Open3DCP: A Public Data Schema for 3D Concrete Printing

Design, rationale, and the measurement gaps of an analysis-ready record for extrusion 3D concrete printing

Nicholas Sonnentag — Sunnyday Technologies, Appleton, WI, USA

2026-06-08

Contents

Abstract

Three-dimensional concrete printing (3DCP) is a coupled material–process problem: a printable cementitious system is defined by composition, rheology, equipment, toolpath, environmental exposure, curing, specimen extraction, test orientation, and measured performance. The literature reports many of these facts, but distributes them across prose, tables, figures, and supplementary files using inconsistent names and unit bases — which makes cross-study comparison, meta-analysis, and machine-learning workflows harder than the underlying measurements require. Open3DCP is an open, flat schema for recording extrusion-based 3DCP mix-design and test records. Version 1.7 defines **244 columns** in its primary record, spanning composition, fibers, admixtures, fresh-state rheology, 3DCP process parameters, hardened mechanical and durability properties, interlayer bond, specimen/test metadata, environmental conditions, and provenance.

Quantities are recorded on a **kg/m³-primary basis** (the field standard), with mass-percent of total wet mix retained as a losslessly-derived secondary representation; missingness is represented as `NULL`, with `0` reserved for explicit source-reported zero. Open3DCP is a *reporting layer*, not a test method or a substitute for ASTM, EN, ACI, ICC, ISO, or RILEM standards: it lets 3DCP records be assembled into interoperable datasets for scientific review, Integrated Computational Materials Engineering (ICME)–style process–structure–property analysis, digital-twin reconstruction, and downstream modeling where sufficient validated data exist. We give the design rationale for each major decision and **demonstrate ingestion** on two openly-licensed public datasets — a cast-concrete benchmark and a real ~30-laboratory 3DCP database — into one schema, including a cross-study print-anisotropy result. We argue — analytically, fitting no predictive model — that classical single-predictor strength models cannot represent 3DCP's process- and orientation-dependence, motivating multi-feature models and the comprehensive record that supports them. We also identify a taxonomy of features that are physically important but cannot yet be measured reliably — an agenda for instrumentation.

Contributions

- An open, **3DCP-native flat data schema** (v1.7, 244 columns) with first-class columns for print process parameters, fresh-state rheology, and interlayer-bond properties — the features that distinguish printed concrete from cast.
- A **design rationale** for each consequential choice: flat table over graph, kg/m³-primary *dual-basis* recording, fineness-modulus over maximum aggregate size, and `NULL` ≠ `0`.
- A **measurement-gap taxonomy** — real-time process monitoring, in-situ material state, and post-process characterization gaps — framed as a research agenda for instrumentation.
- An argument that **classical single-predictor strength models** (Abrams, Bolomey, Féret, Powers) cannot represent 3DCP's process- and orientation-dependence, motivating multi-feature models and the comprehensive record that supports them. This is an analytical argument, not an empirical one: we do not fit or benchmark models, nor claim these 244 columns are the necessary or sufficient set (§7, §12).
- A **worked demonstration** ingesting two openly-licensed public datasets — a cast-concrete benchmark and a real ~30-laboratory 3DCP database — into one schema, yielding a cross-study print-anisotropy result that the schema's orientation field makes expressible (§6).

---

1. Introduction

3D concrete printing has moved from laboratory demonstrations toward construction-scale experimentation; printed residential structures, bridges, and architectural elements have been produced in many countries worldwide using systems ranging from gantry extruders to six-axis robotic arms [1, 2, 3]. The technical literature now spans printable mortars, alkali-activated binders, fiber-reinforced systems, recycled aggregates, interlayer bond, anisotropic strength, rheology, buildability, and durability. The public record is large enough to support comparative analysis, but not yet consistently shaped enough to make that analysis straightforward.

The central problem is not that researchers fail to report data — many papers report substantial detail — but that the detail is not represented in a common structure. A material may appear as "GGBFS," "slag," "ground granulated blast-furnace slag," or a supplier name; a dosage may be reported in kg/m³, percent of binder, or percent of total mass; a compressive strength may refer to a cast cube, a printed prism, a cored specimen, or a coupon loaded across the layer interface. These distinctions matter scientifically, yet they are often not preserved in a machine-readable way. The consequence is direct: datasets from different groups cannot be combined without extensive manual harmonization, so the field cannot easily build the large, multi-source corpora that modern analysis needs. The most widely used concrete dataset for machine learning — Yeh's UCI Concrete Compressive Strength set [4] — captures only composition and age, with no process, rheology, or orientation fields, and predates 3DCP entirely (§3).

What makes 3DCP fundamentally different from cast concrete can be stated in four points:

1. **Process–property coupling.** The same mix, printed at different speeds, layer heights, and time gaps, can produce substantially different mechanical properties. Process parameters are design variables, not noise.
2. **Anisotropy.** Printed concrete is direction-dependent; specimens tested across the layer interface can be 20–40 % weaker than those tested parallel to the layers [5, 6]. A strength value without an orientation is ambiguous.
3. **Rheological demands.** The mix must be simultaneously pumpable and buildable; yield stress, thixotropy, and open time are critical performance parameters absent from conventional concrete datasets [7].
4. **Interlayer bond.** The weakest link in a printed element is usually the interface between layers, where bond depends on surface moisture, time gap, ambient conditions, and degree of hydration at deposition [8, 9].

Classical single-predictor strength models — Abrams' law [10], Bolomey's equation [11], Féret's formula [12], and the Powers–Brownyard gel–space ratio [49] — were developed for cast concrete and predict strength from a single scalar. They encode an assumption that composition dominates and processing is held constant by standard practice. For 3DCP that assumption breaks down (§7), which motivates multi-feature models and, by extension, a record that captures the full feature space those models need.

2. Scope and non-scope

Open3DCP is a schema specification: it defines *how* to record data; it does not provide the data. It is scoped to **extrusion-based (material-extrusion / FDM-style) 3D concrete printing** — the process whose pump, nozzle, layer schedule, and toolpath the schema captures. Particle-bed / binder-jetting, spray, and slip-form methods are out of scope: they have different process variables and would need their own process block.

Open3DCP **is**:

- A public column vocabulary for 3DCP mix-design and test records.
- A kg/m³-primary, SI-unit reporting convention (with a losslessly-derived mass-percent view).
- A standards-aligned cross-reference layer for common material classes and test methods.
- A flat structure for CSV, SQL, Parquet, dataframe, and repository-deposit workflows.
- A citable artifact with a DOI and Apache-2.0 licensing.

Open3DCP **is not**:

- A dataset or benchmark, a database service, or an API.
- A structural-design method or a code-compliance path.

- A replacement for ASTM, EN, ACI, ICC, RILEM, ISO, or jurisdiction-specific requirements.
- Evidence that any particular mix is safe, durable, printable, or construction-ready.

The schema can record data used in qualification or research workflows, but any construction use still requires appropriate laboratory validation, professional engineering review, and approval under the governing jurisdiction.

3. Background and related work

The ICME paradigm and lessons from metals additive manufacturing. The idea that materials can be designed from performance requirements backward through processing–structure–property relationships was advanced by Olson's work on hierarchical, systems-based materials design [18, 19]; this approach was later formalized as Integrated Computational Materials Engineering (ICME). The Materials Genome Initiative [20] sought to extend it through data infrastructure, funding repositories, ontologies, and schema patterns. In parallel, the FAIR principles for scientific data — Findable, Accessible, Interoperable, Reusable — were articulated by Wilkinson et al. [21]; together they shaped the data-stewardship practices Open3DCP follows. Metals additive manufacturing offers a particularly instructive parallel: early laser-powder-bed and directed-energy-deposition work used ad-hoc formats until ASTM F3049 [13] and the NIST AM Bench program established common reporting practices that enabled cross-laboratory comparison. That effort took most of a decade; 3DCP is at the start of a similar trajectory. The key insight carried over is that **the manufacturing process is a design variable, not a constraint** — printable concrete should be purpose-designed for layer-by-layer deposition, and capturing that requires a record of what makes extrusion different from casting.

