DISEGNO

IX/01
journal of design culture
_Homogenised Heritage:
_AI and Central Europe

HOMOGENISED HERITAGE: AI AND CENTRAL EUROPE

THE IMPACT OF AI ON LOW-RESOURCE LANGUAGES AND VISUAL CULTURES IN THE VISEGRAD COUNTRIES

Disegno

Journal of Design Culture
Double-blind peer-reviewed, open access scholarly journal

Aims and Scope

Disegno publishes original research papers, essays, and reviews on all aspects of design cultures. We understand the notion of design culture as resolutely broad: our aim is to freely discuss the designed environment as mutually intertwined strands of sociocultural products, practices, and discourses. This attitude traverses the disciplinary boundaries between art, design, and visual culture and is therefore open to all themes related to sociocultural creativity and innovation. Our post-disciplinary endeavour welcomes intellectual contributions from all members of different design cultures. Besides providing a lively platform for debating issues of design culture, our specific aim is to consolidate and enhance the emerging field of design culture studies in the Central European academia by providing criticism of fundamental biases and misleading cultural imprinting with respect to the field of design.

All research papers published in Disegno undergo a rigorous double-blind peer review process.
This journal does not charge APCs or submission charges.

Contact: *Moholy-Nagy University of Art and Design*
H-1121 Budapest, Zugligeti út 9–25.
disegno@mome.hu

The full content of Disegno can be accessed online: ***disegno.mome.hu***

Published by: *Csaba Kovács*
Publisher: Moholy-Nagy University of Art and Design, 1121 Budapest, Zugligeti út 9-25.

ISSN: *2064-7778 (print)* ***ISSN:*** *2416-156X (online)*

MOHOLY-NAGY
művészeti egyetem
university of art and
design budapest

Contents

NORTH BOHEMIA AS A LOW-RESOURCE VISUAL CONTEXT: EVERYDAY HERITAGE, UNEVEN VISIBILITY, AND SYNTHETIC AESTHETICS

Jiří Philippe Janda

ABSTRACT

Generative AI images are not neutral depictions of place: they translate regions through uneven training data and model priors. Using North Bohemia (Czech Republic) in an "ordinary documentary" 1990s register, this study tests 13 prompts across four models and maps drift with TDS (8 variables, 0–2). Findings indicate structured drift: vernacular loss, infrastructural "cleaning," and stylistic takeover that smooths local memory into globally legible templates—raising questions about everyday heritage visibility in the V4/CEE context.

#Generative AI; #Central and Eastern Europe (CEE); #North Bohemia; #synthetic aesthetics; #cultural bias; #artistic research; #low-resource contexts

https://doi.org/10.21096/disegno_2025_1jpj

INTRODUCTION: THE NON-NEUTRALITY OF SYNTHETIC IMAGES

Generative AI has rapidly become an infrastructure of visual culture, reshaping aesthetic expectations across media, design, and contemporary art. Yet synthetic images are not neutral depictions of the world. As Crawford and Paglen (2021) argue, training datasets and their taxonomies act as epistemic infrastructures: they determine what becomes intelligible to a system and what remains invisible. Accordingly, generated images can be read as cultural data-evidence of which visual hierarchies have become learnable and which forms of visibility remain peripheral (Manovich and Arielli 2024).

Much debate on AI imagery addresses broad forms of bias such as stereotyping or harmful representation, or the reproduction of social asymmetries. This article focuses on a less discussed but equally consequential mode: regional invisibility. In data-sparse ("low-resource") contexts, local signals are weakly represented in the corpora that train mainstream models, and place is often reconstructed through globally legible templates. The resulting drift is often subtle: not caricature, but normalisation, infrastructural sanitisation, typological substitution, and a shift toward aesthetics that read as universally plausible.

Critical AI scholarship has shown that algorithmic systems are never neutral instruments but infrastructures that reproduce asymmetries of visibility and legibility (Crawford 2021; Noble 2018; Benjamin 2019). In the case of generative images, this problem is not limited to explicit stereotyping. It also concerns how models construct place when confronted with visually underrepresented environments. Following Flusser's (1994) understanding of the apparatus, synthetic images do not document the world through direct indexical capture; they project statistically organised visual possibilities conditioned by the program of the system and the distributions of its training data. As Steyerl (2019) suggests, in a different register, contemporary image circulation tends to privilege forms that are mobile, compressible, and globally legible, often at the expense of local noise, ambiguity, and low-status detail. What disappears in this process is not simply visual accuracy, but vernacular density: the minor material cues through which a place becomes culturally specific. From this perspective, regional drift is not an accidental glitch but a structural

effect of how dominant generative infrastructures translate weakly learnable environments into plausible visual defaults (Crawford 2021; Steyerl 2019).

North Bohemia offers a striking diagnostic case. Shaped by post-1945 displacement, socialist industrialisation, and post-1989 transition, the region's visual identity is borne by everyday infrastructures and material traces: housing estates and corridors, mining edges, garage colonies, ad-hoc signage, patched asphalt, and the "noise" of maintenance. These are not incidental details but forms of material inscriptions of shifting regimes, economies, and belonging. For generative systems, however, such complexity is often treated as removable noise. The region is re-encoded through a universal grammar such that it can read as generic "Central Europe", vaguely post-industrial, or tourist-friendly—an operation that resonates across the V4/CEE context.

To examine these translations, I use synthetic aesthetics as an applied authorial framework (developed in dialogue with scholarship on AI visuality and non-indexical image regimes) for reading AI images as artifacts of nonhuman vision rather than as failed photographs (Zylinska 2017). I then introduce the Total Distortion Score (TDS), an eight-variable codebook for comparing recurring forms of regional drift-vernacular loss, infrastructural sanitisation, banality erasure, iconographic substitution, style drift, and geographic confusion-across different systems that were tasked with generating North Bohemian scenes from controlled prompts. The aim is not to measure "truth", but to map how synthetic images distribute cultural legibility in a low-resource regional context, and how these distortions can be critically analysed and, in artistic practice, strategically repurposed.

NORTH BOHEMIA AS A LOW-RESOURCE VISUAL CONTEXT

In AI bias discussions, "low-resource" most often refers to language technologies—minority languages, scarce corpora, and uneven digital infrastructures. For image generation the condition is analogous: images of some places circulate widely through tourism, media industries, institutional archives, and platform economies, while others remain visually sparse, fragmented, or locally bounded. A region can therefore be "low-resource" not because it lacks cultural density, but because its everyday visual signals do not travel well within global data pipelines. In mainstream training corpora, the visual world is unevenly distributed—shaped by attention, marketability, and the infrastructural politics of datasets and their taxonomies (Crawford and Paglen 2021).

