Title

WHITE PAPER

Model-as-Trigger, Appraisal-as-Arbiter

A hybrid architecture for valuing fine art and collectibles: why automated valuation models cannot price unique assets, and what works instead

George Fortin, MBA · Pascal Kropf, PhD

Title Collections Inc. · July 2026

Contents

Key takeaways

KEY TAKEAWAYS

1. AVMs only work where many near-identical assets trade often and every sale lands in a dataset. Fine art fails all three conditions, categorically.

2. Even the best case broke: Zillow lost $881 million in a single year transacting on its own industry-leading model, and commercial real estate never adopted AVMs at all; institutional CRE is still valued by appraisal, by necessity.

3. Art valuations already fail when tested: the IRS adjusted 47% of the art values it reviewed in 2023, on nearly $800 million of claimed value, with errors running in both directions.

4. The rules require a person. USPAP, the RICS Red Book, IVS 105, and US Treasury regulations each independently disqualify unassisted model output as a valuation.

5. We built a fine art AVM and landed on a different answer: let the model flag probable value drift, and let a qualified appraisal, routed and captured at scale, set the value. Model-as-Trigger, Appraisal-as-Arbiter.

1. Executive summary

Every financial decision that touches fine art and collectibles rests on a valuation someone must trust. Premiums and claims, net-worth statements and allocation decisions, estate settlements, matrimonial divisions, charitable deductions, and collateral advances all consume the same input, and the input is only as good as the process behind it. The gold standard of valuation, in the absence of an actual transaction, is a qualified human appraisal. That standard is hard to scale, and the resulting friction explains why many attempts at automating the valuation process have been pursued and keep being tried in the market.

This paper examines whether automated valuation models (AVMs) can produce viable valuations for fine art and collectibles, and answers from an unusual position: we tried to build one.

AVMs require three conditions simultaneously: asset homogeneity, high transaction frequency, and densely digitized transaction records. The three combine into a single governing metric this paper calls comparable transaction velocity, the rate at which arm's-length, machine-readable transactions occur within a given asset's comparable set. Residential real estate, the one asset class where AVMs achieved scale, clears the threshold for marketing purposes and still fails it for transaction purposes: Zillow's estimates were good enough to display and disastrous to transact on, and its home-flipping business lost $881 million in 2021 proving the difference. Commercial real estate sits below the threshold entirely, which is why institutional CRE valuation remains appraisal-based to this day. Fine art sits far beyond CRE on every dimension: unique objects, holding periods measured in decades, sales that commonly take a year or two to realize, and a market whose decisive information is undigitized and, in important respects, deliberately never digitized.

Professional standards and tax law independently disqualify unassisted model output. USPAP's Advisory Opinion 18 holds that an AVM's output is not, by itself, an appraisal, because an appraisal is an opinion of value and only individuals can form opinions. The RICS Red Book provides that a model's output becomes a written valuation only after a valuer applies professional judgment to it, the International Valuation Standards state the same rule worldwide, and US Treasury regulations require that a qualified appraiser be a person. Behind all of these sits a harder

requirement no model can meet: accountability. An algorithm cannot explain its weighting to an IRS examiner, testify in court, or bear professional liability for its conclusion.

Title Collections was born out of the Lloyd's of London accelerator, and we build valuation infrastructure for the professionals who carry this risk: insurers, wealth managers, fiduciaries, counsel, and the appraisers they all depend on. We conducted a research program to design an AVM for fine art assets, with methodology review by the Institut intelligence et données at Université Laval. Mapping the model's data prerequisites against what the art market can actually supply led us to an architectural conclusion rather than a modeling one: algorithmic estimates in this asset class are viable as revaluation triggers, and only a qualified appraisal can serve as the arbiter of value. We call the resulting architecture Model-as-Trigger, Appraisal-as-Arbiter, and this paper presents both the evidence that compelled it and the infrastructure we built around it.

2. The problem: financial decisions rest on trusted valuations

The trust problem is the same across every profession that touches these assets; only the moment of reckoning differs.

For insurers, valuation is the pricing input, the aggregation input, and the claims benchmark simultaneously, and it binds both parties: the insured pays premium on it, and the carrier settles against it. Overstated values produce moral hazard and disputed claims at the moment of loss; understated values produce underinsurance discovered exactly when the relationship is tested.

For wealth managers and family offices, collections increasingly sit inside net-worth statements, allocation decisions, and lending collateral. A stale scheduled value misstates the balance sheet it feeds, and a bank advancing against art requires a current, defensible appraisal before a dollar moves.

For attorneys handling asset separation, in estate settlements and in divorce, the valuation is the settlement. Estate tax is assessed on appraised values under a nine-month clock, and in a matrimonial division where one party retains the art at an agreed figure, every dollar of valuation error transfers wealth between the parties silently. Disputed values mean dueling experts; defensible values mean documents.

Staleness is the ambient condition behind all three. In practice, schedules and statements typically carry objects at their purchase price, and some have been running on those figures for a decade or more while the market moved. The stakes are not hypothetical: when valuations are eventually tested by the most scrutinized review process in the market, nearly half do not survive intact, a record Section 5 examines in detail.

The temptation to solve the scaling problem with an AVM is understandable. The traditional appraisal process is time-consuming, expensive, and logistically heavy, while AVMs promise portfolio-scale revaluation at negligible marginal cost, a promise partially delivered in residential mortgage lending. The question this paper addresses is whether that success transfers to unique assets. The statistical answer, the regulatory answer, and the accountability answer are all no. We know because we went looking for yes.

3. What AVMs are and where they provide reliable valuations

3.1 Three approaches to value, and what survives contact with art

Valuation practice across asset classes rests on three canonical approaches, combined by valuators of every specialty to establish a credible, defensible opinion of value:

- **The market approach:** value inferred from comparable transactions and precedent sales.
- **The income approach:** value derived from the cash flows the asset generates, discounted to present.

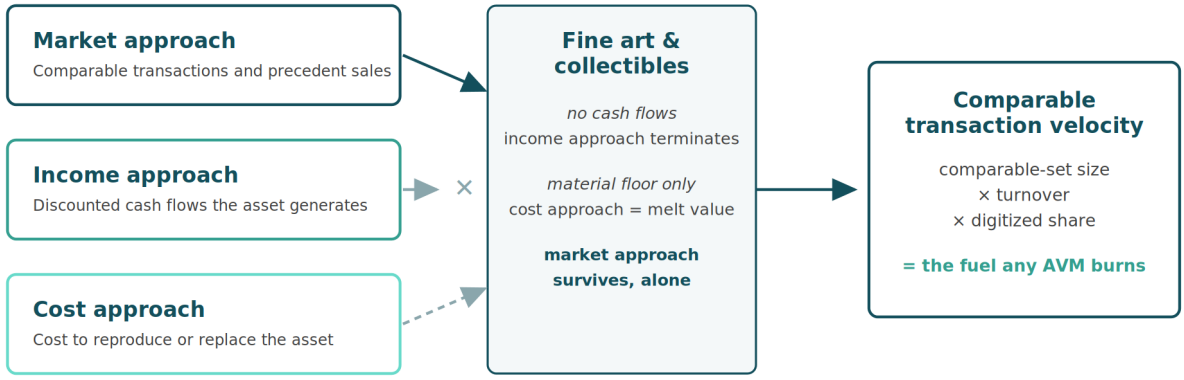
- **The cost approach:** value derived from the cost to reproduce or replace the asset.

