



POLYTYPIC THINKING

When one correct answer becomes a framing error

Humans have a very old superpower: we see something new and quickly decide what it resembles. Without it, we would be terrible at surviving our own mornings. A cup looks like a cup, a door like a door, a storm cloud like something under which a picnic is probably a poor strategic decision, and a strange sound in a dark stairwell deserves attention before we open a philosophical seminar on the nature of sound. The brain compresses the world. It takes an impossible number of details and produces a working version: edible, familiar, dangerous, human, tree, apparently Monday again.

This ability is not the enemy of thought. It is one of the conditions that make thought possible. The problem begins later, almost invisibly, when “this resembles X” becomes “this is X,” and then something stronger still: “everything important about this can already be understood through X.” A cheap preliminary label acquires the status of an essence. The map stops helping us navigate and starts issuing orders to the territory.

In this essay MET[À]CADEMY proposes a working name for another mode of seeing: **polytypic thinking**. It is not an attempt to abolish types, categories or generalization. Nor is it a way of saying that “everything is everything,” so nothing needs to be checked anymore. Polytypic thinking begins with a simpler and more uncomfortable proposition: **a type can be a useful projection of an object without being its final essence**. The same process may lawfully look different at different times, scales, relations, contexts and positions of observation. Sometimes one type is enough. Sometimes several are required. And sometimes the most important result is the recognition that the problem is not the phenomenon but the box we prepared for phenomena.

This is a working MET[À]CADEMY concept, not a claim that humanity has somehow failed to notice the complexity of categories until now. The sciences of classification already contain polythetic and polytypic approaches, prototype theory, family resemblance and other ways of describing classes without one mandatory feature shared by every member. Our question begins at the next step: what happens when different types depend not only on sets of features but on time, scale, observer, relation and context, and when the phenomenon itself sometimes shows that the old grammar of types is no longer adequate? Here, ‘polytypic thinking’ names not a new sticker for old categorical fuzziness, but a discipline that refuses to let classification close an investigation before the phenomenon does.

The world never signed our classifier

Imagine a forest. For a biologist it is an ecosystem. For someone who lives beside it, part of home. For a firefighter, an environment of risk. For a company, a resource. For a child, a territory of adventure. For a community, memory, place names, routes and stories. For a satellite, spectral signals. For a river, if rivers had academic departments, the forest might be part of a regime of water, shade and soil.

Which of these forests is the real one?

The question has already started to malfunction. Not because truth does not exist, but because we have mixed several different questions and demanded that one label perform the work of an entire world. The forest does not become arbitrary because it can be truthfully described in different ways. On the contrary, the multiplicity of relations shows how poor a description can become once it declares itself the only legitimate one.

The MET[À]CADEMY Manifesto states the point more broadly: “No discipline owns the question. No institution owns knowledge. No way of seeing receives a monopoly on the whole.” This is not an invitation to blend every worldview into

one friendly semantic soup. On the contrary, the Manifesto asks us not to lose contradictions, negative results or abandoned paths, and to let different knowledge systems carry their own Knowledge / Evidence Passports. Different forms of evidence and verification enter a shared graph only with their provenance, limits and methods visible. Biology does not have to become mythology, and mythology does not have to become bad biology. Lived experience need not cosplay as a laboratory experiment. But a laboratory experiment does not acquire the right to declare everything for which it lacks an instrument meaningless by default. The task is not to make all ways of seeing identical. The task is to preserve the difference between them and prevent one from quietly becoming customs control for the whole of reality.

This matters for simpler things too. A person can be strong and vulnerable, generous in one relation and cruel in another, an excellent manager and a terrible friend, brave in physical danger and helpless before the need to say one honest sentence. Stereotypical thinking loves administrative cleanliness: strong or weak, good or bad, leader or victim. But a person is not required to maintain compatibility with our input form. Otherwise we do not get psychology, we get an Excel sheet with a personality: the cell is filled, the person is missing.

Typification becomes harmful not when we notice a pattern. It becomes harmful when **the pattern replaces the particular thing**. The most convenient person for a system is the one who has already agreed to become their own profile. The most alive person usually ruins the statistics.