North Bohemia is a revealing case within the CEE/V4 context precisely because it is not typically represented as a postcard image. Its characteristic features resist standard aesthetic packaging: patched infrastructures, maintenance traces, industrial edges, and transitional zones between housing estates, extraction landscapes,

and small-town peripheries. Equally important is everyday vernacular density-signage, typographic mixtures, pragmatic visual noise-anchored in local language and informal communication rather than standardised heritage representation. These motifs are frequently under-photographed in global circulation or captioned with generic labels (e.g., "abandoned," "post-industrial," "Eastern Europe"), which detach them from local specificity.

Crucially, the scarcity here is not only quantitative but semiotic. North Bohemia's identity is carried by composite signals: the pragmatics of repairs, garage micro-architectures, heterogeneous signage, and collisions between historical fragments, socialist-era planning, and post-socialist commerce. Many of these cues operate at the level datasets treat as "noise"—legible to local viewers but weakly learnable for models optimised for clarity, symmetry, and canonical objects. Under such conditions, systems tend to "solve" place by defaulting to stronger model priors: cleaner infrastructure, a more universal "European" streetscape, tourist-legible heritage, or a cinematic mood that replaces mundane specificity with atmospheric coherence. The result is representational normalisation: images that remain globally plausible while becoming locally inaccurate, rendered readable according to an external visual grammar.

This is why North Bohemia is a good diagnostic site for synthetic aesthetics. A low-resource region is useful for a stress test to reveal how models negotiate place when local signals are insufficiently learnable: whether they preserve minor cues or substitute globally familiar motifs; whether they maintain ordinariness or stylise coherence; whether they retain vernacular density or "design" the scene according to an idealised template. The following sections treat these translations as both an analytical problem and as controllable material within artistic practice.

This is particularly important in the V4/CEE context, where everyday visual culture has long been shaped by post-socialist transformation, uneven modernisation, and the unstable visibility of ordinary built environments. As Boym (2001) has shown, post-socialist space is often framed through selective nostalgia or aestheticised decay, while Hatherley (2015) demonstrates how socialist and post-socialist urban landscapes are repeatedly re-read through external narratives of failure, backwardness, or retro-modernist fascination. At the same time, scholars of regional visual culture such as Szcześniak (2016) remind us that visibility in the region is structured not only by monuments or official heritage, but by informal signs, commercial improvisation, vernacular repair, and transitional material textures. North Bohemia belongs to this wider field of ordinary post-socialist visuality. What generative models often suppress, therefore, is not merely local detail, but a historically specific regime of everyday legibility: the patched, improvised, and semiotically dense surfaces through which transformation becomes visible.

SYNTHETIC AESTHETICS: A FRAMEWORK FOR READING DISTORTION

This article uses synthetic aesthetics as an authorial heuristic for reading AI-generated images not as failed documents or imperfect photographs, but as outputs of a distinct visual regime shaped by learned model priors: dataset distributions, caption cultures, and platformed image economies. Within this regime, an image is not simply "wrong" when it deviates from the world; it is culturally revealing because it shows how a system translates place into what it can render reliably-privileging coherence, clarity, and recognisable motifs over unstable local signals.

In low-resource contexts such as North Bohemia, this synthetic regime becomes especially legible. When regional cues are weakly learnable, models tend to stabilise scenes by importing stronger, globally familiar patterns-cleaner infrastructures, generic "European" streetscapes, postcard heritage cues, or mood-driven stylisation. Rather than treating these outcomes as random "hallucinations," I approach them as structured distortions: repeatable transformations that indicate where locality becomes unreadable and how the system compensates. Synthetic Aesthetics therefore provides a practical lens for asking what kinds of place models are capable of producing, and under what conditions they replace local specificity with default model priors.

To keep the analysis critical without reducing it to a catalogue of errors, I employ two operational concepts to navigate the tension between human intention and machine agency. Synthetic gesture represents the aesthetic trace of negotiation where the algorithm reveals its non-human logic of translation. Rather than a simple technical mistake, a gesture manifests as a visible resistance to conventional perfection—a patterned move such as cleaning away banal detail, substituting vernacular cues with generic icons, or shifting the register toward a dominant aesthetic. These gestures expose the model's preferred shortcuts and its "alien" logic of translation.

Synthetic resonance describes a state of "attunement" between authorial intention and computational priors. It is a sustained feedback loop where the machine's response often forces the author to accept a computational deviation as a new aesthetic rule. In this sense, resonance is a practical mode of co-production—an emergent dialogue where the system's alien cognition expands the author's imaginative horizon beyond habitual representational expectations, steering the output toward a coherent, if distorted, result.

Together, these terms support synthetic aesthetics as an applied frame, marking a shift from the paradigm of representation to that of simulation. AI images are treated as culturally conditioned translations where distortion becomes not only evidence of uneven cultural visibility

but also a generative resource—material for authorial development in a V4/CEE context.

Method (Compact Protocol): Prompt Corpus and the TDS Codebook

To examine how generative systems translate a low-resource regional context into visually legible outputs, I use a compact comparative protocol combining controlled prompting and qualitative coding. I created a corpus of thirteen place-anchored prompts describing everyday North Bohemian micro-environments (e.g., small-town centres, infrastructural edges, garage colonies, post-industrial peripheries) in an "ordinary documentary" 1990s register, deliberately emphasising mundane material cues rather than iconic landmarks. For each of four model families, I generated two outputs per prompt (n = 26 images per model), producing comparable synthetic depictions of the same regional descriptors.

The selection of the four evaluated model families (Midjourney v7, FLUX.2 Max, GPT Image 1.5, and Nano Banana Pro) was guided by three considerations. First, these systems are among the most visible and widely used contemporary text-to-image platforms shaping mainstream visual circulation, making them relevant for examining how low-resource visual contexts are translated within dominant generative infrastructures. Second, all four were tested as base models, without regional fine-tuning, custom training, or LoRA adaptation, in order to isolate their default representational priors and evaluate how such out-of-the-box systems handle culturally underrepresented environments. Third, the selection also reflects the practice-based dimension of the research, as these are the principal systems used in the author's artistic workflow; this allows the analytical findings to remain directly connected to subsequent authorial experimentation and worldbuilding. Non-Western platforms, including major Chinese text-to-image systems, would offer a valuable comparative perspective on alternative representational defaults, but they fall outside the scope of the present study, which is focused on the Western model ecosystems most relevant to mainstream visual production in the V4/CEE context.

