Finding Knowledge
Beyond the Algorithm

Ten Reasons for DDC and WebDeweySearch in 2026

Piero Cavaleri

AIB · Editor, WebDewey Italiana

The 2026 Information Landscape

What has changed — and what we risk losing

- Search engines and generative AI (GenAI) now dominate how people discover knowledge
- These tools are fast, accessible, and genuinely useful for millions of everyday tasks
- But for structured, rigorous information retrieval — especially book-level knowledge — they carry documented structural limitations
- The question is not: should we abandon AI? The question is: what are we giving up when we rely on it exclusively?

The Problem with GenAI for Literature Search

Three structural — not incidental — limitations

Hallucinated Citations

GenAI predicts plausible text, not verified sources. It generates convincing author names, journal titles, and DOIs — that do not exist.

Coverage Gaps

Models are trained on fixed datasets. Recent publications, restricted archives, and non-digitized works are systematically absent.

The Matthew Effect

High-profile works are amplified; niche and specialized research is suppressed. Popularity ≠ relevance.

The Problem with Keyword Search

Speed is not the same as understanding

- Keyword search returns a flat list — it does not represent relationships between concepts
- Homonyms, polysemy, and jargon introduce noise that users struggle to manage
- Results are ranked by engagement metrics (clicks, backlinks, SEO) — not intellectual quality
- You find what you search for, and nothing else: serendipity is systematically suppressed
- Search for 'identity' and receive a mix of philosophy, psychology, marketing, and computer science — with no way to navigate the difference

DDC is not a relic of the print era.

It is a sophisticated epistemological infrastructure
whose value increases as the limitations
of unmediated digital retrieval become more apparent.

Ten reasons why — for information professionals.

1. Structured Serendipity

The shelf effect — virtual and physical

- DDC positions related works in a contiguous classification space — by design, not by algorithm
- Users can browse the classification tree to discover what they don’t know in advance
- This adjacency is structural — not personalized, not commercial, not manipulable
- Users often cannot articulate what they need until they find it — subject classification creates the conditions for that encounter

Example

Search 'dog' in WebDeweySearch
→ 190 results

From 636.7 (Animal Husbandry)
to 599.772 (Zoology)
to 363.2 (Police Dog Services)
to 809.93 (Dogs in Literature)

Each opens a structured world.

2. A Universal Language for Knowledge

Numbers that transcend language, alphabet, and culture

30+

linguistic editions

500

= Natural Sciences in Tokyo, Nairobi and Paris

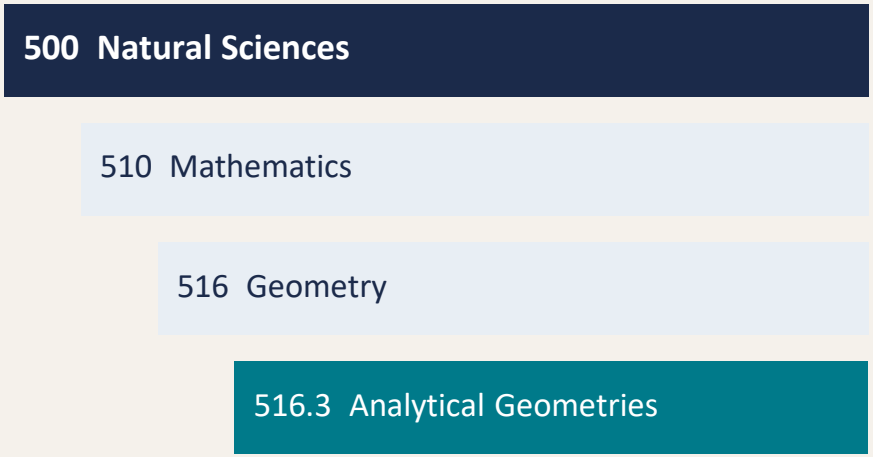
150

years of global interoperability

- Classification numbers are language-neutral — no translation required at the semantic level
- Dewey URIs can serve as stable, globally resolvable identifiers in linked data environments
- DDC can function as a Universal Schema — a semantic interoperability infrastructure that no keyword system can replicate

3. Hierarchical Logic vs. Keyword Retrieval

From the general to the specific — by design



- Navigate from general to particular without knowing the right term in advance
- Define a domain by its DDC range for systematic literature reviews

4. Countering the Fragmentation of Knowledge

Classification as an epistemological statement

- AI-based systems increasingly return decontextualized excerpts, summaries, and generated passages — not documents
- DDC operates on the opposite logic: every classification number is simultaneously a retrieval key and a disciplinary indicator
- A work on trains assigned to 385 (Rail Transport) belongs to economic and social history; another work assigned to 625.1 (Railway Engineering) belongs to applied mechanics
- That disciplinary meaning is explicit, stable, and searchable — no keyword tag can reliably encode it
- For high-quality metadata professionals: subject analysis is an intellectual act, not a bureaucratic one

5. WebDewey: A Living, Evolving System

Continuously updated — not frozen in 1876

Artificial Intelligence

006.3 and dozens of application-specific notations added — DDC keeps pace with emerging fields

LGBTQ+ & Sustainability

Previously underrepresented areas now indexed with precision thanks to expert community oversight

Global Cooperative Benefit

WebDewey updates propagate across cooperative cataloguing networks — a systemic advantage no local taxonomy can match

In DDC the tension between stability and continuous updating is a strength, not a contradiction.