Seminal work in 3D concrete printing. Automated construction with cementitious materials dates to Pegna's solid-freeform work [22]; Khoshnevis developed Contour Crafting [23] and Cesaretti et al. demonstrated binder-jetting of regolith simulant [24]. The modern extrusion era began with two groups in parallel: at Loughborough, Buswell, Lim, Le and colleagues published early mix-design and construction-scale studies [25, 26]; at TU Eindhoven, Bos, Wolfs, Ahmed and Salet produced early structural printed elements — including a 3D-printed concrete bridge designed by testing [1, 27] — with Wolfs et al. on early-age behaviour [5] and Suiker on wall stability [28]. Roussel established the rheological framework for printable concrete [7], and Perrot et al. quantified structural build-up [29]. The NTU Singapore group contributed systematic rheology studies including geopolymers [30, 31], printability regions [32], and fresh/hardened characterization [33]. Interlayer bond — the defining weakness — has been studied by Kruger, van Zijl and colleagues on porosity [6], Sanjayan et al. on surface moisture [8], Moelich et al. on quantitative bond models [9], Van Der Putten et al. on surface modification [34], and Marchment and Sanjayan on mesh reinforcement [35]. Structural implications were addressed by Asprone et al. [36] and Gebhard et al. [37]. Broader reviews include Mechtcherine et al. on production physics [38], Buswell et al.'s research roadmap [3], Nerella et al. on strain-based build-up [39], De Schutter et al. on technical/economic/environmental potentials [2], large-scale UHPC work [40], a classification of building systems [41], particle-bed printing [42], fiber-reinforced and recycled-aggregate formulations [43, 44], the 3DCP.fyi citation network [45], and Wangler et al.'s digital-concrete review [46]; the RILEM TC 276-DFC state-of-the-art report consolidates digital-fabrication practice [50].

Standards development. RILEM TC 304-ADC has driven test-method standardization for 3DCP; its interlaboratory study coordinated testing across roughly thirty laboratories and quantified inter-laboratory variability [14, 15, 16], with an accompanying database system for sharing experimental data [17]. Vasilić reviewed the state of standardization, identifying the gap between established *test methods* (which RILEM provides) and a unified *data schema* for analysis-ready storage (which no formal standard yet provides) [47]. Conventional standards — ACI 318 for structural design, ASTM C39 (compressive), C496 (splitting tensile), C1583 (pull-off direct tension), C78 (flexural), C469 (modulus), C191 (setting), C1611 (slump flow), and the EN 197/206/12390 series — supply the measurement framework 3DCP inherits, but were written for cast concrete; the test methods remain valid while the *conditions* of production and the *metadata* needed to interpret results differ.

Existing datasets and their limitations. The most cited ML concrete dataset, Yeh's UCI set [4], captures only composition and age in kg/m³ (requiring a density assumption for formulation-level comparison) and no process, rheology, orientation, specimen, or provenance fields. Repositories such as Mendeley Data and Zenodo host specialized concrete datasets, but — to our knowledge — none provide a general-purpose 3DCP-native schema with first-class process parameters. The closest prior art is the RILEM TC 304-ADC interlaboratory database system [17], the first multi-laboratory 3DCP dataset with standardized protocols; its schema, however, is scoped to that interlaboratory study (its entity model covers materials, specimens, and tests but not a general printing-process vocabulary), whereas Open3DCP aims at a reusable, analysis-ready record across studies. We position Open3DCP as complementary to [17], not a replacement.

4. Why 3DCP needs a 3DCP-native record

Conventional concrete datasets focus on composition, age, and one or more hardened properties — useful for cast concrete, where placement and compaction are treated as standardized. 3DCP makes that assumption unsafe: the manufacturing process is part of the material definition. At minimum, a 3DCP record must distinguish what was weighed into the mix; how it was prepared and modified over time; how it was pumped and extruded; what geometry was deposited; how much time elapsed between adjacent layers; the environmental conditions during deposition; how the specimen was cured and extracted; the loading direction relative to the layer interface; the test method or local protocol; and whether each value was measured, calculated, estimated, or merely reported. Without these attributes, two identical-looking compressive strengths may describe physically different experiments — a cast cube and a printed coupon loaded across interlayers should not collapse into one data point merely because both report MPa.

5. Schema design principles

Open3DCP is governed by a small set of principles, each motivated by the practical requirements of analysis and the lessons of data standardization in adjacent fields.

5.1 Flat schema. Every stored feature is a named column in a single table — no JSON nesting, no graph structure, no join required for basic analysis. This prioritizes adoption by the researchers, curators, and ML practitioners who overwhelmingly work with tabular data (pandas, CSV, SQL) over the representational elegance of graph models such as GEMD, which capture provenance chains and measurement hierarchies but add friction for the common "load a table and train a model" use case. A graph view can be constructed from the flat schema for the structure it captures; conversely, the flat row deliberately omits the full relational/provenance tree (§10), so flat→graph reconstruction is faithful only for the subset the row holds — the two are complementary, not equivalent.

5.2 Dual-basis, kg/m³-primary. Open3DCP records material quantities on a **kg/m³ primary basis** — the convention the concrete industry and field actually use — while retaining **mass-percent of**

total wet mix as a losslessly-derived secondary representation. Three columns make the two interconvertible without assumption: `original_basis` records what the source reported (`kg_m3` | `mass_pct` | `volume` | `lb_yd3`), and `mix_density_kg_m3` and `total_binder_kg_m3` carry the density and binder totals needed to convert exactly in either direction. This resolves a real tension: kg/m³ is what practitioners report and need, but normally requires a density assumption for cross-dataset comparison when density is unreported; mass-percent is self-normalizing ($\Sigma \approx 100$ %) and ideal for pooled ML, but is foreign to field practice. Storing the source basis plus the density bridge serves both without discarding information. (Schema versions ≤ v1.5 used mass-percent as the primary basis; v1.6 made kg/m³ primary and added the bridge columns. Admixtures are recorded as **solids content** by mass: a PCE superplasticizer dosed at 1.0 % liquid with 30 % solids is recorded as 0.3 %.)

5.3 `NULL` is not zero. Open3DCP distinguishes missingness from absence. `NULL` marks a value that is unknown, not reported, not applicable, not measured, or not recoverable without an assumption; `0` is reserved for an explicit source-reported zero or absence. `steel_fiber = 0` is appropriate when a paper states no steel fiber was used; `steel_fiber = NULL` when the paper is simply silent. The distinction is critical for statistics and model training, because false zeros bias means, correlations, feature importance, and learned absence effects.

5.4 Standards alignment without standards substitution. Column names and descriptions reference established standards where they define material classes or test methods — ASTM C150 (cement types), C618 (fly-ash classes), C989 (slag), C1240 (silica fume), C33 (aggregate grading by fineness modulus), C39 and EN 12390-3 (compressive testing), and RILEM TC 304-ADC orientation terminology. These are interoperability hooks, not endorsement or certification: a column can record that a result was produced under a given method, but the schema cannot verify the method was performed correctly.

5.5 Provenance by design and multi-age support. Every record carries a DOI or source citation, a measurement-confidence flag (measured / calculated / estimated / reported), and a laboratory identifier, so downstream users can filter by data quality and trace results to their source. A companion `strength_measurements` table stores results at multiple ages (one hour through 365 days) linked by formulation, supporting strength-development analysis — early-age strength governs buildability while 28-day strength governs structural adequacy, and most datasets report only the latter.

5.6 Documented trade-offs. Two further choices are deliberate departures from convention. *Fineness modulus over maximum aggregate size:* because 3DCP uses only fine aggregate (constrained by pump and nozzle, generally below 4 mm), Open3DCP classifies sand by **fineness modulus** (FM — the summed cumulative mass-percent retained on the standard sieve series, divided by 100; a single index of overall fineness), which is more discriminating than maximum particle size for the fine aggregate (well below the 4.75 mm sand boundary, often under ~4 mm) that printing uses — two sands of equal maximum size can have very different gradations and packing. Trade sand classes (mason / fine / concrete / coarse), mapped onto ASTM C33 grading, are one realization; EN/ISO grading maps onto the same field. *SCM reactivity factors excluded:* the schema stores what was weighed — the mass of each SCM — and leaves reactivity estimation to downstream feature engineering, because reactivity factors are modeling decisions, not raw data.