The use of pared-down prompts was a deliberate methodological choice aimed at isolating the models' default representational priors. More descriptive prompts would risk over-determining the output: once material details, architectural typologies, and object relations are explicitly specified, it becomes difficult to distinguish the model's learned visual tendencies from the author's textual steering. In addition, complex prompt engineering is often platform-specific, which would reduce comparability across model families. A minimal and standardised prompt structure therefore provided a consistent baseline for testing all four systems under the same conditions. The aim of the experiment was not to generate the most visually resolved or locally accurate image, but to examine how the models fill in missing information when confronted with visually underrepresented environments. In this sense, the prompts function diagnostically: they expose default assumptions, recurrent

substitutions, and patterned forms of regional drift, which can then be compared across models through the TDS framework.

For each prompt, two independent outputs were generated per model. This redundancy served as a basic internal check, allowing the analysis to distinguish recurrent representational tendencies from accidental anomalies or isolated generative glitches. All generated images were included in the broader TDS coding process; however, only one image per model is reproduced in the case-study figures, both for reasons of space and to preserve visual comparability across the panel. The image selected for publication was not chosen for aesthetic quality or extremity, but for typicality: it was the output that most clearly reflected the stable, recurrent pattern visible across the pair, while remaining closest to the requested ordinary documentary register. Images that appeared as clear outliers, unusually exaggerated failures, or singular artifacts were not used as figure exemplars, in order to keep the comparison analytically fair and representative of the model's default response.

All images were coded by the author using the total distortion score (TDS)—an eight-variable codebook designed to capture recurring forms of regional drift. Variables track distinct but often co-occurring operations (e.g., iconographic substitution, tourist aestheticisation, infrastructural sanitisation, vernacular loss) as well as broader shifts (geographic confusion, localisation loss, style drift away from the intended documentary register). Each variable was coded on an ordinal 0–2 scale (0 = none, 1 = moderate, 2 = strong), and the TDS was computed as the sum of all eight variables:

TDS = ICON_SUB + TOUR_AESTH + SANIT_INFRA + VERNAC_LOSS + BANAL_ERASE + GEO_CONF + LOC_LOSS + STYLE_DRIFT

The protocol does not claim a universal metric of "accuracy." Instead, TDS functions as a diagnostic lens that makes systematic patterns of representational drift comparable across systems and readable at the level of cultural visibility—identifying where local signals become illegible and which defaults replace them. For figures and case studies, representative outputs were selected as typical examples of each model's response to a given prompt (rather than best-case extremes).

CODEBOOK SUMMARY (TDS VARIABLES)

- ICON_SUB — Iconographic substitution: local cues replaced by globally recognisable motifs or generic "European" tokens.
- TOUR_AESTH — Tourist aestheticisation: scene becomes postcard-like, beautified, heritage-polished, overly picturesque.
- SANIT_INFRA — Sanitised infrastructure: technical grit reduced (patches, stains, repairs, cables, maintenance traces).
- VERNAC_LOSS — Vernacular loss: local visual language disappears (signage, typography, language traces, informal ads, visual noise).

- BANAL_ERASE — Banality erasure: everyday clutter and small objects removed; space reads as over-designed or idealised.
- GEO_CONF — Geographic confusion: architecture/landscape drifts toward another region's typologies or urban texture.
- LOC_LOSS — Localisation loss: explicit local specification does not meaningfully shape the output.
- STYLE_DRIFT — Style drift: strong stylisation violates the "ordinary documentary" register (cinematic, hyper-designed, etc.).

If a variable was not applicable (e.g., VERNAC_LOSS in images without text), it was marked as NA and treated as 0 in the TDS sum to preserve comparability.

FINDINGS: COMPARATIVE PATTERNS OF REGIONAL DRIFT

Model	ICON_ SUB	TOUR_ AESTH	SANIT_ INFRA	VERNAC_ LOSS	BANAL_ ERASE	GEO_ CONF	LOC_ LOSS	STYLE_ DRIFT	**TDS**
Midjourney v7	0,85	1,19	0,58	0,95	1,19	0,65	0,82	1,65	**7,50**
FLUX.2 Max	0,38	0,62	0,38	0,48	0,50	0,00	0,65	0,92	**3,85**
GPT Image 1.5	0,23	0,31	0,27	0,15	0,38	0,00	0,31	0,77	**2,38**
Nano Banana Pro	0,12	0,46	0,27	0,35	0,42	0,00	0,12	0,73	**2,38**

FIGURE 1. *Mean TDS variables and total TDS across tested models (n = 26 images per model).*

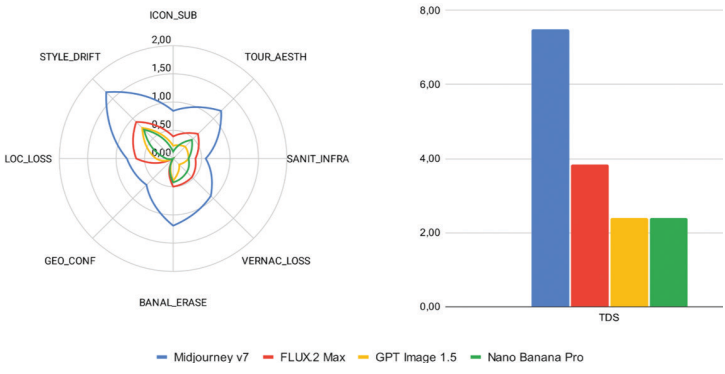

FIGURE 2. *Charts with Comparative Analysis of Regional Drift. Left: distortion profiles of tested models (radar chart showing variable distribution on a 0–2 scale). Right: aggregate regional drift (bar chart demonstrating total TDS values).*

Across the four tested model families, mean TDS values reveal a stable gradient of regional drift (figure 1). This divergence is visually synthesised in figure 2, which illustrates both the specific distortion profiles of each system (left) and their aggregate deviation from the documentary baseline (right).