Fine art and collectibles strip this toolkit down. Artworks generate no cash flows, so the income approach has nothing to discount. The cost approach survives only as a floor for objects whose materials carry independent commodity value: bronze, gold, diamonds in watches and jewelry. The floor is real and occasionally decisive, and it is also the least interesting number in the room, because the gap between material value and market value is precisely where the art lives.

> **The cost approach, executed with an angle grinder.** Maurizio Cattelan's *America*, a fully functional 18-karat gold toilet weighing approximately 98 kilograms, was stolen from Blenheim Palace on September 14, 2019, in a raid that took five and a half minutes. Police believe it was broken up and melted within hours; none of the gold was recovered, and two men were convicted and sentenced in 2025. The work was insured for $6 million as art; its gold was worth about £2.8 million as metal (The Art Newspaper, June 16, 2025). The theft was a forced liquidation at the cost approach: the market premium that made *America* an artwork was destroyed in the melt. Every AVM that prices art from material and dimensional features alone repeats the thieves' method, politely.

What remains for art is the market approach, alone. An automated valuation model is the market approach industrialized: a statistical attempt to predict what the market would bid for an asset, in the absence of any bid, combining hedonic regression, repeat-sales inference, algorithmic comparable selection, and machine-learning ensembles trained on large transaction datasets. Art-market variants substitute inferred reputation features, artist trajectory, exhibition history, market signals, for the square footage and lot sizes of the residential models. All of these techniques share the same fuel: a large volume of arm's-length transactions of assets similar enough that encodable characteristics explain most of the price variance, recorded in machine-readable form. The market approach starves without comparables, and an AVM is the market approach with an industrial appetite.

EXHIBIT 1
Three approaches to value, and what survives contact with fine art



Source: Core valuation approaches: International Valuation Standards (IVS 105). Application to fine art and collectibles: Title Collections analysis.

3.2 The conditions for reliable AVM output

AVMs provide reliable and consistent valuations when three conditions hold simultaneously:

1. **Homogeneity.** Units in the population share a common characteristics vector. Two 3-bedroom houses on the same street are imperfect substitutes, but they are substitutes. Homogeneity sets the size of the comparable set.
2. **Transaction density.** Enough recent, observable, arm's-length sales exist within each segment to fit and continuously recalibrate the model. Density is turnover multiplied by digitized observability: a sale that is never recorded in a machine-readable dataset does not feed the model.
3. **Encodable characteristics.** The attributes that drive price must be expressible as model inputs, whether observed directly (square footage) or inferred by an upstream model (an artist-desirability score estimated over time). Inferred features are legitimate and widely used. They are also not free: the inference is itself a model, so an inferred reputation score needs its own dense history of signals, and inference relocates the comparable-velocity constraint without removing it.

> **DEFINITION**
>
> **Comparable transaction velocity.** The rate at which arm's-length, digitized transactions occur within a given asset's comparable set. It is the product of the three conditions above: how many near-substitutes exist, how often they trade, and what fraction of those trades a model can see. Comparable transaction velocity, and never per-asset turnover, is the variable that determines AVM viability.

The distinction matters because per-asset turnover is misleading. The typical US homeowner now stays put for 12 years, up from 6.5 years in 2005 (Redfin, county-record analysis, March 2026), a slower per-asset rate than institutional commercial real estate at 8 to 11 years (Gau and Wang, 1994; Fisher and Young, 2000). Judged by per-asset turnover alone, residential AVMs should perform worse than CRE AVMs. The opposite is true, and comparable velocity explains why: a tract home in a large metro belongs to a comparable set of thousands of near-substitutes, several of which trade every week, all publicly recorded. Even in the slowest US housing market in three decades, roughly four million existing homes sell each year (NAR, June 2026), captured through mandatory public deed recording and the multiple listing services. The individual house almost never trades and the model never notices, because the comparable set trades constantly on its behalf. Statistical strength is borrowed from the population, and that borrowing is available only where a population exists.

This paper's central claim is falsifiable, and stating the test strengthens it: if comparable transaction velocity within a segment rises above the threshold at which models can be fit and policed at the error tolerance the decision requires, AVM viability returns. Editioned prints and multiples demonstrate exactly this at the margin of the art market, and Section 8 describes the adjacent segment where we tested it directly.

3.3 The comparable-velocity ladder: residential, commercial, art

The strongest evidence that comparable velocity is a hard constraint comes from commercial real estate, because CRE is the controlled experiment. CRE assets are far more standardized than art: square footage, leases, net operating income, cap rates, all quantifiable. If the primary failure reason for AVMs were exotic asset characteristics, one would expect them to perform well in CRE. They do not, and the reason is comparable velocity: the assets do not substitute for one another, and a downtown office tower's comparable set holds perhaps a dozen loosely similar properties, one of which trades in a good year, often confidentially. The academic literature is explicit that this breaks statistical valuation: commercial property markets are characterized by heterogeneous assets and typically too few transactions for statistical approaches to be viable, with failure modes including data lag bias, availability bias, mechanical bias in thin submarkets, and inability to emulate inspection-dependent judgment (Oxford FoRE, 2022).

The institutional market's response is the most instructive fact in this section. Rather than persist with transaction inference, the industry institutionalized the human appraisal: the NCREIF Property Index, the benchmark for US institutional CRE performance, is appraisal-based, with properties externally appraised at least annually precisely because there are not enough transactions to mark the portfolio to market; the live econometric debate concerns how to de-lag the appraisals everyone depends on (Clayton, Geltner and Hamilton, 2001), and nobody proposes replacing them. Even the human gold standard is strained at this frequency: appraisals deviate on average more than 12% from sale prices realized two quarters later (Cannon and Cole, 2011, on 25 years of NCREIF sales). That is the ceiling achieved by qualified professionals with property-level access in an asset class with quantifiable income streams. A model with strictly less information starts below it.

The pattern is confirmed by the people who allocate capital to solving it. In a series of structured interviews Title Collections conducted with investment professionals at proptech-focused venture capital firms, private real estate investment managers, and North American public pension investors, groups that combined manage a quarter of a trillion dollars in real assets, a consistent picture emerged: the most sophisticated institutional investors direct their technology investment at the friction in information sharing, dissemination, and control, and they retain the appraiser at the center of the valuation process.