6. What v1.7 records

Open3DCP v1.7 defines 244 columns in its primary `mix_designs` record (companion tables — `strength_measurements`, `sources`, `test_methods`, `curing_regimes`, `material_aliases` — add linked rows and are not in this count). Figure 1 places the columns in the ICME chain; Figure 2 inventories them under a **Composition–Processing–Conditions–Properties–Provenance (CPPC)** view — a reporting-oriented refinement of the ICME process–structure–property chain that adds explicit Conditions and Provenance legs, not a competing standard. The canonical column-by-column specification is the repository's `Open3DCP_SCHEMA.md` and `sql/create_tables.sql`; this section summarizes the categories.

The 244 is best read as a *vocabulary*, not a per-record dimensionality. It enumerates every reportable field — each cement type, each aggregate size, each fiber, each durability test is its own column — so a typical record leaves most columns `NULL` (the schema is sparse by design). Only four columns are computed from others (`w_c_ratio`, `w_b_ratio`, `a_b_ratio`, `fiber_aspect_ratio`); the rest are independent recorded fields, with five carrying measurement uncertainty, five referencing external files, and the remainder provenance/identity metadata (Figure 4). About three in four columns describe material and performance properties shared with conventional cast concrete; the printing-process and interlayer columns are the additive, 3DCP-specific part (Figure 3).

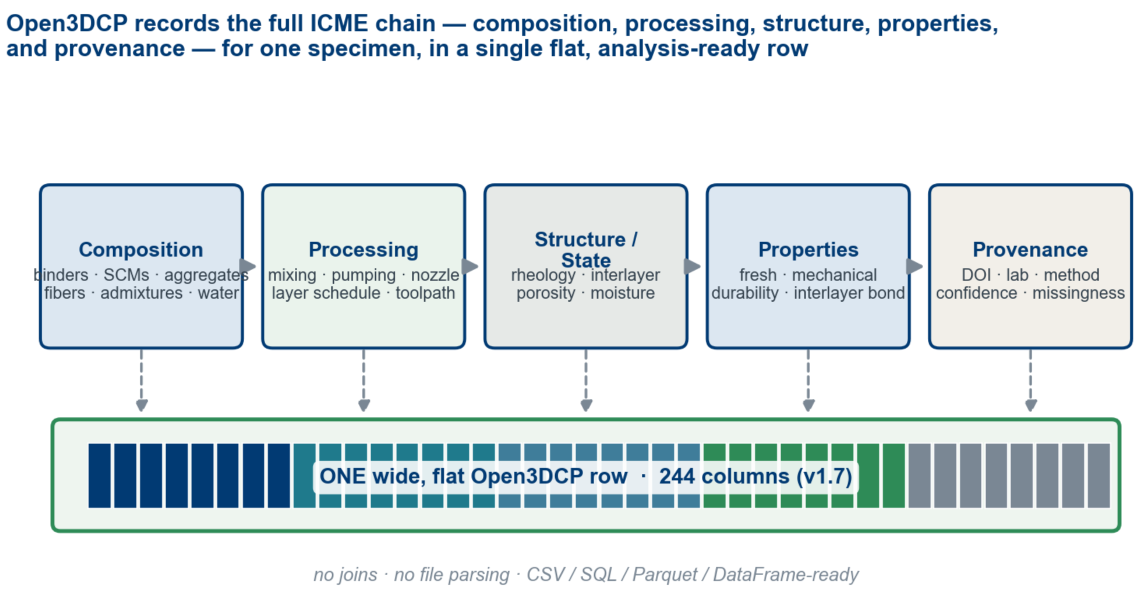

Figure 1: Open3DCP records the full ICME chain — composition, processing, structure/state, properties, and provenance — for one specimen in a single flat, analysis-ready row. The schema is the substrate for ICME-style process–structure–property analysis: it preserves the inputs and outcomes of a printed experiment without joins or file parsing.

Composition. Portland and blended cements, calcium-aluminate and calcium-sulfoaluminate cements, fly ash (generic, Class F, Class C as separate columns per ASTM C618 oxide-sum classification), slag, silica fume, metakaolin, limestone, pumice, bottom ash, rice-husk ash, alkali ac-

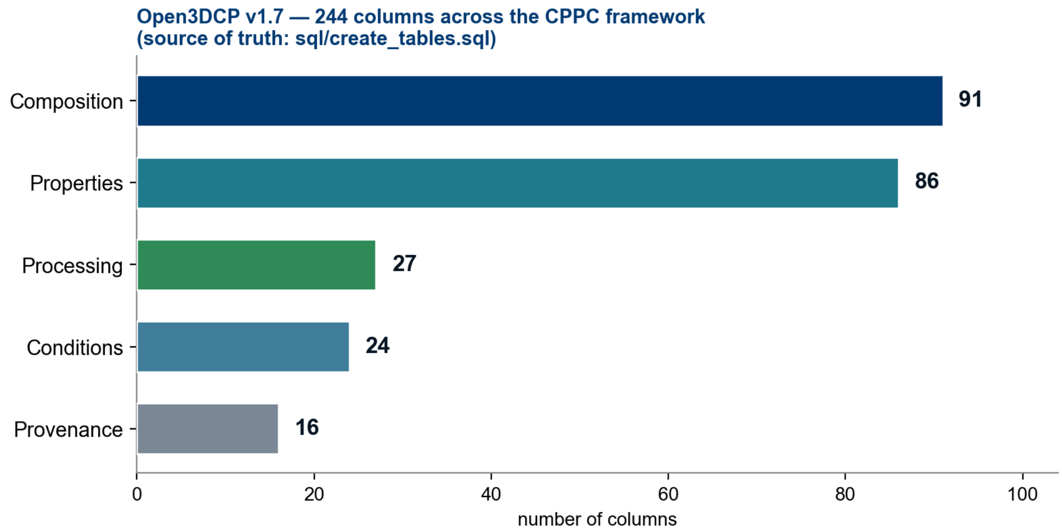

Figure 2: The 244 columns of Open3DCP v1.7 mapped to the CPPC framework. Counts are parsed from `sql/create_tables.sql`, the schema's source of truth. *Composition* (91): binders, SCMs, activators, aggregates, fibers, admixtures, pigments, water, ratios, aggregate conditioning. *Properties* (86): fresh-state, mechanical, interlayer bond, durability, thermal, microstructure. *Processing* (27): 3DCP process parameters, pumping, mixing. *Conditions* (24): test conditions / specimen, environment, exposure class. *Provenance* (16): identity / versioning, source DOI, quality flags.

tivators, nanoscale modifiers, mineral powders, recycled sand, pigments, aggregates, fibers, admixtures, clay rheology modifiers, water, and derived ratios. The schema preserves chemically meaningful distinctions rather than collapsing all cementitious material into a single “binder” field, because fly-ash class, slag, silica fume, limestone, and metakaolin are not interchangeable in hydration, packing, rheology, or long-term performance.

Fibers. Eight core families — `steel_fiber`, `pp_fiber`, `pva_fiber`, `glass_fiber`, `basalt_fiber`, `carbon_fiber`, `nylon_fiber`, `aramid_fiber` — plus `cellulose_fiber` for natural-fiber compatibility (relevant to emerging 3DCP wall-qualification frameworks such as ICC 1150 [51]), with geometry captured separately (`fiber_length_mm`, `fiber_diameter_mm`, `fiber_aspect_ratio`, `fiber_tensile_strength_mpa`). Aspect ratio is the single strongest predictor of fiber contribution to post-crack toughness and is rarely derivable from papers that report only type and dosage.

Fresh-state and rheology. Slump, spread, J-ring, V-funnel, L-box, setting times, air content, fresh unit weight, bleeding, yield stress (static and dynamic), plastic viscosity, thixotropy / structuration rate, open time, and green strength. Printability is not a single property: pumpability, extrudability, shape stability, and open time can move in different directions as water, superplasticizer, VMA, accelerator, grading, and ambient conditions change.