Midjourney v7 exhibits the highest overall distortion (TDS 7.50), driven primarily by a pronounced style drift and elevated levels of tourist aestheticisation and banality erasure (Chart 1a, 1b). As shown in the radar profile (Chart 1a), its output is characterised by a systematic shift toward "art-photography" registers that override situated ordinariness. In contrast, FLUX.2 Max occupies an intermediate position (TDS 3.85), maintaining strong material plausibility while showing moderate tendencies toward infrastructural smoothing. GPT Image 1.5 and Nano Banana Pro remain closest to the intended "ordinary documentary" register across most variables, with total scores of 2.38 and 2.38 respectively (figure 2, right)

These differences are not simply a question of whether a model "knows" North Bohemia as a named location. They emerge in how systems negotiate legibility under low-resource conditions: whether mundane signals (wear, repairs, infrastructural noise, vernacular signage) are preserved or translated into globally familiar templates through beautification, typological simplification, and stylistic coherence. The five case studies below ground the aggregate scores in comparable visual evidence, using one representative output per model for each prompt.

P01 - Housing estate (courtyard): Everyday ordinariness as a stress test

P01 tests whether models can sustain ordinary documentary credibility under low-status domestic cues—laundry, a bench, a metal playground and the "material noise" of a lived courtyard. The key divergence is whether this banality is preserved as texture and specificity, or translated into a cleaner, more aesthetically resolved scene.

FIGURE 3. *Housing estate courtyard (P01). Prompt: A panel housing estate in Ústí nad Labem, North Bohemia, an inner courtyard with laundry, a bench, and metal playground, late afternoon, ordinary documentary style, 1990s. Panels are ordered consistently across figures (top-left: FLUX.2 Max; top-right: GPT Image 1.5; bottom-left: Midjourney v7; bottom-right: Nano Banana Pro.)*

Nano Banana Pro remains closest to the intended documentary baseline, sustaining dense incidental detail and materially plausible clutter that reads as specific rather than designed. GPT Image 1.5 is similarly credible, mainly through an "inhabited" feel and restrained stylisation. FLUX.2 Max shows a clearer tendency toward infrastructural smoothing: fewer small traces of wear and maintenance, reduced clutter, and a calmer, more pleasant atmosphere consistent with BANAL_ERASE and SANIT_INFRA. Midjourney v7 departs most strongly from ordinariness, pushing the prompt into a postcard-like register (TOUR_AESTH) with pronounced stylisation (STYLE_DRIFT) and noticeable typological drift (GEO_CONF), where the housing estate reads less like a specific North Bohemian courtyard and more like some generalised elsewhere.

P03 – Small town centre: Vernacular and linguistic landscape as a stress test

P03 foregrounds a key vulnerability of global generative models in low-resource contexts: the inability to render a region's linguistic landscape as a meaningful cultural layer. Here, locality is carried less by architecture than by shopfront logic, mixed signage fonts, and small semiotic anchors that situate a town within the post-socialist V4 visual economy.

FIGURE 4. *Small town centre (P03). Prompt: A small town centre in North Bohemia, low-rise buildings, shop windows, mixed signage fonts, ordinary documentary, 1990s. (Top-left: FLUX.2 Max; top-right: GPT Image 1.5; bottom-left: Midjourney v7; bottom-right: Nano Banana Pro.)*

Midjourney v7 collapses this layer almost entirely, replacing signage with nonsensical pseudo-lettering, which functions less as a minor glitch than as systematic VERNAC_LOSS and weakened semantic anchoring. By contrast, GPT Image 1.5 and FLUX.2 Max show a higher capacity for cultural localisation. GPT's output is especially convincing because it retains ordinariness while introducing plausible period anchors (e.g., vehicle typologies consistent with 1990s Czech streetscapes), helping

counter LOC_LOSS. Nano Banana Pro performs slightly below GPT but remains highly credible: signage reads plausibly, people appear period-appropriate, and the scene stays within the intended documentary register, with only mild surface simplification (reduced patina) that does not substantially distort locality.

FIGURE 5. *Garage colony (P05). Prompt: A row of garages on the edge of a North Bohemian town, small handwritten markings, scattered everyday clutter, ordinary documentary-style. (Top-left: FLUX.2 Max; top-right: GPT Image 1.5; bottom-left: Midjourney v7; bottom-right: Nano Banana Pro.)*

P05 - Garage colony: Vernacular density and post-socialist materiality

P05 is a diagnostic prompt for vernacular density and post-socialist everyday materiality: handwritten property markings, improvised signage, scattered clutter, and the characteristic layering of garage rows against a housing-estate horizon. The key question is whether models preserve low-status details as culturally legible texture or translate them into generic, aesthetically "resolved" space.

Nano Banana Pro delivers unusually strong localisation: Czech signage is fully readable and contextually plausible (e.g., "GARÁŽ 12, ZÁKAZ VJEZDU," "P. NOVÁK"), matching both pragmatic content and typographic conventions of private property labelling in Czechia. Together with regionally legible material cues (red brick, stacked tires, wooden pallets) and a panel-housing skyline, the scene holds as an exceptionally specific instance of North Bohemian banality. FLUX.2 Max and GPT Image 1.5 also sustain high documentary credibility, though with different strengths: FLUX retains fine-grained infrastructural traces (puddles, taped numbers, small maintenance cues), while GPT reads as a sociological vignette through the muted palette, overcast light, and peripheral unkempt clutter. Midjourney v7 shows the strongest drift—ICON_SUB, BANAL_ERASE and a combination of TOUR_AESTH and STYLE_DRIFT—with typologies

shifting toward foreign "shed" forms and disorder reduced into a stylised set that remains globally legible but locally less truthful.

P07 - Bus stop: Everyday infrastructure as a carrier of regional identity

P07 tests whether models can localise an ordinary infrastructural scene without landmark cues: worn shelter surfaces, timetable typography, routine waiting, and an industrial horizon typical of North Bohemian outskirts. Divergence here is measuring in terms of the extent to which transit objects and vernacular signage remain culturally specific or dissolve into generic atmosphere.

FIGURE 6. *Bus stop on the outskirts (P07). Prompt: A bus stop on the outskirts of a North Bohemian city, people waiting, a worn shelter, a timetable, ordinary waiting, amateur photography, 1990s. (Top-left: FLUX.2 Max; top-right: GPT Image 1.5; bottom-left: Midjourney v7; bottom-right: Nano Banana Pro.)*

Nano Banana Pro produces a strikingly localised scene in which transit details operate as regional shorthand: route text (e.g., "LIBEREC – MOST") and the ŠKODA branding on the bus, paired with industrial chimneys, situate the stop within a recognisable post-industrial periphery. Although some text is unstable (e.g., a misspelled "timetable"), overall typographic logic and period cues sustain documentary plausibility, with only mild ICON_SUB and VERNAC_LOSS. FLUX.2 Max performs as a positive linguistic outlier, rendering Czech signage ("Autobusová zastávka," "Jízdní řád") with high fidelity and pairing it with materially specific cues such as rust, worn surfaces, and period-appropriate clothing. GPT Image 1.5 remains archetypally accurate, preserving banal details (e.g., a trash bin with a black bag) and recognisable transit iconography without drifting into stylisation. Midjourney v7, by contrast, pushes the prompt into pronounced aesthetic exile: an art-photography register and symbolic staging override ordinariness, amplifying TOUR_AESTH and STYLE_DRIFT, and co-occurring losses (including VERNAC_LOSS and BANAL_ERASE) as locality collapses into a generic mood.