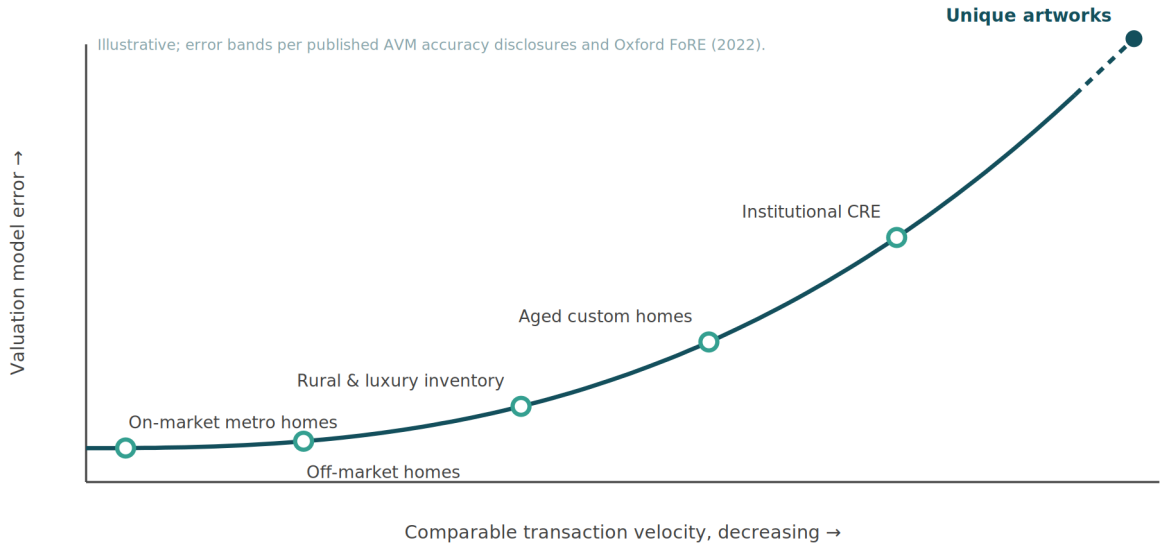
Exhibit 2. The comparable-velocity ladder.

	Per-asset turnover	Comparable set	Comparable velocity	Time to liquidate	AVM outcome
Residential (metro)	~12 yrs (6.5 in 2005)	Thousands of near-substitutes	Thousands of digitized comps per segment per year	Weeks to months	Display-grade viability; transaction-grade failure (see case study)
Institutional CRE	~8 to 11 yrs	A dozen loosely similar assets	A handful of trades per segment per year, often confidential	Months	No institutional acceptance; appraisal-based by necessity
Fine art, unique works	Decades	One	Approximately zero	Commonly 12 to 24 months	Below the chart

Sources: Redfin (2026); Gau and Wang (1994); Fisher and Young (2000); Oxford FoRE (2022); Title Collections analysis.

EXHIBIT 3

Model error rises as comparable transaction velocity thins (illustrative)



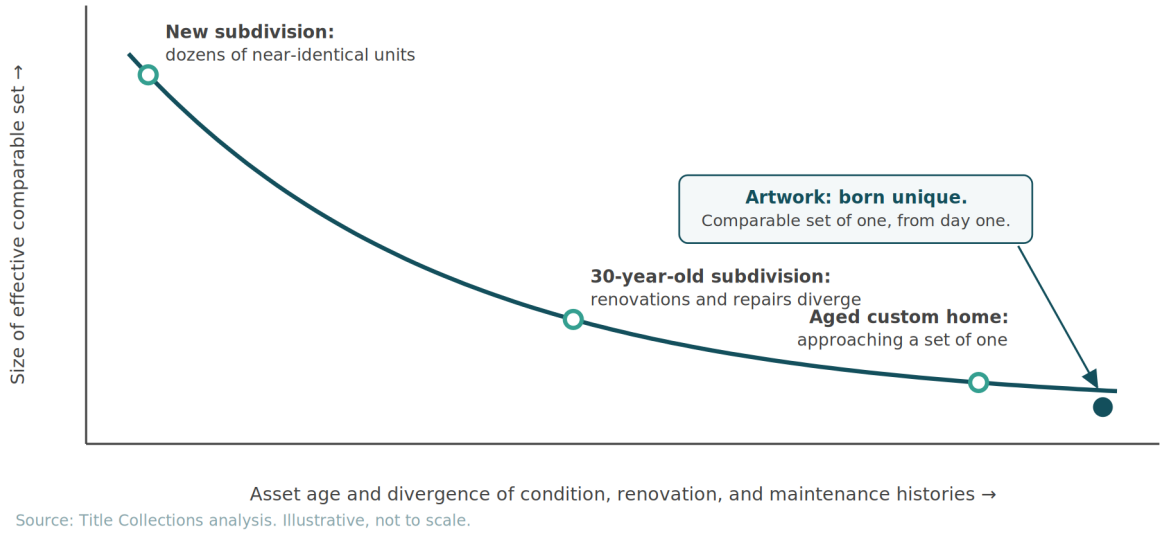
Source: Published AVM accuracy bands; Oxford FoRE (2022); Title Collections analysis. Illustrative, not to scale.

Comparable velocity also decays inside the residential market itself. A freshly built subdivision is the best case an AVM will ever see: dozens of near-identical units differing by lot position and finish package. Then time goes to work: one owner renovates the kitchen, another defers the roof, a third finishes the basement. Thirty years on, divergent maintenance, renovation, and structural histories have made each house progressively unique, and the comparable set that once numbered in the dozens has thinned toward one. Published AVM error rates trace the decay: errors widen for older, rural, and luxury inventory, the segments where divergence has run longest or homogeneity never existed.

A house is born a substitute and becomes unique over decades. An artwork is born unique. It starts where the oldest, most idiosyncratic house ends up, with a comparable set of one.

EXHIBIT 4

Uniqueness accumulates with age; artworks are born unique (illustrative)



Source: Title Collections analysis. Illustrative, not to scale.

3.4 Case study: what happens when a display-grade estimate is asked to transact

CASE STUDY

Zillow Offers: when a display-grade estimate was asked to transact

Zillow Offers was an iBuying operation: Zillow used its valuation models, including the consumer-facing Zestimate, to make cash offers directly to homeowners, take the properties onto its own balance sheet, and resell them. In February 2021, Zillow made the Zestimate itself the live cash offer for eligible homes: the company was market-making residential real estate on the strength of its AVM.

What happened. Zillow purchased roughly 7,000 homes across 25 metropolitan areas at prices its models generated. As markets cooled in 2021, the models kept overvaluing acquisitions. In November 2021 Zillow shut the business down. The home-flipping operation lost $881 million in 2021, driving a consolidated net loss of $528 million for the otherwise profitable company (Parker, The Wall Street Journal, February 10, 2022), and the workforce was cut by roughly 25%. CEO Rich Barton's stated conclusion: the unpredictability in forecasting home prices far exceeded what the company had anticipated.

Why it cratered.

This happened in the asset class with the highest comparable transaction velocity on earth. Fine art has approximately none.

1. **Display-grade accuracy differs from transaction-grade accuracy.** A median error tolerable in a marketing widget is fatal when you transact at the estimate thousands of times on thin margins. The error need only exceed the margin.
2. **Adverse selection concentrates the errors.** Homeowners accept an algorithmic offer most eagerly when the algorithm has overpriced their home. The inventory fills disproportionately with the model's worst mistakes.
3. **Models trained on one regime fail at the turn.** The models forecast resale values months forward through moving interest rates, the unwinding of zero-rate policy, and repricing construction labor. A valuation model with economic forecasts embedded in it inherits their errors and adds its own; when the regime changed, the model was structurally last to know. Art markets are more regime-driven than housing, with taste, fashion, and single-artist markets turning sharply.
4. **Human overrides amplified rather than corrected.** Industry reporting indicates Zillow tweaked its algorithms to win more offers in a hot market, overpaying just as prices began to cool. A model whose output is negotiable in one direction serves as a justification, and no longer as a control.