Process parameters. Print speed, layer height and width, layer time gap, nozzle diameter / shape / area, filament width, extrusion rate, number of layers, path length, infill pattern, contour count, print direction, and pumping/mixing/environmental conditions — the processing leg of the ICME chain, absent from prior concrete datasets. The process columns and their typical downstream outcomes are specified in the repository.

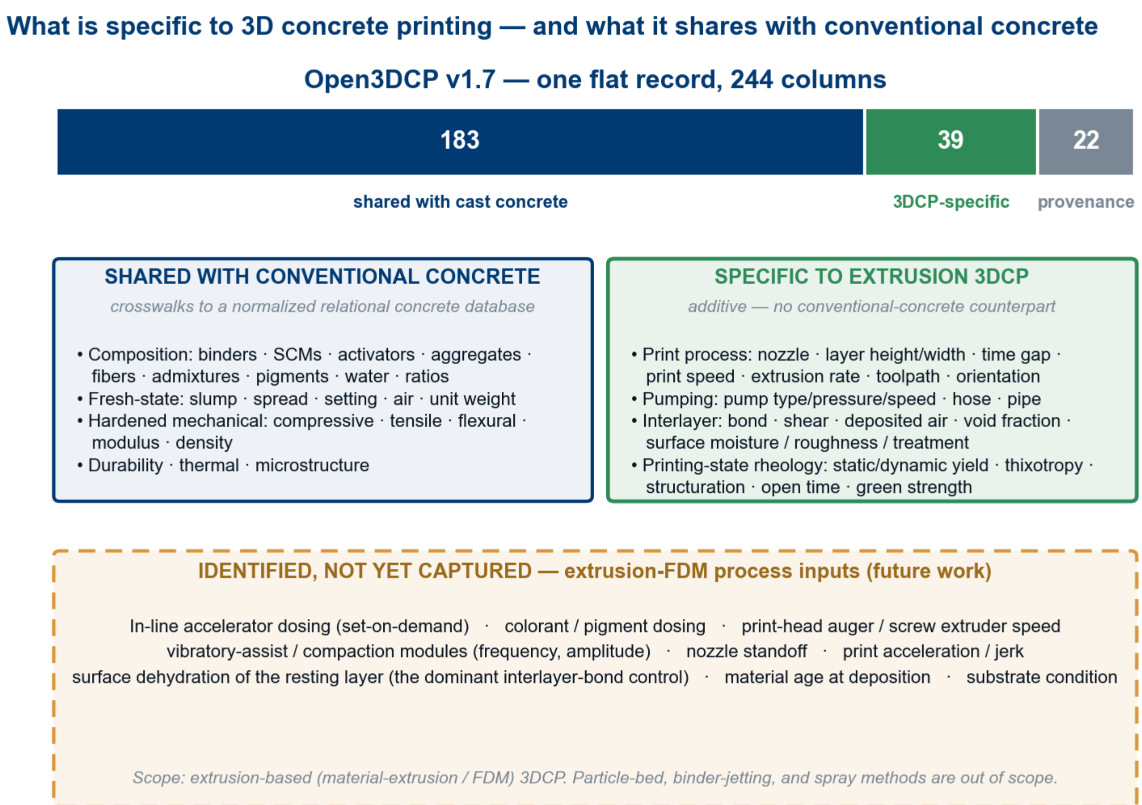

Figure 3: What is specific to 3D concrete printing, and what it shares with conventional concrete. Of the 244 columns, ~183 describe composition and performance properties shared with cast concrete (and therefore crosswalk to a relational concrete database); ~39 are specific to extrusion 3DCP (printing process, pumping, interlayer bond, printing-state rheology) with no conventional-concrete counterpart; the remaining ~22 are provenance. (This is a different cut of the same 244 than the CPPC inventory in Figure 2 — by origin rather than reporting role — and also sums to 244.) The amber annex lists extrusion-process inputs identified but not yet captured — a future-work agenda.

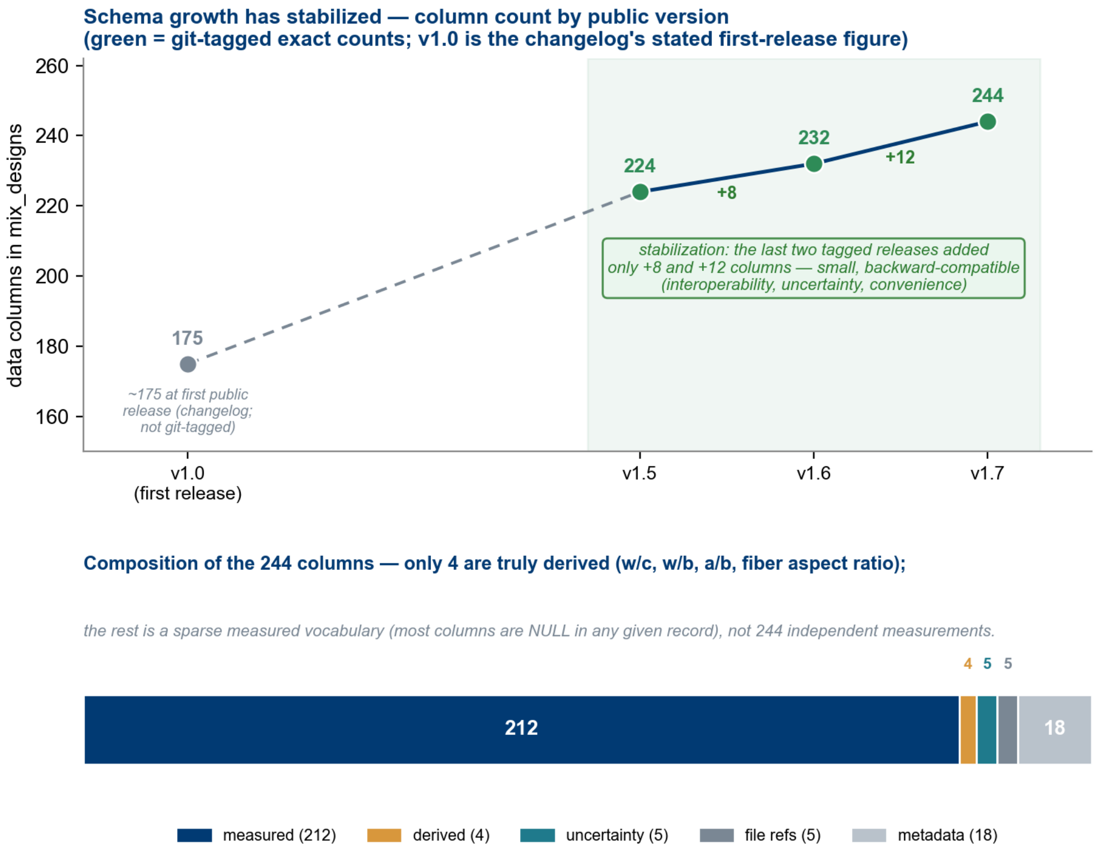

Figure 4: Schema growth has stabilized (top): the last two tagged releases added only +8 and +12 columns, all backward-compatible (interoperability, uncertainty, and convenience fields). Composition of the 244 (bottom): only four columns are derived from others, so the count is a sparse measured vocabulary — the union of every reportable field, most NULL in any record — not 244 independent measurements. Only the git-tagged releases (v1.5–v1.7) carry exact counts; v1.0 is the changelog's stated first-release figure (~175). The 244 are data columns, excluding the SERIAL primary key.

Specimen and test context. Specimen preparation, geometry, dimensions, extraction method, curing conditions, test age, test method, number of specimens averaged, and **test orientation** — especially important because printed concrete is anisotropic. Orientation is a controlled vocabulary (Table 1).

Table 1: Test-orientation controlled vocabulary. Strength ordering is typical; specific values depend on mix design and process parameters [5, 6].