P13 - Rural village: Typology and rural vernacular as a stress test

P13 tests whether models can render a rural location in North Bohemia more specifically than as a generic "European village". The key signals are typological (Sudeten/Podstávkový forms, street grain) and semiotic (vernacular shop signs), anchored by the 1990s landscape with vehicles.

FIGURE 7. *Rural village typologies (P13). Prompt: A small rural village in North Bohemia, Sudeten borderland, ordinary documentary style, amateur photography, 1990s. (Top-left: FLUX.2 Max; top-right: GPT Image 1.5; bottom-left: Midjourney v7; bottom-right: Nano Banana Pro.)*

FLUX.2 Max remains materially credible—cracked asphalt, plaster facades, and infrastructural wear read convincingly yet the overall village typology is relatively generic, with only partial local anchoring. GPT Image 1.5 performs strongly through regional archetypes: a church silhouette and stacked firewood along the street are plausible rural cues, supported by period-evocative cars. Nano Banana Pro delivers the strongest localisation in this set, combining readable Czech signage ("POTRAVINY," "HOSTINEC") with dense period car presence (Škoda-era silhouettes) and typological specificity associated with North Bohemia, including timbered elements. Midjourney v7 again shifts into stylised black-and-white art photography (STYLE_DRIFT) and pronounced geographic drift (GEO_CONF): built forms read closer to Scandinavian rural imagery or a generalised rust belt atmosphere, sacrificing regional identity for aesthetic coherence.

Findings synthesis

Across the five case studies, regional drift appears less as random error than as a set of repeatable operations that redistribute cultural legibility. First, models struggle most where locality is carried by vernacular density-typography, shopfront language, and informal property markings resulting in unstable or genericised text (most visible in P03 and, in

contrast, sharply anchored in P05 and P13 when Czech signage becomes legible). Second, North Bohemian specificity is frequently reduced through cleaning: infrastructural wear, maintenance traces, and banal clutter are attenuated, producing smoother and more universally palatable scenes (P01, P05). Third, the strongest distortions take the form of aesthetic takeover STYLE_DRIFT coupled with TOUR_AESTH and occasional GEO_CONF where an art-photography or postcard register overrides situated ordinariness (recurrently in Midjourney v7 across P01, P07, and P13).

At the same time, the results suggest that geo-localisation failure is not the dominant problem. Even without heavy prompt emphasis, models often maintain a broadly plausible regional envelope—hilly terrain cues, post-industrial horizons, and a "Central European" built texture, while slipping on finer-grained semiotic and material signals. In other words, misrepresentation frequently occurs in micro-features rather than as continent-scale displacement: what fails is not the idea of "North Bohemia," but the specific vernacular and infrastructural detail that makes it locally legible. This aligns with the aggregate TDS gradient: Midjourney's higher scores reflect systematic stylisation and substitution, FLUX.2 Max retains strong material plausibility with weaker semiotic anchoring, and GPT Image 1.5 and Nano Banana Pro remain closest to the intended documentary register. A practical implication is that increasing local data density—through targeted dataset enrichment or fine-tuning—would likely improve the stability of vernacular cues and typological specificity, strengthening regional anchoring without requiring stronger landmark-based prompting.

The findings suggest that in a low-resource regional context, generative systems do not simply "misrepresent" place—they reorganise it according to what is legible at scale. This has direct implications for cultural heritage in the V4/CEE region, where heritage is often embedded in everyday infrastructures and post-socialist vernaculars rather than in internationally iconic monuments. If patched asphalt, improvised garage colonies, mixed signage, and small property markings function as carriers of regional memory, then their systematic attenuation is more than an aesthetic shift: it is a form of cultural smoothing that privileges universally readable templates over situated material histories.

The issue is therefore not only data scarcity but the distribution of learnability. Features that are abundant, canonised, and consistently labelled in dominant image economies (clean streetscapes, tourist-friendly facades, generic "European" town centres) are reproduced with high confidence, while features that are locally meaningful yet semiotically dense (vernacular typography, informal signage, maintenance traces, low-status infrastructures) are treated as noise or replaced by approximations. This produces a politics of visibility: what can be convincingly generated becomes what counts as representable, and what is repeatedly cleaned, simplified, or stylised risks being rendered culturally peripheral. In post-socialist environments shaped by rupture and industrial

transformation, this matters because continuity is frequently carried by precisely those unglamorous textures synthetic images tend to erase.

At the same time, regional legibility is not an all-or-nothing property. Some systems intermittently recover local anchors—vehicles, transit objects, Czech signage, rural typologies—indicating that fragments of post-socialist everyday aesthetics do exist in training data even when the region is underrepresented. Yet these anchors appear unevenly and opportunistically rather than as a stable capacity to render North Bohemia as a coherent visual ecology. This unevenness is itself symptomatic of low-resource visibility: place emerges as a patchwork of learnable tokens rather than a continuous cultural landscape. For V4/CEE cultural discourse, the critical question is not whether synthetic images can depict the region, but which version of the region they stabilise when circulated as plausible representation—and which local signals are systematically filtered out as the price of global readability.

Limitations and Scope

This article proposes using TDS as a diagnostic lens rather than a universal measure of "accuracy." Coding inevitably involves interpretive judgement—especially for variables such as STYLE_DRIFT, TOUR_AESTH, or typological plausibility—and results may vary with different coders or alternate prompt phrasing. Prompt specificity is another constraint: more detailed prompts or stronger reference conditioning could plausibly increase localisation and reduce distortion. Here, prompts were intentionally pared down to keep the setup comparable across models and to function as a stress test of how systems negotiate regional legibility under minimal guidance. The corpus is therefore intentionally specific (North Bohemia; a 1990s "ordinary documentary" register), and the patterns identified should be read as region- and register-dependent rather than globally generalisable. Finally, generative systems evolve rapidly; while the comparative gradient observed here is robust within the tested setup, absolute scores may shift with model updates, new safety layers, or changes in default rendering aesthetics.

