely a word is, on a log scale from one in a billion (left) to certain (right). Up the layers, early at the bottom; the actual output at the top
lensp, p

2 Thick red path: J-lens probability of the chosen word, layer by layer
lensp

3 Thin gray path: the word the J-lens ranks first at each layer, written beside it where it changes
lens1, lensp1

Ring and layer name: from this layer up the chosen word stays in front
lens1, rank

Top: the chosen word at its output probability, the runner-up small beside it
p, alt

4 Bottom: the latest words. Paths of earlier words stay gray and fade
text

Resolution Type 188



1 One loop per token in text order, 12 to a row, rows knitted back and forth; a new answer starts further down
text

2 Loop length = entropy: short and tight = certain, long and loose = uncertain
e

3 Long diagonal thread: the new token hooks into the earlier word of its answer it attends to most, if more than 10 tokens back; brighter and thicker = more attention
att L36, top[0]

4 Faint second thread: the same step without steering. Under steering it slides along the row by how much more likely steering made the chosen tokens; where the two threads part, moiré appears
p, cf.p_tok, cf.e

5 Fader settings, lower left
c

6 A surprising token plucks the mesh; the wave runs through neighboring loops and along the attention threads
s

ABLAZE - a tomography of artificial power

Sylvia Eckermann and Gerald Nestler, in collaboration with Aaron Berman

EXHIBITION: 11.–15. Nov 2026 15–18 h, art space Eckermann Nestler, Neulinggasse 9, 1030 Vienna
www.thefutureofdemonstration.net/ablaze
studio@theoriesinmind.net

THE FUTURE OF DEMONSTRATION Season 6