Code	Axis	Description	Typical strength
X	Longitudinal	Parallel to extrusion direction	Highest
Y	Transverse	Perpendicular, within layer plane	Moderate
Z	Interlayer	Perpendicular to layer interfaces	Lowest
XY_45	In-plane	45° diagonal in layer plane	X–Y intermediate
XZ_45	Cross-layer	45° diagonal across layers	X–Z intermediate
CAST	Isotropic	Moulded reference specimen	Baseline

Hardened, durability, and interlayer. Compressive, tensile, splitting tensile, flexural, modulus, bond, fracture energy, toughness, impact, fatigue, density, and Poisson's ratio; a durability suite covering chloride transport, carbonation, shrinkage, creep, freeze–thaw, sulfate and ASR expansion, permeability, absorption, sorptivity, scaling, corrosion indicators, thermal properties, and fire resistance; and interlayer columns for bond, shear, void-area fraction, deposited air content, surface roughness, surface moisture state, and surface treatment. The schema does not assert that every dataset measures all of these; it provides stable homes when the measurements exist.

Provenance, basis, and uncertainty. Beyond DOI, citation, confidence, lab, and quality flags, v1.6 added per-measurement uncertainty columns (e.g. compressive_strength_stddev_mpa), raw-data references that keep large payloads external (raw_data_doi, stress_strain_file, rheology_curve_file, microstructure_image, raw_data_file), and the basis columns of §5.2. Version 1.7 added a material_class classification, a batch timeline (batch_label, date_of_casting), and aggregate-conditioning columns (aggregate_moisture_state, aggregate_absorption_p aggregate_moisture_content_pct, aggregate_prewetted) that make effective mix water recoverable when aggregates are batched off the SSD reference.

Quantifying the cost of flattening. Projecting a heterogeneous or relational source onto one flat row can lose information. Rather than hide that, the fidelity of the mapping can be scored against the source and reported — a proposed, not-yet-calibrated convention (§11) — so the cost of each conversion is recorded rather than hidden. A machine-readable crosswalk to a normalized relational concrete database, and a specification of the process-parameter columns that have no conventional relational counterpart, are included in the repository.

Worked demonstration: two public datasets in one schema. To show the schema works on real data — not only by design — we ingested two openly-licensed public datasets into Open3DCP and report the result (Figure 5). The UCI/Yeh (1998) concrete dataset [4], a *cast*-concrete benchmark, maps onto the composition and hardened-strength columns and ingests at a fidelity of **96.7/100 (A)** over its 126 source fields. The RILEM TC 304-ADC interlaboratory study on the mechanical properties of 3D-printed concrete [52] — a *real 3DCP* database from roughly thirty laboratories, stored as a normalized relational SQLite export — was curated into Open3DCP and populates the columns UCI leaves empty: 3DCP process, fresh-state rheology, hardened mechanical, **orientation**, and the per-measurement uncertainty columns (mean ± standard deviation ± n). (An automated fidelity score for the RILEM source awaits a SQLite reader; the present ingestion is curated through the relational structure.) Together the two sources light up complementary slices of the record, and Open3DCP spans their union (Figure 5, left).

The demonstration also yields a result that **cannot be expressed without an orientation field, nor compared across laboratories without a shared schema**: print anisotropy. Mapping the RILEM U/V/W loading codes onto Open3DCP's `test_orientation_code` (X/Y/Z/CAST), the same printed mortar is **32–55 % weaker loaded along the layers** than the cast reference, consistently across three independent laboratories (Figure 5, right) — direct empirical support for the argument of §7. This is a demonstration of *ingestion and interoperability*, not a predictive-model benchmark; fitting models on pooled Open3DCP data is future work (§12).

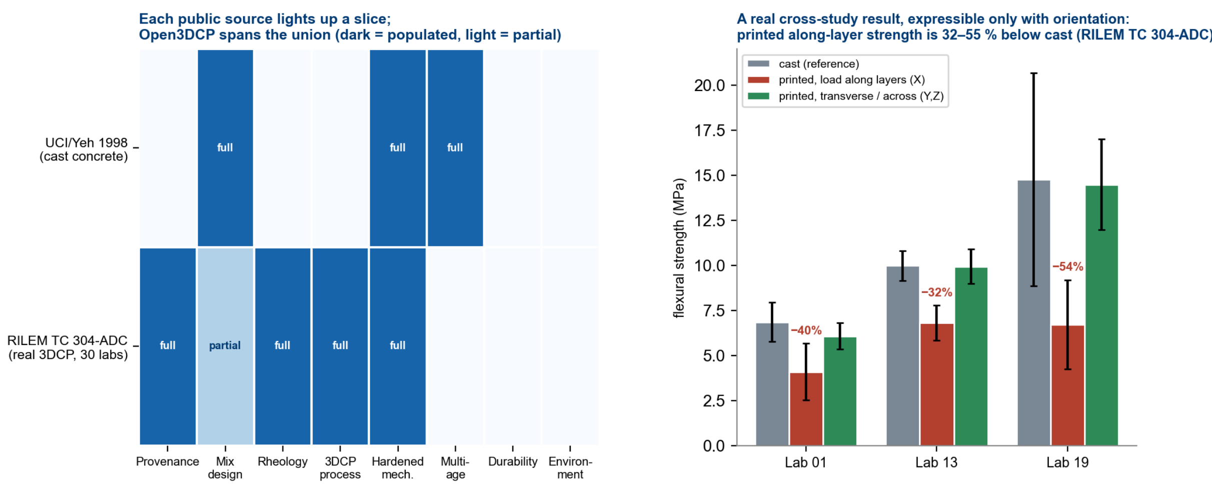

Figure 5: **Worked demonstration.** *Left:* two openly-licensed public datasets ingested into Open3DCP — UCI/Yeh (cast) and the RILEM TC 304-ADC interlaboratory study (real 3DCP, ~30 labs) — each populating complementary slices of the record (dark = populated, light = partial), with Open3DCP spanning the union. *Right:* a real cross-study result from the RILEM data, expressible only because the schema records loading orientation — printed concrete loaded along the layers (X) is 32–55 % weaker than the cast reference, consistently across three independent laboratories (flexural strength, ~28 d; error bars are ± one standard deviation).

7. The case for multi-feature models

The concrete industry has relied on empirical strength models for over a century: Abrams' law [10], Bolomey's equation [11], Féret's formula [12], and the Powers–Brownyard gel–space ratio [49] (the last a function of degree of hydration rather than a pure composition ratio) predict strength from a single scalar, on the assumption that composition dominates and processing is held constant by standard practice. For cast concrete under controlled conditions, this is reasonable. For 3DCP the assumption is violated in at least four ways: **process variation** (two specimens of identical composition but different print speed, layer height, and time gap can differ substantially in compressive strength — the process-parameter literature reports strength changes of order 10–30 % from time-gap and speed alone [7, 38]), **anisotropy** (printed specimens are direction-dependent, with cross-layer strength commonly 20–40 % below in-plane [5, 6], a spread a one-output model cannot represent), **SCM complexity** (modern mixes contain three to five binders whose reactivities differ by orders of magnitude, which a single w/b ratio treats equally), and **rheological constraints** (the admixture cocktail that achieves printability may differ from what optimizes strength at the same w/b). The RILEM TC 304-ADC interlaboratory study [14, 15] supplies evidence at scale: the same mix design, printed across laboratories with different equipment and process parameters, produced inter-laboratory variability substantially larger than cast-concrete round-robins, driven primarily by process differences — exactly the variables single-predictor models ignore. The worked demonstration in §6 (Figure 5) shows this directly: the same mortar is 32–55 % weaker loaded along the print layers than cast, consistently across three laboratories.

This is an *analytical* motivation, not an empirical demonstration: it shows that single-predictor models *cannot represent* 3DCP's process- and orientation-dependence, which makes multi-feature models necessary, and a multi-feature model can only use features that were recorded — hence the value of a comprehensive record. It does **not** establish that *these* 244 columns are the necessary or sufficient set; that is an empirical question the schema is built to let others answer through feature-ablation studies as data coverage improves (§12). We fit no model and benchmark no predictor here.