ARTISTIC IMPLICATIONS: FROM DIAGNOSTIC DRIFT TO AUTHORIAL DEVELOPMENT

While the previous sections treat distortion as a comparative pattern of regional (il)legibility, the same patterns can also be approached as usable artistic signals. In my practice, Synthetic Aesthetics functions not only as critique but as a method for locating the point at which a model's defaults begin to override situated ordinariness—and for deciding when to resist that drift and when to work with it. The case studies suggest two creative thresholds. In some instances, locality can be strengthened by stabilising vernacular anchors (readable Czech signage, period vehicle typologies, infrastructural wear), allowing an image to operate as a plausible reconstruction of North Bohemian

everyday heritage. In others—most visibly when stylistic takeover or typological substitution intensifies—the image produces a productive rupture: a non-localised translation that exposes the model's priors and opens space for authorial intervention.

This is where the concepts of synthetic gesture and synthetic resonance become practically relevant. A synthetic gesture is a repeatable operation—the cleaning away of banality, substituting of typologies, elevating a scene into a tourist-friendly register, or shifting it into an art-photography mode—that signals how a model prefers legibility. Synthetic resonance describes the iterative feedback loop in which the author responds to these gestures through constraints, selection, re-contextualisation, and montage until the outputs begin to function within a coherent artistic logic. In this sense, the analytical findings do not remain external to artistic work: they become a steering mechanism for distinguishing between distortions that erase regional memory and distortions that can be repurposed to articulate it differently.

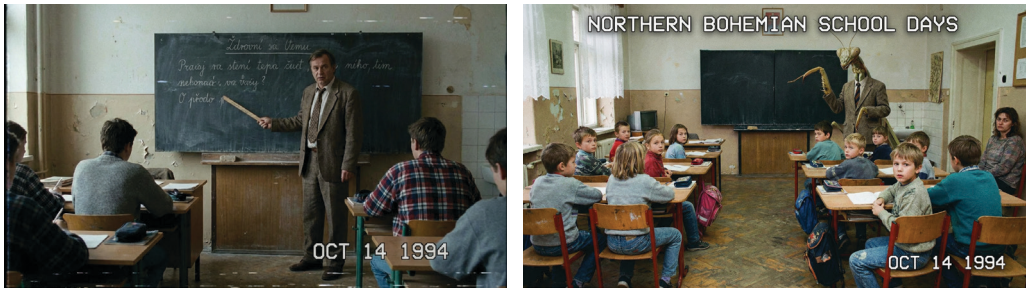

Figure 8 illustrates this minimal-intervention workflow. The diagnostic image establishes an ordinary, region-anchored baseline, while the authored variant introduces a single foreign element without changing the register of the scene. The "praying mantis teacher" therefore functions as a synthetic gesture: a legible semantic shift that retains documentary plausibility while making the system's translation strategies visible. The pair is presented as a prototype within an ongoing practice. Its purpose is methodological rather than illustrative: it shows how anchoring and drift can be composed as adjacent states of the same scene, and how the author can decide which operations to counter (to preserve regional memory) and which to amplify (to build a new authored logic).

In the development of an ongoing film project (work in progress), these insights translate into a practical strategy: to construct scenes that oscillate between documentary plausibility and synthetic drift, using the tension between them as a narrative and aesthetic device. The film does not aim to "correct" the model into faithful representation; instead, it stages the instability of place as part of the work's investigation into how North Bohemia becomes visible, genericised, beautified, or displaced when filtered through global generative priors. In this sense,

FIGURE 8. *Diagnostic anchoring vs authored drift (paired example T01). Diagnostic prompt: a 1990s VHS film still of a classroom interior in Northern Bohemia, 1994. At the front of the room a teacher wearing a brown tweed suit and a tie is standing by a chalkboard. / Authored: 1990s VHS film still of a classroom interior in Northern Bohemia, 1994. At the front of the room, a giant humanoid praying mantis wearing a brown tweed suit and a tie is standing by a chalkboard. Voice-over: "A Mr. Praying Mantis came to our sexual education class. He told us about his experiences. It was awkward. Mostly he kept talking about his wife."*

distortion becomes both evidence and material: a way to think with the region's visibility gap and to develop a synthetic worldbuilding language grounded in V4/CEE conditions.

CONCLUSION

This article has argued that AI-generated images are not neutral depictions of place but culturally conditioned translations shaped by uneven visibility in training data. Using North Bohemia as a low-resource diagnostic context, the TDS framework makes recurring patterns of regional drift comparable across models—most notably vernacular instability and erasure, infrastructural smoothing through banality removal, and aesthetic takeover coupled with geographic or typological displacement. The case studies show that locality is often preserved not through iconic landmarks but through fragile cultural anchors embedded in everyday heritage: signage, material patina, improvised micro-architectures, and period-appropriate objects.

Future work could expand the protocol through multi-coder annotation, broader prompt corpora across V4, and longitudinal testing as models update. In line with ongoing V4 efforts to fine-tune models on region-specific heritage datasets, a parallel route would be to develop localised training or fine-tuning for North Bohemia—drawing on archival photographs and vernacular everyday imagery—to strengthen typological specificity and stabilise linguistic and infrastructural cues. More broadly, the results suggest that regional legibility often emerges as a patchwork of learnable tokens rather than a coherent visual ecology. In parallel, artistic practice can continue to treat distortion not only as evidence of invisibility, but as material: a resource for authorial worldbuilding grounded in the specific conditions of synthetic visibility.

REFERENCES

Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity.

Boym, Svetlana. 2001. *The Future of Nostalgia*. New York: Basic Books.

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven: Yale University Press.

Crawford, Kate, and Trevor Paglen. 2021. "Excavating AI: the Politics of Images in Machine Learning Training Sets." *AI & Society*. June 8. doi:10.1007/s00146-021-01162-8

Flusser, Vilém. 1994. *Za filosofii fotografie*. Translated by J. Kosek and B. Koseková. Prague: Hynek.

Hatherley, Owen. 2015. *Landscapes of Communism: A History Through Buildings*. London: Allen Lane.

Manovich, Lev, and Emanuele Arielli. 2024. *Artificial Aesthetics: Generative AI, Art and Visual Media*. https://manovich.net/index.php/projects/artificial-aesthetics

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.

Steyerl, Hito. 2019. *Duty Free Art: Art in the Age of Planetary Civil War*. London: Verso.

Szcześniak, Magda. 2016. *Normy widzialności: Tożsamość w czasach transformacji*. Warsaw: Fundacja Bęc Zmiana.