We have known for a long time that categories are softer than they look

Polytypic thinking does not emerge from a vacuum. The psychology of categorization has long shown that natural human categories are often not structured like legal forms with necessary and sufficient conditions. In classic work by Eleanor Rosch and Carolyn Mervis, categories were shown to have prototypes and “family resemblances”: some members feel more typical than others, and members of one category need not share a single magical feature present in all cases without exception. Their 1975 paper became one of the foundations of modern prototype approaches to categorization.

Later reviews reveal an even more complex picture. Human categorization does not appear to rely on a single mechanism. We use rules, remember examples, construct prototypes and switch strategies according to the task. J. David Smith’s review makes a particularly useful point: prototypes are valuable precisely because they compress experience, but compression inevitably discards something. In 1975, anthropologist Rodney Needham described polythetic classification: a class may be held together by family resemblances across many properties without any single property being mandatory for every member. In contemporary knowledge organization, Michael Kleineberg argues that monothetic and polythetic modes of classification have different strengths and can coexist as different levels of classificatory cognition; his 2022 overview is available at <https://doi.org/10.5771/9783956509568-159>. This lineage matters. MET[À]CADEMY is not inventing categorical plurality; it is asking about the next boundary, the plurality of the conditions under which a type becomes visible.

That is important background, yet polytypic thinking asks a further question. What if the issue is not merely that category boundaries are fuzzy? What if the same object lawfully belongs to different types in different relations? What if its type changes over time? What if two observers obtain different but compatible descriptions? What if one scale reveals chaos while another reveals stable order? And what if, at some point, not only our answers but **the grammar of available answers** proves insufficient?

That is where what we call polytypicity begins.

A shadow is not light’s enemy

Take the simplest image: light and shadow. Everyday language likes to place them opposite each other. Light here, shadow there. Positive and negative. Visible and hidden. Day and night. A convenient pair on which humanity has built roughly half of its metaphorical real estate.

Physically, however, a shadow is not an independent substance that arrived to fight light. It appears in a relation between a light source, a body, a space and an observer. The same object can be illuminated on one side and dark on another. Shadow can hide detail, but it can also reveal an object’s form more clearly than uniform lighting. In photography and painting, shadow is not merely “less light” but a carrier of volume. In heat it may be protection. In danger it may be cover. In symbolic language it may become what a culture refuses to see and yet cannot remove without making its image of itself incomplete.

So what are we looking at: light or shadow?

Sometimes the answer really is simple. But sometimes the better questions are: light for whom, shadow relative to what, at what moment, on which surface, in which function? And then we discover that the contradiction was not living in the phenomenon at all. It was living in a question that was too small.

Polytypic thinking does not teach us to call shadow light. It teaches us not to confuse an opposition in language with independence in the world.

A seed is not a tree. Yet “not-a-tree” is a terrible description

Time does even stranger things to types. A seed is not a tree. A sapling is not a mature tree either. A felled tree no longer lives as a tree, but it may remain material, habitat for other organisms, part of the soil, part of the history of a place. If we classify every moment separately, we obtain several different bodies. If we see only continuity, we erase real transformations. We need both views: what changed and what travelled through the change.

The same is true of a wound and a scar. A scar is a trace of damage and a trace of repair. At the moment of injury it belongs to one story; years later, another. For a doctor it may be tissue with specific properties; for the person, memory; for an investigator, evidence; for an artist, part of an image. None of these descriptions should automatically consume the others.

History suffers from our love of single moments too. We place a date beside an event because calendars are useful. But the time of manifestation is not always the time of genesis. Revolutions develop before the day that later reaches the textbook. Technological shifts begin long before mass adoption. A cultural form may receive a familiar name far later than some of the relations, practices and recurring tasks from which it assembled.

This problem became especially visible in our research line on Ukraine’s long cultural memory, “The Country of Edges.” The binary question “is there direct genetic continuity or not?” proved insufficient not because genetics is unimportant, but because cultural continuity may have more than one carrier. Geography, recurring frontier conditions, forms of mobility, warrior practices, symbols, social forms, language, ritual and memory may follow different trajectories. Some can break, some merge, some reappear through similar tasks, some pass into a different cultural body. “Direct descendant” and “no connection whatsoever” can become two equally poor buttons.