6. DDC as a Cognitive Map

You don't just find documents — you learn where knowledge lives

- Keyword search gives you a list. DDC gives you a map.
- Disciplines are the backbone of the knowledge system, not a mere additional element
- Placing a topic under Sociology (301) vs. Psychology (150) is an intellectual act — it reveals methodology, language and a point of view on “reality”
- For doctoral students and researchers entering an unfamiliar field, this is invaluable guidance
- In an era of confident AI responses that conceal their epistemological assumptions, DDC's transparency about disciplinary boundaries is a corrective mechanism

> *"The classification number is informative in and of itself, regardless of the specific document it refers to."*
>
> *— Bowker & Star, Sorting Things Out (2000)*

7. Resistance to Misinformation

A curated corpus in an era of content pollution

The Open Web

- SEO-optimized noise
- Disinformation and AI-generated filler
- Engagement metrics ≠ intellectual quality
- No selection criteria

DDC-indexed Collections

- Professional selection and acquisition
- Cataloguing as an implicit quality filter
- Curated corpus, not unmediated web
- Structural mitigation against content pollution

8. Support for Interdisciplinary Research

Map the topology of your research problem — systematically

- The notes on "comprehensive works", the "see" and "see also" notes, as well as the auxiliary tables, guide the user to understand that works on disciplinary intersections centered on the researched subject may exist and where they are located.
- Cross-references in the Relative Index facilitate the mapping process across related disciplines
- Environmental law + economics? → 344.046 (Environmental Law) + 333.7 (Economics of Natural Resources)
- Auxiliary tables encode geographic, linguistic, temporal, and formal facets directly in the notation: 305.4209492 = Women (305.42) in the Netherlands (09492)
- For systematic reviews and scoping studies, DDC-based strategies offer reproducibility and transparency that keyword searches cannot guarantee

9. Independence from Terminological Change

Numbers outlast words

- Keyword searches are sensitive to terminology: search today's jargon and miss decades of literature published under yesterday's preferred terms
- DDC class numbers are assigned based on conceptual essence, not surface vocabulary
- In DDC, outdated terms are commonly documented for a long time after they have been changed
- A work indexed as 'racial relations' and a modern study on 'structural racism' are both anchored to 305.8 — regardless of era or language
- Cross-cultural consistency: English DDC uses 'White race' (305.809); Italian DDC uses 'Popolazione bianca' — same number, different terminology, same concept
- Profound implications for retrospective bibliographic research in historical and legal studies

10. Accountability and Transparency

You can teach DDC. You cannot teach a black-box algorithm.

- AI chatbots generate responses without revealing sources, training data, or confidence levels — the user receives an answer but cannot verify the process
- DDC's tables are published, editorial decisions are traceable, hierarchical logic can be systematically taught and critiqued
- When a number is assigned, the reasoning is reconstructible by trained users
- Teaching DDC gives users a transferable mental model of knowledge organization — a gift that lasts beyond any single search

A Proposal: DDC as a Semantic Anchor for A

Not just an alternative — a powerful ally

- In Retrieval-Augmented Generation (RAG) architectures, DDC notations in document metadata can anchor retrieval to a curated, stable semantic structure
- Instead of relying solely on vector proximity (prone to thematic drift), the DDC hierarchy can filter and weight retrieved fragments by disciplinary context
- A 'classification-aware' AI could state: "This answer draws primarily from Engineering [620] rather than Economics [330]" — restoring contextual transparency
- DDC does not eliminate the fragmentation inherent in AI retrieval, but it can structure and semantically link that fragmentation
- This positions library science not as a legacy profession, but as a critical partner in the development of trustworthy AI systems

At a Glance: A Comparative Overview

Functionality	GenAI Search	Keyword Search	DDC / WebDeweySearch
Primary objective	Generate plausible text	Retrieve specific items	**Map a disciplinary domain**
Underlying logic	Statistical inference	Relevance algorithms	**Hierarchical structure**
Disciplinary context	Absent / weakly inferred	Absent / user-defined	**Explicit, stable, foundational**
Resistance to manipulation	Low (training bias, ad inference)	Low (SEO, commercial ranking)	**High (editorially controlled)**
Systematic literature review	High risk of hallucinated citations	Dependent on user expertise	**Comprehensive, disciplinary coverage**

What we are defending

is not tradition.

It is epistemological rigor.

DDC offers what no algorithm can replicate:
principled classification · hierarchical navigation · disciplinary context · semantic interoperability · transparency

For information professionals, promoting these capabilities
is not about defending the past — it is about shaping the future.