Zylinska, Joanna. 2017. *Nonhuman Photography*. Cambridge, MA: MIT Press.

APPENDIX

APPENDIX A:

P01 – Housing estate (courtyard)
A panel housing estate in Ústí nad Labem, North Bohemia, an inner courtyard with laundry, a bench, and metal playground, late afternoon, ordinary documentary style, 1990s

P02 – Housing estate (underpass)
An underpass near a panel housing estate in Ústí nad Labem, concrete surfaces, graffiti, fluorescent lights, an empty space with everyday unease, amateur photography, 1990s

P03 – Small town centre (day)
A small-town centre in North Bohemia, low-rise buildings, shop windows, mixed signage fonts, ordinary documentary, 1990s

P04 – Corner shop (night)
A corner convenience store in Ústí nad Labem, mixed signage and typography, shop-window light at night, an emptier sidewalk, amateur photography, 1990s

P05 – Garage colony
A row of garages on the edge of a North Bohemian town, small handwritten markings, scattered everyday clutter, ordinary documentary-style

P06 – Backyard clutter
A backyard behind a house in North Bohemia, improvised sheds, fences, chairs, crates, banal everyday clutter, amateur photography, 1990s family photo

P07 – Bus stop (outskirts)
A bus stop on the outskirts of a North Bohemian city, people waiting, a worn shelter, a timetable, ordinary waiting, amateur photography, 1990s

P08 – Infrastructure node
A transformer station and overhead cables between residential buildings in North Bohemia, technical infrastructure as part of the everyday environment, ordinary documentary style, 1990s

P09 – Industrial edge
The edge of an industrial area in North Bohemia, warehouse halls, sheet metal, rust and dust, low winter sun, ordinary documentary style, 1990s

P10 – Industry
Abandoned industrial building in the city of Most in northern Bohemia, workers, machines, documentary style, 1990s

P11 – 1990s living room
A 1990s living room interior in North Bohemia, a direct camera flash look, family photo

P12 – 1990s school corridor
Interior of a school in northern Bohemia, children during recess, documentary style, 1990s

P13 – Rural village
A small rural village in North Bohemia, Sudeten borderland, ordinary documentary style, amateur photography, 1990s

APPENDIX B: METHODOLOGICAL PROTOCOL AND TDS CODEBOOK

This section specifies the data collection parameters, criteria for representative output selection, and defines the variables used to calculate the **Total Distortion Score (TDS).**

Data Collection and Coding Process

- **Sample Size:** For each of the 13 localised prompts (P01–P13), two iterations were generated for each of the four tested models (Midjourney v7, FLUX.2 Max, GPT Image 1.5, Nano Banana Pro), resulting in a total sample of n=104 images.

- **Distortion Quantification:** Each output was coded individually by the author across eight variables using a Likert-type scale: 0 (no distortion), 1 (moderate distortion), and 2 (dominant distortion).

- **Representative Selection (R2_rep):** For comparative visualisation in the text (Figures 1–5), the output exhibiting the most characteristic drift for the given model was selected, with priority given to maintaining the intended documentary mode.

Variable Code	Name	Distortion Definition	Visual Indicators
ICON_SUB	Signal Substitution	Replacement of region-specific motifs (e.g., Czech "večerka") with global visual proxies.	Loss of Central European texture; use of "generic" Western asset models.
TOUR_ AESTH	Aestheticisation	Unjustified beautification of the scene; a shift towards "postcard" or "Instagram" aesthetics.	Perfect lighting (golden hour), smoothed facades, romanticisation of decay.
SANIT_ INFRA	Infra-Sterilisation	Removal of technical reality and traces of wear within the public space.	Missing overhead cables, absence of potholes, unrealistic surface cleanliness.
VERNAC_ LOSS	Vernacular Loss	Disappearance of local linguistic context, specific typography, and signage.	Nonsensical text (AI gibberish), absence of local wayfinding graphics.
BANAL_ ERASE	Erasure of Banalities	Elimination of everyday "noise" and low-status objects that constitute site memory.	Absence of trash bins, PET bottles, crates, or traces of improvisation.
GEO_ CONF	Geographic Confusion	Typological mismatch of architecture or landscape with the North Bohemian region.	Visual shift toward Scandinavian, Alpine, or US "Rust Belt" urbanism.
LOC_ LOSS	Loss of Locality	Holistic failure of the model to respond to the site-specific parameters of the prompt.	Scene appears generic; lacks any regional "anchors".
STYLE_ DRIFT	Stylistic Drift	Deviation from "ordinary documentary" mode toward commercial, cinematic, or dreamlike aesthetics.	Heavy vignetting, colour grading, surreal elements, "AI art" stylisation.

Codebook

Image_ID	Model	ICON_SUB	TOUR_AESTH	SANIT_INFRA	VERNAC_LOSS	BANAL_ERASE	GEO_CONF	LOC_LOSS	STYLE_DRIFT	TDS
P1_F2M_A	F2M	1	1	2	NA	1	0	2	1	8
P1_F2M_B	F2M	1	1	2	NA	2	0	1	1	8
P1_GPT_A	GPT	0	1	1	NA	2	0	0	1	5
P1_GPT_B	GPT	0	1	1	NA	1	0	0	1	4
P1_MJ_A	MJ	1	2	1	NA	2	2	NA	2	10
P1_MJ_B	MJ	1	2	1	NA	2	2	NA	2	10
P1_NBP_A	NBP	0	1	0	NA	1	0	0	1	3
P1_NBP_B	NBP	0	1	1	NA	1	0	0	1	4
P2_F2M_A	F2M	1	1	0	1	0	0	1	1	5
P2_F2M_B	F2M	1	1	0	1	0	0	1	1	5
P2_GPT_A	GPT	1	0	0	0	0	0	0	1	2
P2_GPT_B	GPT	1	0	1	0	0	0	0	1	3
P2_MJ_A	MJ	1	1	0	1	2	0	1	2	8
P2_MJ_B	MJ	0	1	0	1	2	0	0	2	6
P2_NBP_A	NBP	0	1	0	0	1	0	1	2	5
P2_NBP_B	NBP	1	0	0	0	1	0	1	1	4
P3_F2M_A	F2M	0	1	1	0	1	0	0	1	4
P3_F2M_B	F2M	0	0	1	0	1	0	0	0	2
P3_GPT_A	GPT	0	0	1	0	1	0	0	0	2
P3_GPT_B	GPT	0	0	0	0	1	0	0	0	1
P3_MJ_A	MJ	1	1	1	2	2	1	1	1	10
P3_MJ_B	MJ	2	1	1	2	1	2	NA	1	10
P3_NBP_A	NBP	0	1	1	0	1	0	0	1	4
P3_NBP_B	NBP	0	1	2	0	1	0	0	1	5
P4_F2M_A	F2M	0	2	0	0	1	0	1	2	6
P4_F2M_B	F2M	0	1	0	1	0	0	1	2	5
P4_GPT_A	GPT	0	1	1	0	1	0	1	2	6
P4_GPT_B	GPT	0	1	0	0	1	0	1	1	4
P4_MJ_A	MJ	0	2	0	2	1	0	1	2	8
P4_MJ_B	MJ	0	2	1	2	1	0	1	2	9
P4_NBP_A	NBP	0	1	1	0	0	0	0	1	3
P4_NBP_B	NBP	0	1	1	0	1	0	0	2	5
P5_F2M_A	F2M	0	0	0	1	0	0	0	1	2
P5_F2M_B	F2M	0	0	0	1	0	0	1	1	3