Polytypic thinking does not provide the answer in advance. It simply refuses to kill the question before we have distinguished **which kind of continuity we are discussing.**

The scale on which disorder suddenly becomes order

A single ant is not an ant colony. A single cell is not an organism. One car is not a city’s transportation system. One conversation is not a culture. One pixel is not a photograph. Yet we often behave as if a type found at one scale were automatically inherited by every other scale.

Up close, a process may look chaotic. At a larger scale, a stable form may appear. The noise of individual fluctuations may carry a slow rhythm. Local behavior may be simple while collective behavior produces properties absent from every participant individually. And the reverse is also true: a beautiful macroscopic order may depend on millions of local conflicts, errors and compensations.

So the sentence “this is chaos” may be correct and still insufficient. Chaos at which scale? Is there recurrence at another? Does a form persist over time? Is there structure in the distribution of what appears random locally? Have we called something “noise” merely because our instrument does not yet know how to read it?

For science this is not exotic. New levels of description continually create new objects: temperature is not a property of one molecule in the same sense in which it is a property of a macroscopic system; an economy is not reducible to the psychology of one buyer; a language does not exist inside a single word. Yet in everyday and machine reasoning we still move types across scales as though scale were merely a camera zoom.

Polytypic thinking reminds us that **changing scale can change not only the amount of detail but the type of phenomenon that becomes visible.**

The observer is not standing outside the world

Another habit comes from the dream of an absolutely external observer: we tend to think that a description belongs only to the object. Yet many meaningful properties exist in relations.

For a doctor, a cough may be a symptom. For the patient, an exhausting part of the day. For an epidemiologist, a population signal. For an employer, a reason for absence. For a child in the next room, a sound that means a parent is ill. The event is one, but its role changes depending on the system in which the question is asked.

This does not mean “everyone has their own truth” and verification can be thrown out the window. Quite the opposite. Multiple perspectives require more discipline. We must state who is observing, what can be measured, which question is being asked and where the observer’s authority ends. The Manifesto offers a compact principle: **“Everything may enter the graph. Nothing enters without provenance.”** Provenance is needed not only for documents. It is needed for viewpoints.

Two observers may disagree because one is wrong. But they may also disagree because they are observing different aspects. Polytypic thinking does not resolve contradiction by declaring that everyone is right. It first asks **whether they are answering the same question.**

That small distinction can save entire research programmes.

The stereotype as useful laziness

It would be easy at this point to declare stereotypes the enemy. That would be an impressively stereotypical decision.

In a broad cognitive sense, a stereotype is often a cheap prior. It is useful in roughly the same way as a sketch on a napkin: the trouble begins when the napkin is laminated, stamped and declared the constitution of reality. If a configuration has been X a thousand times, beginning with X on the thousand-and-first encounter is rational. Without such compression neither humans nor machines could act in time. The problem is not the prior. The problem is the moment when a prior becomes a verdict.

“Things like this are usually X” is a reasonable beginning. “Therefore this particular case is X in every important respect” is a much stronger claim. “Anything that does not fit X is noise or error” is where a tiny statistical inquisition begins to form around our convenience.

Human stereotypes are especially dangerous when they concern people, because a type quickly becomes an expectation, the expectation becomes treatment, treatment becomes opportunity, and opportunity returns as “evidence” for the original type. A person is judged incapable, given fewer chances, produces a weaker result, and the system records triumphantly: “see, we were right.”

This is no longer merely a description of the world. It is a way of helping the world become more like the prediction. If historical inequality enters the data and we rename it an “objective signal,” automation has not purified the injustice. It has simply given it a faster interface.

Then we trained machines on humanity

Modern AI did not grow in a sterile laboratory outside culture. It learned from human texts, photographs, decisions, archives, rankings, linguistic habits and human histories of whom we noticed more often, whom we described as dangerous, whom we called a genius, whom we photographed in laboratories, whom we imagined as leaders and whom as servants.