Regulators reached a parallel conclusion. Six US federal agencies, the OCC, the Federal Reserve, the FDIC, the NCUA, the CFPB, and the FHFA, jointly adopted quality-control rules for AVMs in mortgage lending, in effect since October 1, 2025 (89 Fed. Reg. 64538), mandating five standards: confidence in the estimates, protection against data manipulation, avoidance of conflicts of interest, random sample testing, and nondiscrimination compliance. The regulators' checklist for residential AVMs is a list of the ways the best-known deployment failed.

4. Why fine art sits outside of every threshold

Fine art and collectibles fail the comparable-velocity test categorically, and the finding carries a forty-year academic pedigree. Baumol (1986), analyzing three centuries of auction data, concluded that art prices are structurally unanchored: a unique artwork has no equilibrium price for a model to converge toward, and if prediction is a losing game in stock prices, it is certainly unlikely to be a winner in the market for works of art. The tools have improved since 1986; the constraints have not moved. It is like upgrading from a bow to a rifle, but still shooting arrows.

4.1 Uniqueness: comparable sets of one

A painting is not a unit drawn from a population of near-substitutes. It is, in the limit, its own population. Even within a single artist's oeuvre, price variance is driven by period, subject matter, palette, provenance, exhibition history, condition, literature, and freshness to market, exactly the characteristics a hedonic model struggles to encode. Two works by the same artist, same year, same dimensions, same medium can transact an order of magnitude apart.

The boundary of the claim deserves honesty, because honesty sharpens it. The luxury goods sector, where objects share technical specifications and reference numbers, is workable terrain for algorithmic comparison; works of art are inherently unique. Editioned and quasi-unique material sits closer to homogeneity, and routine valuation of lower-value, high-homogeneity objects may eventually be handled algorithmically with appropriate disclosure. That corner of the market is real. It is also the corner where schedule risk does not concentrate: the concentration risk in any collection sits in exactly the objects that are one of one.

4.2 Illiquidity: the market that takes a year to answer

Where a metro house sells in weeks and an institutional building in months, a significant artwork commonly takes 12 to 24 months to sell at fair value, in Title Collections' experience and that of the vast number of art dealers and appraisers we discuss the topic with. Illiquidity of this depth compounds against any model three ways. The quick-sale discount is severe: compressing the timeline means accepting a materially lower price, which is why the estate-planning literature treats art as a canonical liquidity problem, with US estate tax of up to 40% typically due within nine months of death and life insurance, deferral, and lending strategies existing specifically to bridge the gap while assets are sold properly (Erskine; Charles Schwab, 2026; Comerica, 2026). The estate clock runs out before the market answers. Transaction costs are high and lumpy, so realized net prices scatter around any modeled gross. And timing is unpredictable: "value as of a date" for an asset that needs a year of marketing is a different epistemic object than a house price.

Repeat sales of the same work are correspondingly rare and separated by decades, and the scarcity has been measured: Baumol's dataset covering 1652 to 1961 contained only 640 repeat-sale price pairs, and Anderson's index for 1715 to 1986 contained 3,329 (Mei and Moses, 2002, reviewing the literature). Three centuries of the Western art market yielded a few thousand repeat-sale observations; the US residential market records more comparable transactions than that before lunch on an average Tuesday. Auction data adds selection bias, buy-ins are handled inconsistently, and the middle of the market is thinning: sales of works priced $50,000 to $250,000, the band where most scheduled collections live, are down 29% since 2010 (McAndrew, 2026).

4.3 Sparse digitization: the market the models cannot see

Even if artworks were homogenous and traded frequently, the data to train an AVM would still not exist, because the fine art and collectibles market is one of the least digitized major asset markets in the world.

Exhibit 5. The data infrastructure gap.

Data layer	US residential real estate	Fine art and collectibles
Transaction recording	Mandatory public deed and price recording at the county level	No recording obligation anywhere; auction results public, dealer and private sales confidential by design
Asset registry	Parcel numbers, assessor rolls, title chains	Standards exist (Getty Object ID, CONA, CIDOC-CRM, Spectrum) but adoption is voluntary and coverage reaches a small fraction of objects in private hands
Attribute schema	Standardized MLS fields	No enforced schema; condition, provenance, and exhibition history live in PDFs, letters, and institutional files
Coverage	Near-total	A fraction of the market, skewed toward auction

Five consequences follow:

1. **The visible market is a minority of the actual market.** In 2025, global art sales reached an estimated $59.6 billion: dealers accounted for $34.8 billion (58%), auction houses' own private sales for $4.2 billion, and public auctions, the only segment that produces the data AVMs train on, for $20.7 billion (McAndrew, 2026). Roughly two-thirds of the market by value transacts with no public price record, and practitioners corroborate: private prices are unreported or disclosed only through trusted relationships (Czudej, 2026). A model trained on auction records is a model of the consigned-to-auction market, which differs systematically from the market as a whole.
2. **Even the public data is reconstructed, not recorded.** There is no art-market equivalent of a deed registry or an MLS feed. The commercial aggregators rebuild the transaction record after the fact from published catalogues and results, with coverage floors and completeness that varies with each house's cooperation. Conversations with former employees in the field suggest the operational reality has long been more artisanal than the automated image, with results historically gathered through phone conversations and manual reconciliation, though practices may have modernized recently. And the source material is eroding: as houses shift to born-digital catalogues on ephemeral platforms, many online-only sales no longer produce a formal catalogue at all. The most digitized corner of the art market is becoming harder to archive, not easier.
3. **The records that exist are not machine-readable.** Provenance chains, condition reports, conservation history, and literature citations are documents. Digitizing them is object-by-object work requiring expertise.
4. **The decisive information is structurally withheld.** In the art market, information asymmetry is a competitive advantage: the participants who hold the most price-relevant knowledge monetize precisely the fact that it is undisclosed. No dataset of it exists because no one who has it has an incentive to create one. A market can lag on digitization by inertia; this market withholds by design.
5. **Prices do not travel.** An artist's recognition, and therefore price level, differs across countries; sales data from one national market produces suggestions unachievable in another. For global collections, an AVM trained on one geography's exhaust misprices the same object as it crosses borders.

For the majority of objects a professional schedules, an AVM would be estimating from no relevant data while presenting the output with the same confidence either way.

4.4 Value drivers that resist encoding

The characteristics that most move price in art are the ones that resist structured capture. Condition: real estate prices it through a standardized, licensed inspection trade in percentage adjustments, while art condition has no standard schema, interacts with attribution and conservation history, and moves price discontinuously. Provenance and title can

add or destroy value in step functions; a restitution claim is not a percentage adjustment. Attribution is binary and catastrophic: “attributed to” versus “by” is a difference between price levels. Authenticity, literature, exhibition history, and freshness to market price in through expert judgment. Some of these can be approximated by inferred features, per Section 3.2, and the inference inherits the data problem. An AVM does not degrade gracefully on these inputs. It is blind to them, and it returns a confident number anyway. A wrong number with no error flag is worse than no number.

4.5 Recency, the tails, and where model confidence actually lives

Model reliability in art tracks recency of comparables more than fame. A blue-chip work by an artist whose market cleared multiple auction sales in the past three years sits in the densest data neighborhood art offers, and a model’s estimate there is at its most defensible. The same artist’s work last priced twenty years ago sits in the dark: the intervening decades contain taste shifts and single-artist booms no cluster average reliably bridges. Fame concentrates data, and time disperses it.