8. Digital twin and ICME framing

A full digital twin of a 3DCP process spans at least four layers of information: a **material definition** (composition, product identity, particle characteristics, water/admixture basis, fiber geometry); a **process definition** (mixing, pumping, nozzle geometry, motion, extrusion rate, layer schedule, environment, curing); **state and structure** (fresh rheology, thixotropic recovery, interlayer surface condition, porosity, moisture, temperature, hydration, fiber orientation, microstructure); and **performance and provenance** (mechanical and durability properties, test method, specimen geometry, orientation, lab, confidence, source). Open3DCP covers much of the first, second, and fourth, and a useful subset of the third (Figure 1). **Open3DCP is not a digital twin.** A digital twin, as the digital-fabrication community uses the term, requires live, bidirectional coupling to the running process; a static, scalar, single-row schema has none. Open3DCP is better described as a *structured experiment record* — a substrate from which many aspects of published 3DCP work can be reconstructed and compared, and on which future digital twins could be built. Full process twins will additionally require time-series machine logs, synchronized sensor data, imaging, and richer links to raw files.

9. What we cannot capture — and why

This section identifies features that are physically important for 3DCP but cannot currently be measured reliably. The taxonomy is intended to guide instrumentation research and future schema extensions, not to claim the schema is complete.

9.1 Real-time process-monitoring gaps. The schema stores rheology as point measurements (typically a pre-print laboratory rheometer reading), but the material's rheological state changes continuously from mixer through pump, hose, and nozzle; no inline rheometer exists for cementitious materials at production scale. Pump pressure is a single scalar, yet in practice it fluctuates with consistency and toolpath back-pressure in ways that correlate with segregation and flow discontinuities. Nozzle standoff is a nominal value that varies with robot positioning and substrate deformation. And print-head dynamics — acceleration, deceleration, cornering — cause local variation in deposition rate not captured by a single print-speed value.

9.2 In-situ material-state gaps. The actual interlayer moisture at the moment of deposition is a continuous field depending on ambient humidity, wind, time gap, and diffusivity; Sanjayan et al. [8] showed it affects bond by 20–40 % and Moelich et al. [9] modeled it quantitatively, yet no standardized protocol measures it in production. The degree of hydration at each interface forms a gradient across layers, but the schema stores a single destructively-measured value. Fiber-orientation distribution critically affects directional strength but can only be measured destructively by micro-CT after hardening. Aggregate packing within the filament and the internal temperature field during exothermic curing are likewise not measurable in-situ.

9.3 Characterization-protocol gaps. Interlayer void fraction requires destructive cross-sectional imaging with no standardized analysis; surface roughness at layer interfaces has no standard protocol for 3DCP; and ambient conditions reported as room averages may differ from the point of deposition in large-scale or outdoor printing.

9.4 Process inputs not yet captured. Beyond the real-time signals above, several extrusion-process *inputs* an operator sets are not yet schema columns: in-line accelerator dosing (set-on-demand), colorant / pigment dosing, print-head auger / screw extruder speed (distinct from the bulk-mixer speed the schema records), and vibratory-assist or compaction modules — alongside nozzle standoff, print acceleration / jerk, surface dehydration of the resting layer (the dominant interlayer-bond control), and material age at deposition (Figure 3, annex). These are an explicit future-work agenda; the schema extends additively as they are specified.

9.5 The digital-twin horizon. A complete twin would require on the order of 300+ parameters, including time-series data that cannot be represented as scalar columns. Open3DCP captures the formulation, process, and performance layers of a record well, plus a useful part of the in-situ state layer; the largest remaining gap is the real-time process data that current hardware does not routinely capture. As sensor technology improves — inline rheometers, thermal imaging tied to print-head motion, machine vision for filament geometry — the flat schema can extend additively without breaking existing queries or models. The gap analysis is not a criticism of the schema's completeness: **the features we cannot measure today are, in many cases, the features that would most improve predictive models**, and closing these gaps requires instrumentation, not schema design.

10. Adoption path and community

Adoption does not require populating every column — null columns are ignored during analysis. A laboratory can adopt the schema by (1) mapping local column names to canonical Open3DCP names; (2) recording material quantities with the source basis preserved (`original_basis`) so kg/m³ ↔ mass-percent remains exact; (3) preserving missing values as `NULL` and using `0` only for explicit zeros; (4) filling provenance fields before analysis fields; (5) recording test method, specimen geometry, age, and orientation for every mechanical result; (6) depositing the dataset in a public repository when rights allow; and (7) citing the schema and the original test methods. The flat schema is database-agnostic (PostgreSQL, SQLite, CSV, Parquet, pandas/polars/R) and pairs naturally with standard ML libraries for an end-to-end, open-source pipeline. By design it is **FAIR-aligned** [21]: every record carries a DOI or citation (Findable), the schema is open under Apache-2.0 (Accessible), naming follows ASTM/EN/RILEM with SI units and controlled vocabularies (Interoperable), and the license plus provenance metadata supports reuse (Reusable).

Because most of the schema is shared with conventional concrete (§6), Open3DCP interoperates closely with a normalized relational concrete database (Figure 6). Material composition and test results map both ways — exactly for ages, geometry, and ratios; with a recorded density for the kg/m^3 ↔ mass-percent step. The correlation is close but the round-trip is *not* lossless: relational structure that a flat row cannot hold (parametrized geometry, reinforcement layouts, devices, loading histories) collapses to a side record, and the extrusion-3DCP process and interlayer columns have no conventional relational table to map into. Stating these boundaries is what makes the mapping auditable rather than asserted.

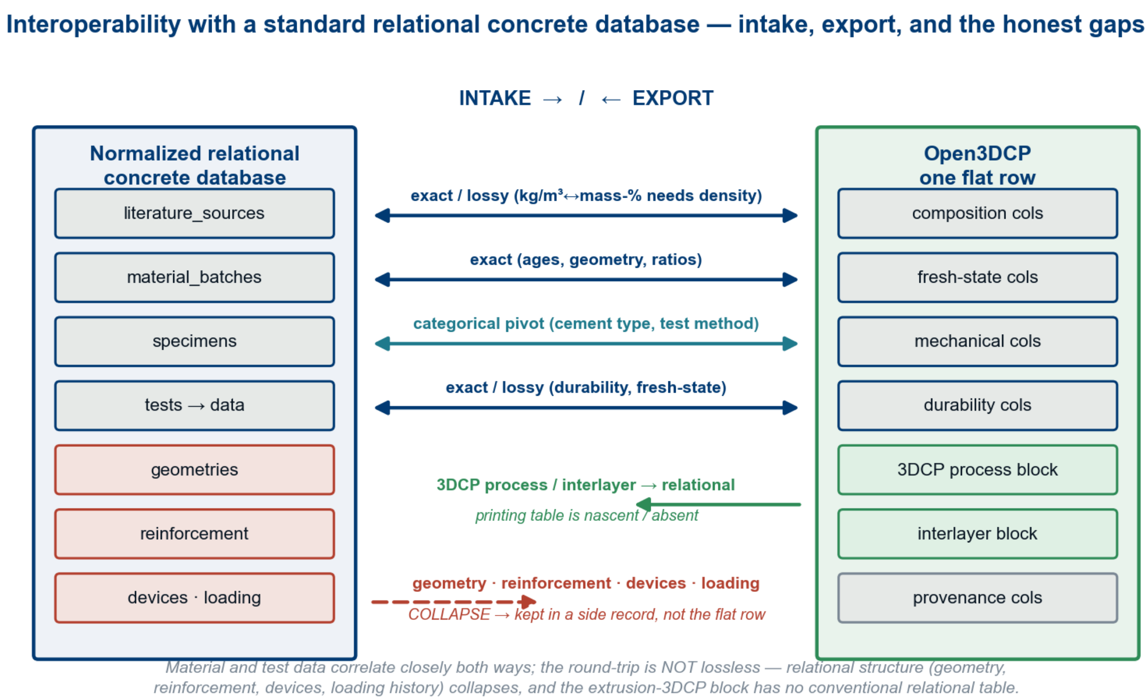

Figure 6: Interoperability with a normalized relational concrete database. Material and test data map both ways by fidelity class (exact / lossy / categorical); the round-trip is not lossless — relational structure (geometry, reinforcement, devices, loading history) collapses to a side record, and the extrusion-3DCP block has no conventional relational counterpart.