Along with human knowledge, machines inherit human shortcuts.

The problem is no longer theoretical. In 2024, Valentin Hofmann, Pratyusha Ria Kalluri, Dan Jurafsky and Sharese King showed that language models can make significantly different social judgements about people on the basis of dialect features alone. In “AI generates covertly racist decisions about people based on their dialect”, models produced negative associations with African American English even without explicit mention of race. Particularly disturbing was the gap between overt and covert behaviour: newer models could produce more positive explicit answers while deeper associations remained strongly negative.

Another large analysis of generative language models and social identity found systematic ingroup/outgroup biases across many models. This is a useful warning: a safer tone does not guarantee that the system has stopped using inherited cultural coordinates. Sometimes a stereotype has merely learned to behave better in an interview.

But even that is not the most interesting part.

The stereotype comes back wearing an exoskeleton

When a neighbour at dinner says, “I think people like that usually...,” we hear the neighbour’s opinion. We know its provenance, can roughly estimate the experience behind it, argue, laugh, or ask them to stop talking nonsense.

When a complex AI system produces a similar conclusion in polished language, with a calm tone and the appearance of access to an enormous body of knowledge, the same pattern can be perceived very differently. It returns not as “someone’s opinion” but as **analysis.**

Here we encounter a phenomenon that MET[Ä]CADEMY can usefully call **epistemic laundering of a stereotype**. An old human template goes to the data centre, takes a shower, puts on a tie and comes back saying, “according to the analysis...” A human cultural prior enters data. A system learns it, sometimes amplifies it, synthesizes it and returns it to a human. At the output, an old human habit acquires a new provenance: “it is not me who thinks this; the AI saw it.”

In a series of experiments with 1,401 participants, Moshe Glickman and Tali Sharot showed that human–AI feedback loops can amplify human bias. Small distortions that entered an algorithm could be amplified; people interacting with the biased system gradually became more biased themselves; and they underestimated the degree to which the AI had influenced their judgements. The effect was demonstrated across perceptual, emotional and social tasks, including generative imagery.

That changes the scale of the problem. The most dangerous AI is not necessarily the one that invents new nonsense. Sometimes the more dangerous system is the one that returns our old nonsense fluently, calmly and with the air of having just checked the universe against a spreadsheet. A stereotype no longer simply passes from person to person. It can pass through a machine, acquire statistical polish, return to a person, then re-enter culture as a result apparently validated by machine interaction.

Humanity creates a prior. The machine learns from it. The machine returns it. Humanity sees confirmation. New decisions, texts and images become material for the next cycle.

This is no longer a mirror. It is a mirror with an amplifier.

When personalization becomes a cage

In the earlier MET[Ä]CADEMY Research Essay “The Interface That Knows You,” we asked a related question: what happens when a system becomes increasingly good at predicting a person? Personalization can be convenient, but knowing a person’s past behaviour is not the same as knowing their possible future. A system that perfectly knows what I clicked yesterday may quietly narrow what I will be allowed to discover tomorrow.

Polytypic thinking adds another layer. Any user profile is a typification: “likes this,” “not interested in that,” “belongs to this audience,” “likely to respond in this way,” “resembles users in this cluster.” For recommendation this may be useful. For a person it becomes dangerous when the system forgets that a profile is a history of observations, not a passport of essence.

Humans possess an inconvenient talent for ruining good predictive models: they can change their minds.

They can suddenly fall in love with music they switched off for ten years. Change profession. Leave a community. Stop drinking. Begin reading what “people of their type” usually do not read. Fall in love outside the recommendation. Forgive. Refuse to forgive. Grow up. Break. Recover. Invent another life.

If a system sees a person as a stable type, each such change first appears to it as prediction error. The ethical question begins one step later: does a system have the right to help a person remain predictable simply because predictability is convenient for the system?

And if that system controls a large part of what the person sees, prediction error can gradually become possibility error.

We may be building cultural autocomplete

Generative AI is entering not only recommendation but writing, music, design, education, programming, science, advertising, translation and the everyday formulation of thought. This creates enormous possibilities. But when one class of systems helps millions of people finish sentences, propose plots, structure presentations and find “good examples,” a new question appears: what happens not to the quality of one output, but to the diversity of culture as a whole?