Two structural problems remain even in the data-rich neighborhoods. Statistical estimators compress toward the center of their training distribution, so the exceptional works within an artist’s market are systematically underestimated even when the artist is data-rich, and those works are where concentration risk lives. And the returns literature shows the top of the market obeys dynamics the middle does not: expensive works systematically underperform the art-market index (Mei and Moses, 2002), so projecting population-level appreciation onto top-end works overstates their drift. The objects a professional most needs to get right are the objects where the model’s confidence is least earned.

5. The regulatory and accountability floor

Suppose, counterfactually, that an AVM could estimate a Basquiat within a tolerable band. It still would not produce an official value, for three compounding reasons.

EXHIBIT 6
The regulatory and accountability floor: three layers, one requirement

Each layer independently disqualifies an unassisted model output; together they leave no arbitrage.

Source: USPAP; RICS; IVS; Treasury regulations; case law as cited in Section 5. Title Collections analysis.

5.1 Professional standards require a professional

The standards hierarchy draws the same line at every level. USPAP addressed AVMs directly in Advisory Opinion 18, adopted in 1997 and still in force: an AVM's output is not, by itself, an appraisal, because an appraisal is an opinion of value, which only an individual exercising judgment can form; AO-18 applies to real, personal, and intangible property, so it covers art and collectibles on its face. Advisory Opinion 37 (2018) extends the principle to computer-assisted tools generally, and a proposed Advisory Opinion 41, in public comment as of 2026, consolidates the guidance for AVMs, statistical software, and AI, on the same principle: technology can assist analysis and cannot replace the appraiser's independent judgment or responsibility. The RICS Red Book Global Standards, effective January 31, 2025, provide in mandatory standard PS 1 that a valuation based wholly or partly on the output of an AVM or AI-assisted model is regarded as a written valuation only where it has been subject to the additional application of professional judgment by a valuer. And the International Valuation Standards state the rule for the world: per IVS 105, no model without the valuer applying professional judgment can produce a compliant valuation.

5.2 Tax and legal frameworks require a person

- A qualified appraisal must be conducted by a qualified appraiser, and Treasury Regulation §1.170A-13(c)(3) requires that the qualified appraiser be a person. Charitable contributions of property above $5,000 require such an appraisal (IRC §170(f)(11)).
- The review record shows why the requirement has teeth, and why a professional who can support the valuation in a legal and tax argument is the asset the process actually demands. The IRS Art Advisory Panel, which reviews audited returns containing appraisals of art valued at $50,000 or more, reviewed 195 items in fiscal 2023 with an aggregate claimed value of $795.5 million. It accepted 53% and adjusted 47%, in both directions: 32% of items revised downward, 15% upward, with gross movements exceeding $120 million (IRS Publication 5392). The two directions largely net out for the government; they never net out for the individual estate, donor, or claim.
- Courts enforce methodology and independence. In Estate of Kollsman v. Commissioner, No. 18-70565 (9th Cir. June 21, 2019), affirming T.C. Memo. 2017-40, the courts sustained a $781,488 estate tax deficiency where the estate's valuation, prepared by a premier auction house specialist, included no comparable sales data, and where the expert's simultaneous solicitation of the works' consignment gave him a direct financial incentive toward lowball estimates. The bar is a defensible, documented, methodologically sound valuation by an independent professional. A credentialed human with weak process fails it. An algorithm cannot even take the test.
- The same trail is presumed wherever values have legal effect: agreed-value insurance provisions and proof-of-loss processes at claim time, collateral files in art-secured lending, and the appraisals exchanged in matrimonial division, where courts expect qualified experts whose methodology survives cross-examination.

US AND CANADA

The architecture of the requirement is the same on both sides of the border. In the United States, the qualified-appraiser regime described above governs charitable, estate, and gift valuations. In Canada, the Canada Revenue Agency expects fair market value determinations for donations and deemed dispositions to be supported by qualified, well-documented appraisals, and gifts of certified cultural property flow through the Canadian Cultural Property Export Review Board, which reviews the fair market value determinations underlying tax certificates. In both jurisdictions, a person with a defensible methodology stands behind every value that has legal effect.

5.3 The accountability gap: an algorithm cannot be deposed

Behind the standards and the statutes sits the requirement they both encode: someone must stand behind the number. Czudej (2026), writing as a practicing appraiser and board member of the Appraisers Association of America, states the problem in its sharpest form. An algorithm cannot explain why it weighted one comparable over another. It

cannot respond to an IRS examiner's challenge to its methodology. It cannot testify in court. It cannot bear professional liability for a conclusion that exposes an estate to penalties, an insurer to a bad-faith claim, or a divorce settlement to reopening.

And when two algorithmic tools produce different valuations for the same work, no mechanism exists for deciding which is correct, because neither can give an account of itself. The question lands at the worst possible moment, at claim time or at settlement: whose number governs, and who stands behind it? A value with no accountable author is a liability whose owner has not yet been determined.

The three floors compound with the statistics of Sections 3 and 4: the assets where AVMs fail statistically are the same assets where standards, tax law, and accountability demand human expertise. There is no arbitrage between them.

6. How to evaluate any automated valuation claim for art and collectibles

Professionals in every seat will keep receiving valuation-technology pitches. Three questions triage all of them, and the evidence in this paper supplies the grading key:

1. **What is the comparable transaction velocity of the segment the model was trained on, and of the objects it will be asked to price?** A model demonstrated on editioned multiples or heavily traded luxury goods has proven nothing about unique works. Ask for the comparable-set size and the count of digitized transactions per year behind the objects that carry your concentration risk.
2. **Who bears professional accountability for the output?** If the answer is no one, the output cannot be a valuation for any purpose that matters: USPAP, the RICS Red Book, IVS 105, and Treasury regulations each disqualify it independently. If the answer is an appraiser who reviews the output, ask what their scope of work adds beyond a signature, because Kollsman shows that a signature over weak methodology fails in court.
3. **What happens when the model is wrong in the tail?** Every model is most confident near its training mean and least reliable on stale, exceptional, or record-setting objects, which is where concentration risk lives. Ask for the error distribution, never the average, and ask what the system does when uncertainty is high. A system that escalates to a qualified human has an answer. A system that returns a number anyway is the Zillow architecture with a smaller balance sheet.

Any tool that survives all three questions will look, structurally, like the architecture described in Section 8, because these questions are the design constraints the market imposes.

7. How value is actually discovered: the market-maker network

If the data does not determine the price, what does? In thin markets, price discovery is a network. The people who know what a work would bring are the people who would make the market for it: the two or three dealers who handle the artist, the specialist who ran the last comparable consignment, the collectors who underbid it. An appraisal, properly done, is a structured survey of that network combined with documented comparables and the appraiser's own market participation.

This claim has peer-reviewed support. Fraiberger et al. (2018), analyzing the exhibition histories of nearly half a million artists, found that artistic reputation and market success are driven by position within a network of institutions rather than by any measurable property of the work itself. The single most price-relevant variable in art is a network variable. Networks can be studied; they cannot be scraped into a hedonic feature.