We invite the 3DCP community to adopt common column names and units in published datasets — even partial adoption of shared names for fields like `compressive_strength_mpa`, `w_b_ratio`, `layer_height_mm`, and `test_orientation_code` would sharply reduce the effort of combining datasets — and we encourage standards bodies and the community to evaluate Open3DCP (or a derivative) as one input to future 3DCP data-reporting practice, extending the RILEM TC 304-ADC database approach [17] from interlaboratory studies to routine publication.

11. Limitations

These limitations bound the claims above; §12 maps each to planned work.

1. **Coverage is a v1.7 snapshot.** The 244-column count and per-category figures reflect `sql/create_tables.sql` at v1.7; the schema is under active, additive development and these numbers evolve. The canonical reference is always the repository.
2. **Single-maintainer governance and unquantified curation reliability.** The schema and its controlled vocabularies are maintained by one author; column choices and any scoring conventions are an expert convention, not yet ratified by a working group or calibrated against a labelled benchmark. The reliability of human curation — `NULL`-vs-`0` judgements, confidence flags, basis assignment, and the resolution of conflicting or duplicate source values — is likewise not yet quantified (no inter-curator agreement study).
3. **The measurement-gap taxonomy is qualitative.** §9 names gaps and their physical importance but does not quantify how much each would improve a model; that ordering awaits feature-ablation studies on datasets built with the schema.
4. **Typical ranges and orientation orderings are indicative**, not gates, and not yet sourced to a fixed citation set; Table 1's strength ordering is typical, not measured here.
5. **Adoption at scale is unproven.** Cross-dataset interoperability is asserted by design; no large, heterogeneous, multi-source corpus has yet been assembled and modeled end-to-end on the public schema.
6. **The schema is not a substitute for validation.** It records what was done and measured; it does not certify structures, validate models, or establish that a mix is printable, durable, or safe.

12. Future work

1. Assemble and publish a **multi-source corpus** on the public schema and report cross-dataset model performance and the data-coverage distribution across the CPPC categories.
2. **Validate the source-to-schema mapping** — tie any fidelity scoring of the mapping to a labelled benchmark or a measured downstream consequence, and add a sensitivity analysis.
3. **Quantify the measurement gaps** of §9 through feature-ablation studies, converting the qualitative taxonomy into a ranked instrumentation agenda.
4. Move governance toward a **working-group / standards process** (with RILEM, ACI, or ASTM) so the vocabulary and grade bands are community-ratified rather than single-author conventions.
5. Extend additively as instrumentation matures (inline rheometry, thermal imaging, machine vision), keeping the flat-schema backward-compatibility guarantee.

13. Data and code availability

- **Schema & tooling:** Open3DCP v1.7 [48] — `github.com/sunnyday-technologies/Open3DCP` (Apache-2.0); Zenodo concept DOI `10.5281/zenodo.19647470` (always resolves to the latest version of record). The canonical column list is `Open3DCP_SCHEMA.md` / `sql/create_tables.sql`; companion tables include `strength_measurements`, `sources`, `test_methods`, `curing_regimes`, and `material_aliases`.
- **Reproducing the figures:** the column counts in Figures 2 and 4 are parsed from `sql/create_tables.sql`, the schema's single source of truth; the figures are regenerated by the committed figure scripts (matplotlib, PNG ≥ 200 dpi + SVG).

- **Supporting material:** a machine-readable crosswalk to a normalized relational concrete database and a specification of the 3DCP process-parameter columns are included in the repository.

14. Manuscript license and author statements

Author contributions. Nicholas Sonnentag: conceptualization, methodology, software, data curation, writing — original draft, review & editing.

Competing interest. The author is the founder of Sunnyday Technologies, which develops Open3DCP and related 3DCP tools. Open3DCP is released under an open-source license (Apache 2.0) with no commercial restriction on use.

Acknowledgments. The author thanks the RILEM TC 304-ADC committee for foundational work on 3DCP test standardization and interlaboratory data sharing, and the broader 3DCP research community whose published data motivated and informed the schema.

15. References

1. F. Bos, R. Wolfs, Z. Ahmed, T. Salet (2016). Additive manufacturing of concrete in construction: potentials and challenges of 3D concrete printing. *Virtual and Physical Prototyping* 11(3) 209–225. https://doi.org/10.1080/17452759.2016.1209867
2. G. De Schutter, K. Lesage, V. Mechtcherine, V. N. Nerella, G. Habert, I. Agustí-Juan (2018). Vision of 3D printing with concrete — technical, economic and environmental potentials. *Cement and Concrete Research* 112 25–36. https://doi.org/10.1016/j.cemconres.2018.06.001
3. R. A. Buswell, W. R. Leal de Silva, S. Z. Jones, J. Dirrenberger (2018). 3D printing using concrete extrusion: a roadmap for research. *Cement and Concrete Research* 112 37–49. https://doi.org/10.1016/j.cemconres.2018.05.006
4. I-C. Yeh (1998). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research* 28(12) 1797–1808. https://doi.org/10.1016/S0008-8846(98)00165-3
5. R. J. M. Wolfs, F. P. Bos, T. A. M. Salet (2018). Early age mechanical behaviour of 3D printed concrete: numerical modelling and experimental testing. *Cement and Concrete Research* 106 103–116. https://doi.org/10.1016/j.cemconres.2018.02.001
6. J. Kruger, A. du Plessis, G. van Zijl (2021). An investigation into the porosity of extrusion-based 3D printed concrete. *Additive Manufacturing* 37 101740. https://doi.org/10.1016/j.addma.2020.101740
7. N. Roussel (2018). Rheological requirements for printable concretes. *Cement and Concrete Research* 112 76–85. https://doi.org/10.1016/j.cemconres.2018.04.005
8. J. G. Sanjayan, B. Nematollahi, M. Xia, T. Marchment (2018). Effect of surface moisture on inter-layer strength of 3D printed concrete. *Construction and Building Materials* 172 468–475. https://doi.org/10.1016/j.conbuildmat.2018.03.232

9. G. M. Moelich, J. Kruger, R. Combrinck (2021). Modelling the interlayer bond strength of 3D printed concrete with surface moisture. *Cement and Concrete Research* 150 106559. https://doi.org/10.1016/j.cemconres.2021.106559
10. D. A. Abrams (1918). *Design of Concrete Mixtures.* Bulletin 1, Structural Materials Research Laboratory, Lewis Institute.
11. J. Bolomey (1935). Granulation et prévision de la résistance probable des bétons. *Travaux* 19(230) 228–232.
12. R. Féret (1892). Sur la compacité des mortiers hydrauliques. *Annales des Ponts et Chaussées* 4 5–164.
13. ASTM International (2014). *ASTM F3049 — Standard Guide for Characterizing Properties of Metal Powder Feedstock.* ASTM International, West Conshohocken, PA.
14. RILEM TC 304-ADC (2025). Interlaboratory study on 3D printed concrete: approach and main results. *Materials and Structures.* https://doi.org/10.1617/s11527-025-02686-x
15. RILEM TC 304-ADC (2025). Compressive strength and modulus of elasticity of 3D printed concrete: interlaboratory study results. *Materials and Structures.* https://doi.org/10.1617/s11527-025-02688-9
16. RILEM TC 304-ADC (2025). Flexural and tensile strength of 3D printed concrete: interlaboratory study results. *Materials and Structures.* https://doi.org/10.1617/s11527-025-02687-w
17. RILEM TC 304-ADC (2025). Design and implementation of a database system for querying, sharing, and analyzing experimental data from interlaboratory studies on 3D printed concrete. *Materials and Structures.* https://doi.org/10.1617/s11527-025-02650-9
18. G. B. Olson (1997). Computational design of hierarchically structured materials. *Science* 277(5330) 1237–1242. https://doi.org/10.1126/science.277.5330.1237
19. G. B. Olson (2000). Designing a new material world. *Science* 288(5468) 993–998. https://doi.org/10.1126/science.288.5468.993
20. National Science and Technology Council (2011). *Materials Genome Initiative for Global Competitiveness.* Executive Office of the President, Washington, DC.
21. M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3 160018. https://doi.org/10.1038/sdata.2016.18
22. J. Pegna (1997). Exploratory investigation of solid freeform construction. *Automation in Construction* 5(5) 427–437. https://doi.org/10.1016/S0926-5805(96)00166-5
23. B. Khoshnevis (2004). Automated construction by contour crafting — related robotics and information technologies. *Automation in Construction* 13(1) 5–19. https://doi.org/10.1016/j.autcon.2003.08.012
24. G. Cesaretti, E. Dini, X. De Kestelier, V. Colla, L. Pambaguian (2014). Building components for an outpost on the lunar soil by means of a novel 3D printing technology. *Acta Astronautica* 93 430–450. https://doi.org/10.1016/j.actaastro.2013.07.034
25. T. T. Le, S. A. Austin, S. Lim, R. A. Buswell, A. G. F. Gibb, T. Thorpe (2012). Mix design and fresh properties for high-performance printing concrete. *Materials and Structures* 45(8) 1221–1232. https://doi.org/10.1617/s11527-012-9828-z
26. S. Lim, R. A. Buswell, T. T. Le, S. A. Austin, A. G. F. Gibb, T. Thorpe (2012). Developments in construction-scale additive manufacturing processes. *Automation in Construction* 21 262–268. https://doi.org/10.1016/j.autcon.2011.06.010
27. T. A. M. Salet, Z. Y. Ahmed, F. P. Bos, H. L. M. Laagland (2018). Design of a 3D printed concrete bridge by testing. *Virtual and Physical Prototyping* 13(3) 222–236. https://doi.org/10.1080/17452759.2018.1476064
28. A. S. J. Suiker (2018). Mechanical performance of wall structures in 3D printing processes:

theory, design tools and experiments. *International Journal of Mechanical Sciences* 137 145–170. https://doi.org/10.1016/j.ijmecsci.2018.01.010
29. A. Perrot, D. Rangeard, A. Pierre (2016). Structural built-up of cement-based materials used for 3D-printing extrusion techniques. *Materials and Structures* 49(4) 1213–1220. https://doi.org/10.1617/s11527-015-0571-0
30. B. Panda, C. Unluer, M. J. Tan (2018). Investigation of the rheology and strength of geopolymer mixtures for extrusion-based 3D printing. *Cement and Concrete Composites* 94 307–314. https://doi.org/10.1016/j.cemconcomp.2018.10.002
31. B. Panda, C. Unluer, M. J. Tan (2019). Extrusion and rheology characterization of geopolymer nanocomposites used in 3D printing. *Composites Part B: Engineering* 176 107290. https://doi.org/10.1016/j.compositesb.2019.107290
32. Y. W. D. Tay, Y. Qian, M. J. Tan (2019). Printability region for 3D concrete printing using slump and slump flow test. *Composites Part B: Engineering* 174 106968. https://doi.org/10.1016/j.compositesb.2019.106968
33. S. C. Paul, Y. W. D. Tay, B. Panda, M. J. Tan (2018). Fresh and hardened properties of 3D printable cementitious materials for building and construction. *Archives of Civil and Mechanical Engineering* 18(1) 311–319. https://doi.org/10.1016/j.acme.2017.02.008
34. J. Van Der Putten, G. De Schutter, K. Van Tittelboom (2019). Surface modification as a technique to improve inter-layer bonding strength in 3D printed cementitious materials. *RILEM Technical Letters* 4 33–38. https://doi.org/10.21809/rilemtechlett.2019.84
35. T. Marchment, J. Sanjayan (2020). Mesh reinforcing method for 3D concrete printing. *Automation in Construction* 109 102992. https://doi.org/10.1016/j.autcon.2019.102992
36. D. Asprone, C. Menna, F. P. Bos, T. A. M. Salet, J. Mata-Falcón, W. Kaufmann (2018). Rethinking reinforcement for digital fabrication with concrete. *Cement and Concrete Research* 112 111–121. https://doi.org/10.1016/j.cemconres.2018.05.020
37. L. Gebhard, J. Mata-Falcón, A. Anton, B. Dillenburger, W. Kaufmann (2021). Structural behaviour of 3D printed concrete beams with various reinforcement strategies. *Engineering Structures* 240 112380. https://doi.org/10.1016/j.engstruct.2021.112380
38. V. Mechtcherine, F. P. Bos, A. Perrot, W. R. Leal da Silva, V. N. Nerella, S. Fataei, R. J. M. Wolfs, M. Sonebi, N. Roussel (2020). Extrusion-based additive manufacturing with cement-based materials — production steps, processes, and their underlying physics: a review. *Cement and Concrete Research* 132 106037. https://doi.org/10.1016/j.cemconres.2020.106037
39. V. N. Nerella, M. A. B. Beigh, S. Fataei, V. Mechtcherine (2019). Strain-based approach for measuring structural build-up of cement pastes in the context of digital construction. *Cement and Concrete Research* 115 530–544. https://doi.org/10.1016/j.cemconres.2018.08.003
40. C. Gosselin, R. Duballet, P. Roux, N. Gaudillière, J. Dirrenberger, P. Morel (2016). Large-scale 3D printing of ultra-high performance concrete — a new processing route for architects and builders. *Materials & Design* 100 102–109. https://doi.org/10.1016/j.matdes.2016.03.097
41. R. Duballet, O. Baverel, J. Dirrenberger (2017). Classification of building systems for concrete 3D printing. *Automation in Construction* 83 247–258. https://doi.org/10.1016/j.autcon.2017.08.018
42. D. Lowke, E. Dini, A. Perrot, D. Weger, C. Gehlen, B. Dillenburger (2018). Particle-bed 3D printing in concrete construction — possibilities and challenges. *Cement and Concrete Research* 112 50–65. https://doi.org/10.1016/j.cemconres.2018.05.018
43. G. Ma, Z. Li, L. Wang (2018). Printable properties of cementitious material containing copper tailings for extrusion-based 3D printing. *Construction and Building Materials* 162 613–627.

https://doi.org/10.1016/j.conbuildmat.2017.12.051

44. T. Ding, J. Xiao, S. Zou, Y. Wang (2020). Hardened properties of layered 3D printed concrete with recycled sand. *Cement and Concrete Composites* 113 103724. https://doi.org/10.1016/j.cemconcomp.2020.103724
45. 3DCP.fyi contributors (2024). 3DCP.fyi — a comprehensive citation network graph on the state of the art in 3D concrete printing. *Fourth RILEM International Conference on Concrete and Digital Fabrication (DC2024).* https://doi.org/10.1007/978-3-031-70031-6_62
46. T. Wangler, N. Roussel, F. P. Bos, T. A. M. Salet, R. J. Flatt (2019). Digital concrete: a review. *Cement and Concrete Research* 123 105780. https://doi.org/10.1016/j.cemconres.2019.105780
47. K. Vasilić (2025). Standardization aspects of concrete 3D printing: state of the art, requirements and first steps towards standardization. *RILEM Technical Letters* 9 98–105. https://doi.org/10.21809/rilemtechlett.2024.201
48. Sunnyday Technologies (2026). *Open3DCP: Open Data Standard for 3D Concrete Printing* (v1.7). https://github.com/sunnyday-technologies/Open3DCP. Apache License 2.0; Zenodo concept DOI `10.5281/zenodo.19647470`.
49. T. C. Powers, T. L. Brownyard (1946). Studies of the physical properties of hardened Portland cement paste. *Journal of the American Concrete Institute* 43 101–132.
50. N. Roussel, D. Lowke (eds.) (2022). *Digital Fabrication with Cement-Based Materials: State-of-the-Art Report of RILEM TC 276-DFC.* RILEM SOAR Vol. 36, Springer. https://doi.org/10.1007/978-3-030-90535-4
51. International Code Council (2026). *ICC 1150-2026 Standard for Automated Construction Technology for 3D Printing Walls.* ICC. ISBN 978-1-971077-70-3.
52. RILEM TC 304-ADC (2024). *Database of the RILEM TC 304-ADC interlaboratory study on mechanical properties of 3D printed concrete* [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.12200570 (CC BY 4.0).

Corresponding author: Nicholas Sonnentag, Sunnyday Technologies (nick@sunn3d.com).