In an experiment by Anil Doshi and Oliver Hauser, generative AI improved evaluations of individual short stories but made AI-assisted stories more similar to each other. It is a characteristic result: the individual user can receive a better work while the collective space becomes less diverse.

Later debate about brainstorming has revealed a similar conflict between the quality of an individual idea and the diversity of a pool of ideas. An analysis in *Nature Human Behaviour* directly raises the problem of reduced diversity even where average performance improves. In 2026, another study of more than 880,000 texts associated widespread LLM use as writing assistants with declining linguistic diversity and stronger dominance of common patterns.

None of this proves that AI is destined to make culture homogeneous. Other experiments suggest interaction design can preserve or even increase diversity. But the evidence destroys a comfortable illusion: measuring the quality of each individual output is not enough. A million local improvements can change the global landscape in a way that no single screen reveals.

We risk creating cultural autocompletes: every sentence good, every image professional, every presentation clean, every plot functional, while the world slowly begins to speak in the language of an extremely competent average. As if humanity finally prepared the perfect deck for a meeting at which nobody has anything left to say.

The tails of the distribution are part of humanity too

There is another reason not to treat typification as an innocent service. Statistical systems naturally see what is abundant more clearly than what is rare. The rare case has fewer examples. The new may have one by definition. A failed hypothesis may never enter a textbook. A small culture may be represented more weakly than a dominant one. A person with an unusual trajectory will always lose to the “typical user” in the number of available witnesses.

This is why the MET[Å]CADEMY Manifesto gives so much attention to abandoned paths, negative results and unresolved questions: “A negative result is a result. An abandoned path is knowledge. An unresolved question is inheritance.” This is not archive romanticism. It is protection against a world in which only what has already become large enough to be visible survives.

Technically, the danger even has a literal analogue. In work on model collapse, Ilia Shumailov and colleagues showed that when generative models are recursively trained carelessly on data generated by previous models, the tails of the distribution can disappear early: rare events, less probable variants, what lies far from the centre. That research concerns machine learning and synthetic data, not human culture directly, so transferring the conclusion one-to-one would be wrong. But the metaphor is too important to miss: **when a system repeatedly learns from its own projection of reality, it risks losing what that projection represented weakly.**

Now place that beside the human–AI loop.

What happens to cultural “tails” if people increasingly formulate thoughts through systems trained on previous corpora, while future systems increasingly learn from a world already partly shaped by previous systems?

This is not a prophecy of catastrophe. It is a reason to protect provenance, diversity and the right to remain atypical.

The most serious problem is the boundary of the new

Discussion of stereotypes in AI naturally begins with injustice. Who gets hired, who receives credit, whose language is treated as professional, whom a system imagines as criminal, doctor or secretary. This is critical.

But research systems face another danger that receives less attention: **stereotyping can destroy novelty before we have realized that we encountered something new.**

Imagine a system that has read almost everything humanity knows about a subject. It sees a new observation. Its great strength is finding related structures in the past. It quickly proposes the nearest known category, assembles supporting arguments and writes an excellent explanation. The researcher receives a brilliant answer.

But what if the most valuable part of the observation was precisely what did not fit the category?

The most powerful analogy machine may become the most efficient mechanism for premature closure.

The irony is nearly perfect. We build a machine capable of remembering an enormous portion of the past and then risk using that memory to grant the future permission to exist only in forms already known to the past.

The system becomes smarter while the horizon may become shorter.

“This is an error” sometimes means “where is the error?”

This becomes especially visible with strange, paradoxical or apparently impossible examples. A conventional system meets something that does not fit its picture and applies the label ERROR. Often that is correct. Data can be bad, instruments can fail, claims can contradict each other, and beautiful hypotheses are sometimes merely beautiful.

But the word “error” needs an address.

Is the error in the input data? The measuring instrument? Our interpretation? The scale? Did we mix a metaphor with a physical claim? Are two observers seeing different aspects? Did we transfer a rule from one world to another? Or does our set of categories simply have no place for the phenomenon in front of us?