The United States government's own valuation-review mechanism operates on the same premise: the IRS explains that its Art Advisory Panel of dealers, scholars, and curators exists because the panelists provide knowledge of private sales drawn from personal experience, information that cannot be obtained effectively from within the IRS (IRS Publication 5392). When the tax authority of the world's largest art market needs a value checked, it convenes the market-maker network. Practitioners describe the same mechanism from inside: Czudej (2026) recounts price variation across seemingly identical works by the same artist, resolved only by a specialist dealer's knowledge of a physical property assessable in person, worth a significant share of the price, and shared through a relationship built over years. Our own research reached the identical conclusion from the modeling direction: mapping what an appraiser actually contributes yielded non-public pricing information obtained through gallery relationships, and comparable selection based on exhibition correlation, style, and reputation within the relevant buyer categories. The appraiser's method contains the market. The dataset is its lagging, partial exhaust.

The implications: digitized auction records are the exhaust of the network, the relevant expertise is unevenly distributed and highly specific (the right appraiser for an Abstract Expressionist canvas is the wrong appraiser for a Patek reference or a Ming bronze), and therefore the scarce resource is matched expertise. The scaling problem is a routing and logistics problem. That reframing is the hinge of this paper, and Section 8 describes what we built when we accepted it.

8. What we learned building one: Model-as-Trigger, Appraisal-as-Arbiter

8.1 The research program

Title Collections did not arrive at this paper's conclusion from the sidelines. We conducted a research program to design an automated valuation model for fine art assets: a full architecture covering standardized artwork representation, reference clusters of comparable works, cluster price momentum between valuation events, and uncertainty estimation. The program's machine-learning methodology was developed and reviewed under a mandate with the Institut intelligence et données, Université Laval (NRC-IRAP mandate report, March 13th, 2026).

Designing the model forced the question this paper has been answering: what data would it need? The prerequisites were specific: a persistent object identity for every artwork; a populated schema of price-relevant attributes; a sufficient density of observed pricing events within each reference cluster to estimate momentum and, critically, uncertainty bands around it. Sections 3 through 7 describe what we found when we mapped those prerequisites against what the market supplies. The requirements for a viable AVM output on unique works cannot be satisfied by the art market's current data infrastructure, and the most decisive inputs are structurally withheld.

SCOPE AND LIMITATIONS

Automated valuation for art is an important area of research, and we support everyone pursuing it: the academic hedonic and repeat-sales literature, the art-index projects, and the commercial data firms have invested years where our program invested months. We state our own scope plainly. Our fine art statistical results were inconclusive in the direction that matters: the data infrastructure of the unique-works market cannot support decision-grade point estimates, for the reasons documented in Section 4. To test whether that conclusion reflected our pipeline or the asset class, we ran the same methodology on luxury watches, a far more liquid and densely digitized adjacent category, and reached a similar conclusion: point estimates reached useful precision for ranking and screening, while market-time effects and decision-grade confidence intervals remained out of reach, the same failure mode that matters for valuation. Despite decades of collective effort across far better-resourced projects, a valuation system that a practitioner could rely on for decision-grade values on unique works does not yet exist. Our fine art conclusions rest on the data-infrastructure and market-structure analysis of Sections 3 through 7; our architectural conclusion rests on both.

> *"AVMs are only interesting for uninteresting assets."*
>
> — Pascal Kropf, PhD, concluding Title Collections' AVM research program

8.2 The architectural conclusion

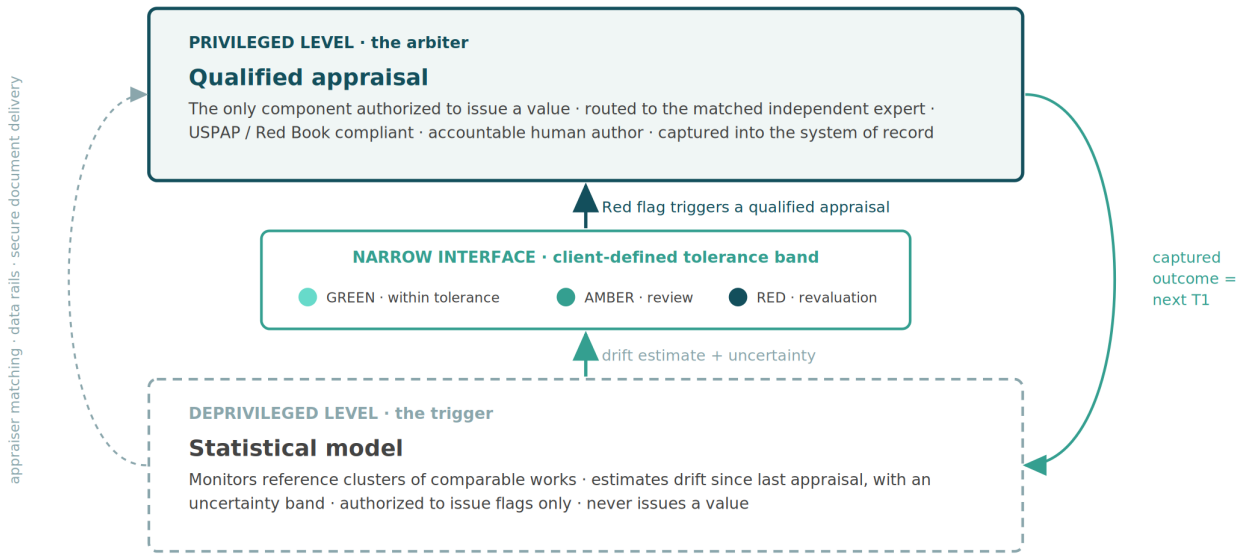
The productive output of the research was not a model. It was an architecture. If algorithmic estimates in this asset class cannot be valuations, what can they responsibly be? Our design converged on a specific answer: **the model is a trigger, and the appraisal is the arbiter.**

DEFINITION

Model-as-Trigger, Appraisal-as-Arbiter. A hybrid valuation architecture in which a statistical model monitors reference clusters of comparable works and estimates directional drift in an object's value since its last qualified appraisal, with an uncertainty band, and in which every determination of value is made by a qualified appraisal, routed, executed, and captured by the surrounding infrastructure.

The system tracks reference clusters and estimates the drift in an object's value since its last qualified appraisal, with an uncertainty band. It then does the one thing a statistical estimate is actually qualified to do: classify. Green, the last appraised value likely remains within tolerance, no action. Amber, possibly outside tolerance, review recommended. Red, likely outside tolerance, revaluation needed. The tolerance band is defined by the client, and never by the model. A Red flag triggers a qualified human appraisal, routed, executed, and captured.

EXHIBIT 7
Model-as-Trigger, Appraisal-as-Arbiter: separation of authority



Source: Title Collections.

The division of labor follows each party's actual competence: statistical models are genuinely good at detecting that something has probably changed, and in this asset class they are structurally unfit to say what the new value is; appraisers are the only actors who can produce a defensible value, and their scarce time should be spent where drift is

probable. Every objection documented in this paper is answered by the placement of the decision. The compression and recency problems of Section 4.5 do not threaten a trigger, because a flag needs to detect drift, and never to price the tail. The accountability requirement of Section 5.3 is satisfied, because every value has a qualified human author. The standards requirements of Section 5.1 are satisfied by construction, because the model never issues a valuation at all. This is also where the appraisal profession itself has landed: practitioners writing on AI converge on tools that supplement the appraiser's judgment (Czudej, 2026; Smith, 2024), with physical examination remaining vital and AI-assisted appraisers producing more rigorous documentation and more defensible opinions of value.