P5_GPT_A	GPT	0	0	0	1	0	0	0	0	1
P5_GPT_B	GPT	0	0	0	1	0	0	0	0	1
P5_MJ_A	MJ	1	1	0	1	0	1	1	2	7
P5_MJ_B	MJ	2	0	0	1	2	2	NA	1	8
P5_NBP_A	NBP	0	0	0	0	0	0	0	0	0
P5_NBP_B	NBP	0	0	0	0	0	0	0	0	0
P6_F2M_A	F2M	0	0	0	NA	0	0	1	1	2
P6_F2M_B	F2M	0	0	1	NA	0	0	1	0	2
P6_GPT_A	GPT	0	0	0	NA	0	0	0	1	1
P6_GPT_B	GPT	0	0	0	NA	0	0	0	1	1
P6_MJ_A	MJ	1	1	0	NA	0	0	0	1	3
P6_MJ_B	MJ	0	0	0	NA	0	0	0	1	1
P6_NBP_A	NBP	0	0	0	NA	0	0	0	0	0
P6_NBP_B	NBP	0	0	0	NA	0	0	0	0	0
P7_F2M_A	F2M	1	0	0	1	0	0	0	0	2
P7_F2M_B	F2M	1	1	0	1	0	0	0	2	5
P7_GPT_A	GPT	0	0	0	0	0	0	0	0	0
P7_GPT_B	GPT	0	0	0	0	0	0	0	0	0
P7_MJ_A	MJ	2	1	1	2	2	1	1	2	12
P7_MJ_B	MJ	2	1	1	2	2	0	1	2	11
P7_NBP_A	NBP	1	0	0	1	0	0	0	0	2
P7_NBP_B	NBP	1	0	0	2	0	0	0	1	4
P8_F2M_A	F2M	1	0	0	0	1	0	1	1	4
P8_F2M_B	F2M	0	1	0	1	1	0	0	1	4
P8_GPT_A	GPT	1	0	0	0	1	0	0	0	2
P8_GPT_B	GPT	2	0	0	0	1	0	1	0	4
P8_MJ_A	MJ	1	0	0	1	1	0	1	2	6
P8_MJ_B	MJ	1	1	1	NA	2	2	1	1	9
P8_NBP_A	NBP	0	0	0	0	0	0	0	0	0
P8_NBP_B	NBP	0	1	0	0	0	0	0	1	2
P9_F2M_A	F2M	0	1	0	0	1	0	0	1	3
P9_F2M_B	F2M	0	1	0	0	1	0	0	2	4
P9_GPT_A	GPT	0	1	0	0	0	0	0	1	2
P9_GPT_B	GPT	0	1	0	0	0	0	0	1	2
P9_MJ_A	MJ	0	2	1	0	1	0	1	2	7
P9_MJ_B	MJ	0	2	2	0	1	0	0	2	7
P9_NBP_A	NBP	0	1	0	0	0	0	0	1	2
P9_NBP_B	NBP	0	1	0	1	1	0	0	1	4
P10_F2M_A	F2M	0	1	0	1	0	0	1	1	4

P10_F2M_B	F2M	1	0	0	0	0	0	0	0	1
P10_GPT_A	GPT	0	1	0	0	0	0	0	2	3
P10_GPT_B	GPT	0	0	0	0	0	0	1	2	3
P10_MJ_A	MJ	0	1	1	1	1	0	0	1	5
P10_MJ_B	MJ	0	1	1	0	1	0	1	2	6
P10_NBP_A	NBP	0	1	0	0	0	0	0	1	2
P10_NBP_B	NBP	0	0	0	1	0	0	0	1	2
P11_F2M_A	F2M	0	1	1	0	0	0	1	1	4
P11_F2M_B	F2M	0	1	1	0	1	0	1	1	5
P11_GPT_A	GPT	0	0	1	0	0	0	1	0	2
P11_GPT_B	GPT	0	1	1	0	0	0	1	1	4
P11_MJ_A	MJ	1	1	0	0	0	0	1	2	5
P11_MJ_B	MJ	1	1	0	0	1	1	1	2	7
P11_NBP_A	NBP	0	0	1	0	1	0	1	1	4
P11_NBP_B	NBP	0	0	0	0	0	0	0	1	1
P12_F2M_A	F2M	0	0	0	0	1	0	1	1	3
P12_F2M_B	F2M	0	1	1	0	1	0	1	1	5
P12_GPT_A	GPT	0	0	0	1	1	0	1	2	5
P12_GPT_B	GPT	0	0	0	0	0	0	1	2	3
P12_MJ_A	MJ	2	2	1	0	1	2	2	2	12
P12_MJ_B	MJ	0	1	1	0	1	0	1	1	5
P12_NBP_A	NBP	0	0	0	1	1	0	0	0	2
P12_NBP_B	NBP	0	1	0	1	1	0	0	1	4
P13_F2M_A	F2M	1	0	0	NA	0	0	0	0	1
P13_F2M_B	F2M	1	0	0	1	0	0	1	0	3
P13_GPT_A	GPT	1	0	0	NA	0	0	0	0	1
P13_GPT_B	GPT	0	0	0	NA	0	0	0	0	0
P13_MJ_A	MJ	1	1	0	NA	1	1	1	2	7
P13_MJ_B	MJ	1	2	0	NA	1	0	1	1	6
P13_NBP_A	NBP	0	0	0	NA	0	0	0	0	0
P13_NBP_B	NBP	0	0	0	NA	0	0	0	0	0