The apparently impossible becomes a useful anti-stereotypical test here. Not because everything strange is true, but because the strange forces a system to reveal **where exactly it draws the boundary of the possible**.

If every unusual claim automatically becomes a new type, we get an intellectual carnival in which verification dies of exhaustion. If every unusual claim is automatically reduced to the nearest familiar thing, we get a sterile laboratory in which novelty is forbidden by fire-safety regulations.

Polytypic thinking is needed precisely between those extremes.

Polytypic does not mean “many labels”

The next mistake is easy. We can take an ordinary classifier, allow it to apply seventeen tags instead of one, and announce a glorious victory for pluralism.

No.

If a system has a fixed set of categories and simply selects more of them, it is still inside the same typology. That can be useful, but it does not address the deeper problem.

Polytypic thinking begins where we allow ourselves to ask: **is the set of types itself adequate?**

Yesterday’s categories may work today and fail tomorrow. A new object may require a new concept. A new relation may reveal that two old categories were partial views of one process. A change of scale may produce a property absent from the old grammar. A change of cultural context may make an old distinction crude or harmful.

In that sense polytypic thinking is not a catalogue of many types but **the right of a typology to be revised by the phenomenon**.

That is much more radical. And much more dangerous to beautiful tables.

Sometimes the boxes owe us an apology

The history of knowledge is full of classifications that once seemed natural. Some allowed science to make enormous progress. Some proved temporary. Others changed so radically that a contemporary reader barely recognizes the original question.

A planet is a type that humanity has revised together with its picture of the Solar System. Disease is a type reshaped through anatomy, microbiology, genetics and psychiatry. Species is a category around which biology still maintains several concepts useful for different tasks. Language and dialect cannot be separated by linguistics alone; history, power and identity live in the distinction. Even “art” repeatedly experiences border crises with craft, design, ritual, commerce, technology and memes.

Typologies are not a shameful weakness of science. Knowledge could not assemble without them. But a mature knowledge system should remember that a typology has a history.

Why would our present categories deserve immunity from the future?

Polytypicity as discipline, not indulgence

There is a temptation to read all this as soft relativism: everyone may call everything whatever they like, as long as we respect perspectives. That is the opposite of what is needed.

If type depends on time, scale, observer and context, then the demands on description become stricter, not weaker. We must state the conditions under which a claim works. We must not transfer evidence from one layer to another. We must not present metaphor as mechanism, correlation as cause, cultural image as genetic fact, or statistical prior as a property of one particular person.

Polytypicity does not reduce verification.

It reduces **the monopoly of one mode of verification over questions it never asked**.

This is why the MET[À]CADEMY Manifesto insists on Knowledge / Evidence Passports and provenance. Not for bureaucratic pleasure, but so that different pictures can coexist without disguising their nature. A scientific result can remain a strong scientific result. A myth can remain a strong cultural witness. Personal experience can remain honestly marked experience. A hypothesis can remain a hypothesis. The unknown does not have to impersonate an answer.

Polytypic thinking is not “everything is true.”

It is “do not erase the type of a claim as it crosses between worlds.”

What would an AI look like if it did not fossilize the past?

We do not need to design a new technical architecture in this essay. First we can change an intellectual habit.

A good research system should not be ashamed of simple answers when they are sufficient. A duck that looks, swims and quacks like a duck does not require the opening of a Department of Post-Waterfowl Ontology every morning. Simplicity is a virtue until it starts destroying distinctions.

But such a system should notice when its confidence comes not from the phenomenon but from the popularity of the category. It should treat the nearest known type as a first hypothesis, not a pre-trial detention cell. It should be able to preserve two incompatible projections long enough to learn whether one is actually wrong. It should remember that a rare variant may be an error, but may also be a new class, a weak signal, a regime shift or simply a person who is under no obligation to be statistically convenient.

And perhaps most valuable of all: it should sometimes be able to say not merely “I do not know the answer,” but **“it appears that our way of asking this question is insufficient.”**

That is another kind of uncertainty.

Not an empty cell in a table. A suspicion that the table is too small.