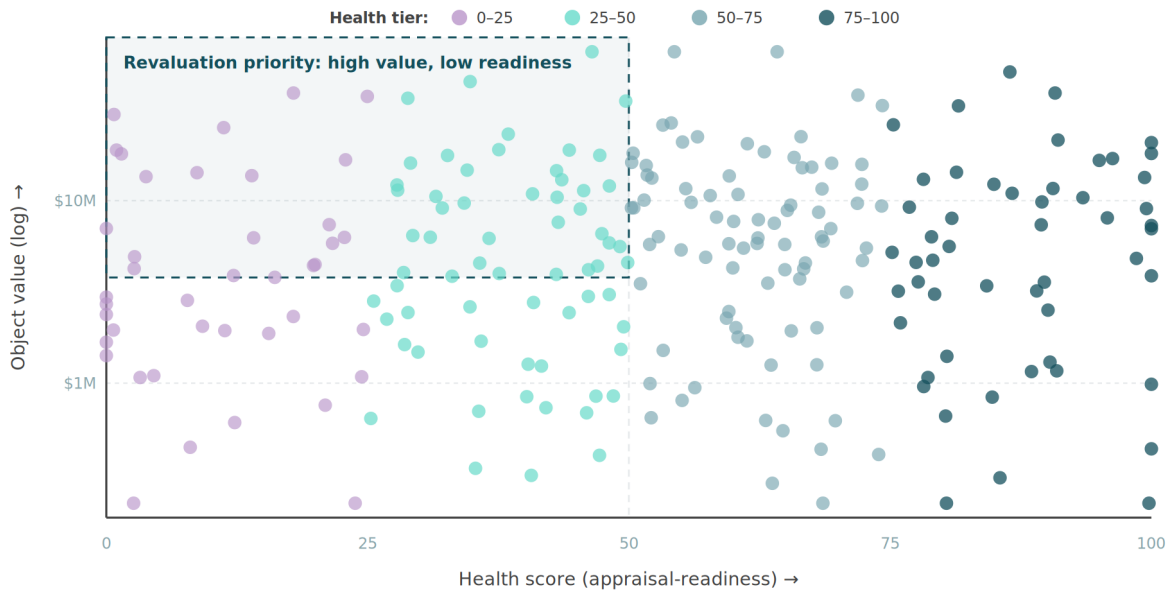
8.3 The infrastructure around the arbiter

An architecture that routes to human appraisal inherits the original problem: appraisal is hard to scale. The friction sits in three places, and the Title Collections platform addresses each:

1. **Expert selection.** Finding the right specialist and matching them to the object class is slow and relationship-dependent. We built an AI-assisted appraiser recommendation system with a feedback loop, spanning the categories of fine art and collectibles our network covers. The AI routes to the expert. Routing to a matched, independent appraiser also removes the conflict Kollsman punished: the expert valuing the object has no stake in transacting it.
2. **Information logistics.** Assembling and securely transmitting the documentation an appraiser needs is a coordination burden that today happens over email, if at all. Our data rails deliver the required information to the selected expert all at once, structured and secured.
3. **Outcome capture.** The appraisal result too often lands as a PDF in a drawer, immediately going stale. Outcomes are captured into a collection management system of record, so values remain live, auditable, and connected to downstream uses, including the trigger model, which consumes each new appraisal as its next baseline.

Beneath all three runs the layer Section 4.3 showed the asset class lacks: an appraisal-readiness discipline. USPAP requires appraisers to maintain documented workfiles and support their conclusions with records; the constraint in this asset class is that the records rarely exist in usable form when the appraiser arrives. Title Collections operates a proprietary scoring system that establishes a health diagnostic of a portfolio of art and collectibles assets, scoring the data quality and compliance state of each object and generating an improvement plan that progressively brings it to appraisal-ready. The rubric is proprietary; its output is the portfolio view below, which doubles as a signal for underwriting, reporting, and settlement readiness alike: where value concentrates while readiness lags, revaluation priority lives.

EXHIBIT 8
Portfolio health tiering: value against appraisal-readiness (illustrative view)



Source: Title Collections platform, health diagnostic view. Illustrative data; scoring rubric proprietary.

One further consequence deserves attention. Independent appraisers, most of whom work in small practices, face a widening structural disadvantage as large institutions build bespoke AI tooling. Shared infrastructure of the kind described here levels that access: the specialist network that actually holds the market's knowledge gets the routing, data rails, and capture systems that no small practice could build alone. For every profession that depends on a distributed appraisal network, strengthening that network is the point.

8.4 What this means in practice: insurers, wealth managers, and counsel

Title Collections works with insurance professionals from the Lloyd's of London market and adjacent specialty markets, alongside the wealth managers and estate practitioners who steward the same collections. The architecture serves each profession where its risk actually lives.

For insurers. Fine art risk management runs on three questions: what is it, where is it, and how much is it. This paper is about the infrastructure behind the third question, and the architecture strengthens the other two as a byproduct, since a captured, current schedule is also an inventory and a location record. The market's current answer is a rational adaptation to price-discovery difficulty: most fine art policies are written on an agreed-value basis, both parties settling on an amount precisely because finding the right price is hard, with the appraisal anniversary as the nominal refresh standard, loosely enforced. None of this is negligence; it is the equilibrium a market reaches when accurate valuation is expensive relative to the premium. That ratio is the constraint any solution must respect: this is a competitive market with abundant capacity and low loss ratios, and premiums are small as a function of underlying asset value, so accuracy must arrive at a cost sized to the premium. Schedule-wide revaluation cycles fail that test; trigger-based revaluation passes it, because monitoring is cheap and the expensive resource is spent only where drift is probable. The deeper correction is about responsibility: accurate values are ultimately the policyholder's, and market dynamics have drifted underwriters into acting as the effective buyers of valuation. A strategy the policyholder can execute simply, a health diagnostic, a flag, a routed appraisal, returns the responsibility to where it sits, and every party benefits. The value gap, finally, is a structural exposure that concentrates at the worst moments. Consider a natural

catastrophe like the recent Los Angeles fires: art and collectibles are easily movable, so establishing where objects actually were at the moment of loss is genuinely hard, and an object scheduled above its market value creates, in that fog, an incentive problem no carrier can audit after the fact. The gap between scheduled and market value is itself the moral hazard, priced into the book whether or not anyone ever acts on it. Current values shrink the gap; a captured location record shrinks the fog. Both are outputs of the same infrastructure.

For wealth managers and family offices. A collection that sits inside a net-worth statement deserves the same reporting discipline as every other line: values with a refresh cadence driven by evidence rather than by the anniversary of a binder. The trigger model supplies exactly that cadence, flagging the objects whose reported values have probably drifted while leaving the rest alone, so the cost of accuracy scales with actual market movement. The same current, defensible values unlock art-secured lending without a scramble for fresh appraisals, and the health diagnostic gives advisors something the asset class has never had: a portfolio-level answer to "how good is our documentation," object by object, before a transaction, a loan, or a generational transfer forces the question.