Why this becomes a question of time

If AI remained an occasional instrument used by a small group of specialists, the problem could be postponed. But AI is becoming an environment. Search, correspondence, education, translation, design, programming, documents, therapeutic conversations, bureaucracy, recruitment, medicine, scientific writing and political communication increasingly pass through models or beside them.

That means model typifications gradually cease to be private internal operations. They become part of the human environment.

If a system treats one writing style as “professional,” people begin to write that way. If it proposes certain plots more often, those plots appear in culture more often. If it rates one profile as a “promising candidate,” that profile receives more opportunities. If a recommender decides a person is “not interested” in a subject, the subject can disappear from their horizon. If millions of people use the same class of systems, the statistical centre of those systems gains unprecedented cultural weight.

At some point we must stop asking only: “does AI classify the world correctly?”

We must also ask: **“what kind of world becomes more probable because of the way AI classifies it?”**

That is no longer merely a question of accuracy. It is a question about the morphogenesis of culture.

The future may lose a vote to the past simply because it has fewer examples

The new is always statistically weak.

The first work of a new artistic movement loses to millions of works in established genres. The first strange hypothesis loses to the corpus of accepted explanations. The first person living in a way for which there is no name loses to every ready-made social type. The first technology that changes the structure of a task may initially look like a poor version of the old technology.

This was already a problem before AI. Now we are building systems whose extraordinary strength lies precisely in generalizing across enormous numbers of previous examples.

The central question for research AI may therefore be paradoxical: how can we use almost the whole memory of humanity without letting that memory vote against everything that does not yet exist inside it?

The MET[Å]CADEMY Manifesto begins: “Humanity knows more than it can remember together.” AI genuinely offers a chance to expand collective memory. But memory without the capacity to revise its own categories can become not a library of the future but an extraordinarily well-indexed past.

We do not need less memory. We need memory that does not confuse frequency with truth.

Polytypic thinking as the world's right to be more complex than our answer

Perhaps the shortest formulation is this: **type is a projection, not an essence.**

Not always. Some categories need rigid definitions. In law, engineering, safety, mathematics and standards, a hard boundary can save a system from chaos. Yet even there it helps to remember the difference between “we define this category this way for this task” and “reality itself consists of our administrative fields.”

Polytypic thinking does not wage war on boxes. The boxes themselves are innocent enough. The trouble begins when a box receives a budget, an algorithm and the authority to decide what happens to whoever does not fit inside it.

It leaves a label on the box: **“created for this question, under these conditions.”**

And it leaves the world the right occasionally not to fit.

This matters especially for MET[Å]CADEMY because its basic idea is not a new universal judge but a meeting place for different ways of knowing without forced loss of provenance. If AI inside such a field simply reproduces the most popular human typology, we get a very modern interface to a very old habit.

Something else is more interesting.

Not teaching a machine never to generalize. Teaching it to notice the boundary where generalization stops helping us see.

After the answer

When AI tells us, “this is X,” who once decided that X exists as a separate type at all? Who assembled the list of categories from which the machine now chooses so confidently, and which people, phenomena, cultures and ways of living were poorly represented among those who shaped that list? How many of our “natural” categories are old agreements that have simply lived long enough to forget their own history?

If the same process is chaotic up close and ordered from afar, where does its true type live? If an event is a cause for one observer and a trace for another, which one is wrong? If what is trauma today becomes memory tomorrow and a source of action the day after, at what moment are we entitled to name it finally? If a cultural form changes carrier, language, technology and symbols, what exactly must remain for us to speak of continuity?

If a person behaves “typically” for ten years and changes in the eleventh, will the system see a transformation or an error in its own prediction? How many possibilities do we already fail to see because an algorithm has become excellent at showing us what “people of our type” usually choose? How many scientific discoveries began as results that first looked like noise, and how many of them would a modern system explain away too quickly?

If millions of people ask AI what to write, draw, listen to, study and think about, does its statistical centre become a new cultural gravity? If future models learn from a world increasingly shaped by previous models, what happens to rare ways of speaking, seeing and imagining — who protects the tails? When a child twenty years from now asks an AI, “who am I?”, what exactly will the system answer: who the child is, who they were, who they resemble, or the average of people whom someone once managed to place into the correct boxes? And in such a world, who will defend the simple right of a person to be statistically inconvenient?