For counsel in estate settlement and matrimonial division. In asset separation, the valuation is the settlement: estate tax is assessed on appraised values under a nine-month clock that runs faster than the 12-to-24-month sale timeline of the underlying assets, and in a division where one party retains the art at an agreed figure, every dollar of valuation error moves wealth between the parties. The Panel's 47% adjustment rate and Kollsman's $781,488 deficiency are what indefensible values cost when tested. The architecture's contribution to counsel is readiness before the triggering event: objects held at appraisal-ready, with provenance, condition, and documentation assembled and a current qualified appraisal on file or one flag away. The worst time to discover a gap in title or a missing condition report is mid-settlement, with the clock running and an adversarial expert across the table. A defensible trail converts disputes into documents.

Questions for the professionals. In the tradition of the London market's emerging-risk work, we close with the questions we believe every seat should be asking, with or without us:

- What share of the values in your schedules, statements, or files still carry the purchase price, and how old is the oldest figure?
- At claim time or at settlement, what fraction of the objects you rely on could produce a standards-compliant appraisal trail without remediation?
- If a division, a loan, or a natural catastrophe demanded defensible values on 90 days' notice, how many objects could deliver one, and could you establish where each object actually was?
- When a valuation-technology vendor presents error statistics, have you seen the tail of the distribution, or only its center?

9. Conclusion

AVMs are a triumph of statistics applied to homogenous, liquid, densely recorded markets, and they are only interesting for uninteresting assets. Commercial real estate already lacks the comparable transaction velocity they require; fine art trades more rarely still, is born unique, takes a year or two to sell, and conducts most of its business outside any dataset a model could see, in a network that withholds its knowledge by design. Standards, tax law, and the demands of accountability independently require a qualified human author behind every value an insurer schedules, a wealth manager reports, or a court accepts. We know these constraints are real because we designed against them: Title Collections built the architecture for a fine art AVM, mapped its prerequisites against the market, and concluded that the honest role for the model is the trigger, and the only possible arbiter is the appraisal. The scalable path to accurate, current, compliant values is better infrastructure around the appraiser.

About Title Collections

Title Collections is a financial services company applying institutional governance frameworks to fine art and collectibles. Born out of the Lloyd's of London accelerator, the company builds valuation infrastructure for insurers, advisors, fiduciaries, and collectors. Recognizing the governance prerogative of valuation in an institutional and regulatory environment, Title developed a network of qualified appraisers spanning over a hundred categories of fine art and collectibles, supported by its proprietary AI-assisted matching and routing system, secure data rails, and a collection management system of record. Its core mission is to increase asset governance for insurers, wealth managers, and institutions, leveraging its proprietary health diagnostic and gap remediation system that brings every object to its required state. Learn more at titlecollections.com

About the authors

George Fortin, MBA, is Managing Director and founder of Title Collections. He brings over twenty years of asset management experience, holds an MBA from University of Toronto, spent part of his career in real estate private equity at a major Canadian pension investment manager, and is an alumnus of the Lloyd's of London accelerator. During his career as a real estate private equity asset manager, he created and managed the fund's allocation to proptech and real assets innovation, where he investigated automated valuation models extensively.

Pascal Kropf, PhD, is Director of data science at Title Collections and led the AVM research program described in this white paper, including the machine-learning mandate reviewed by the Institut intelligence et données, Université Laval. He holds a PhD in Neuroscience from McGill's Neurological Institute-Hospital and a master's in physics from the University of Bern in Switzerland, where he worked on light-induced neural activation. Following his studies, Pascal led many teams in data science and machine learning in fields spanning computer vision to large-scale education.

References

Baumol, W.J. (1986). "Unnatural Value: Or Art Investment as Floating Crap Game." American Economic Review, 76(2), Papers and Proceedings, 10-14.

Cannon, S.E. and Cole, R.A. (2011). "How Accurate Are Commercial Real Estate Appraisals? Evidence from 25 Years of NCREIF Sales Data." Journal of Portfolio Management, 37(5), 68-88.

Charles Schwab (2026). "Should You Add Life Insurance to Your Estate Plan?" schwab.com.

Clayton, J., Geltner, D. and Hamilton, S.W. (2001). "Smoothing in Commercial Property Valuations: Evidence from Individual Appraisals." Real Estate Economics, 29(3), 337-360.

Comerica (2026). "Estate Liquidity Planning Strategies." comerica.com.

Czudej, T. (2026). "AI, Art Valuation and the Question of Accountability." WealthManagement.com, February 25, 2026.

Erskine, M. "Unlocking The Value: Art Appraisal For Gift And Estate Tax Matters." Family Wealth Report.

Estate of Kollsman v. Commissioner, No. 18-70565 (9th Cir. June 21, 2019) (mem.), aff'g T.C. Memo. 2017-40.

Fisher, J. and Young, M. (2000). "Holding Periods for Institutional Real Estate in the NCREIF Database." Real Estate Finance, Fall 2000.

Fraiberger, S.P., Sinatra, R., Resch, M., Riedl, C. and Barabási, A.-L. (2018). "Quantifying reputation and success in art." Science, 362(6416), 825-829.

Gau, G.W. and Wang, K. (1994). "The Tax-Induced Holding Periods of Real Estate Investors." Journal of Real Estate Finance and Economics, 8(1).

Institut intelligence et données, Université Laval. NRC-IRAP mandate report, March 13th, 2026.

Internal Revenue Service. Publication 5392, "The Art Advisory Panel of the Commissioner of Internal Revenue, Annual Summary Report for Fiscal Year 2023" (Rev. 6-2024).

McAndrew, C. (2026). The Art Basel and UBS Global Art Market Report 2026. Arts Economics.

Mei, J. and Moses, M. (2002). “Art as an Investment and the Underperformance of Masterpieces.” American Economic Review, 92(5), 1656-1668.

National Association of Realtors (2026). Existing-Home Sales series, June 2026.

Oxford Saïd Business School, Future of Real Estate Initiative (2022). “The Future of Automated Real Estate Valuations.”

Parker, W. (2022). “Zillow’s Shuttered Home-Flipping Business Lost $881 Million in 2021.” The Wall Street Journal, February 10, 2022.

Quality Control Standards for Automated Valuation Models, 89 Fed. Reg. 64538 (Aug. 7, 2024) (effective Oct. 1, 2025).

Redfin (2026). “The Typical U.S. Homeowner Hangs Onto Their House for 12 Years.” March 3, 2026.

RICS (2025). RICS Valuation – Global Standards (Red Book), effective January 31, 2025, PS 1; International Valuation Standards, IVS 105.

Smith, M. (2024). “The Impact of AI on Art and Antique Valuations.” Quastel Associates, January 19, 2024.

The Appraisal Foundation. USPAP Advisory Opinion 18, “Use of an Automated Valuation Model (AVM)” (1997); Advisory Opinion 37, “Computer Assisted Valuation Tools” (2018); proposed Advisory Opinion 41 (second exposure draft, 2026).

The Art Newspaper (2025). “In the can: two men jailed for Cattelan gold toilet theft.” June 16, 2025.

Title Collections Inc. (2025). Internal research: conceptual valuation model and AVM design documentation; structured interviews with institutional real estate and proptech investment professionals.

Treasury Regulation §1.170A-13(c)(3); Internal Revenue Code §170(f)(11).

Zillow Group, Inc. Form 10-K (Zillow Offers wind-down disclosure).