And if the most powerful intellectual instrument humanity has built learns to recognize flawlessly everything humanity has already seen, who will be first to notice what we have never seen before? What if the largest stereotype of the future is not a stereotype about a person, a culture or a machine, but our quiet conviction that **reality itself is obliged to fit the types we have already managed to invent for it?**

Sources and research anchors

This Research Essay develops principles of the MET[Å]CADEMY OF HUMANITY Manifesto concerning the absence of monopoly by any single way of seeing, the preservation of contradiction and provenance. For the history of categorization it draws on Rodney Needham, “Polythetic Classification: Convergence and Consequences,” *Man* 10(3), 1975, DOI: <https://doi.org/10.2307/2799807>; Michael Kleineberg, “Monothetic Classification and Polythetic Classification: A Cognitive-Developmental Perspective,” 2022, DOI: <https://doi.org/10.5771/9783956509568-159>; Eleanor Rosch & Carolyn Mervis, “Family resemblances: Studies in the internal structure of categories,” *Cognitive Psychology* 7(4), 1975, DOI: [https://doi.org/10.1016/0010-0285\(75\)90024-9](https://doi.org/10.1016/0010-0285(75)90024-9), and J. David Smith, “Prototypes, Exemplars, and the Natural History of Categorization,” *Psychonomic Bulletin & Review* 21, 2014: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3947400/>. On human–AI feedback loops: Moshe Glickman & Tali Sharot, “How human–AI feedback loops alter human perceptual, emotional and

social judgements,” *Nature Human Behaviour* 9, 2025: <https://www.nature.com/articles/s41562-024-02077-2>. On covert social biases in language models: Valentin Hofmann, Pratyusha Ria Kalluri, Dan Jurafsky & Sharese King, “AI generates covertly racist decisions about people based on their dialect,” *Nature* 633, 2024: <https://www.nature.com/articles/s41586-024-07856-5>; Tiancheng Hu et al., “Generative language models exhibit social identity biases,” *Nature Computational Science* 5, 2025: <https://www.nature.com/articles/s43588-024-00741-1>. On creativity and collective diversity: Anil R. Doshi & Oliver P. Hauser, “Generative AI enhances individual creativity but reduces the collective diversity of novel content,” *Science Advances* 10, 2024: <https://doi.org/10.1126/sciadv.adn5290>; Lennart Meincke, Gideon Nave & Christian Terwiesch, “Chat-GPT decreases idea diversity in brainstorming,” *Nature Human Behaviour* 9, 2025: <https://www.nature.com/articles/s41562-025-02173-x>. On linguistic diversity under LLM-assisted writing: “The shrinking landscape of linguistic diversity in the age of large language models,” *Nature Human Behaviour*, 2026: <https://www.nature.com/articles/s41562-026-02550-0>. On recursive training on synthetic data: Ilia Shumailov et al., “AI models collapse when trained on recursively generated data,” *Nature* 631, 2024: <https://www.nature.com/articles/s41586-024-07566-y>.

(MoH) · SUPPORT · 2026-08-25

SUPPORT THE FIELD · ПІДТРИМАТИ ПОЛЕ

Independent research, books, translations and public materials continue through voluntary support.

Незалежні дослідження, книги, переклади й відкриті матеріали продовжуються завдяки добровільній підтримці.

Monobank



UAH · Monobank Jar
send.monobank.ua/jar/2TMaAhchSZ

PayPal



International support
paypal.me/4IevgenKarogod

USDT · TRC20



TRON / TRC20 only
TDaCBB3T94CtMPRHn9pP8Uvdy8CTtWE5SV

d4ttara.github.io/metacademy-of-humanity/support/

Voluntary support creates no ownership, investment, membership or governance rights and does not present MoH as a registered charity. · Добровільна підтримка не створює прав власності, інвестиційних, членських або управлінських прав і не представляє MoH як зареєстровану благодійну організацію.