DATASOURCE

Monitor

PROTOTYPE

Pricebook Digital | Companion to Strategic Portfolio Analysis

PREPARED FOR **Kevin McCarron**

February 2026 | *Companion Document*

PREPARED BY **Matthew Scott**

Table of Contents

Background

The Strategic Portfolio Analysis delivered in February 2026 identified a specific vulnerability in Pricebook Digital's competitive position: the DataSource — the curated database of HVAC manufacturer specs and pricing — requires ongoing manual maintenance across multiple manufacturer websites. When Trane updates a product line or Daikin restructures their catalog pages, someone at Pricebook needs to notice, find the changes, and update the database.

That report proposed a **DataSource Process Audit & Automation Assessment** as one of nine deliverables. This document is a companion to that proposal: a working prototype of the monitoring tool, built to test whether automated change detection is feasible.

The Maintenance Challenge

The current process is manual. Team members check manufacturer websites periodically, compare what's there to what's in the database, and update accordingly. The failure modes are predictable: changes go unnoticed for days or weeks, coverage is inconsistent across manufacturers, and there is no audit trail showing what changed and when.

The scope is concrete: **6 manufacturers** across **12 product page URLs**, each containing specs, pricing, and availability data that Pricebook's dealers rely on for accurate quoting. Manufacturers restructure their websites and update product lines without notification.

The Prototype

The DataSource Monitor is a web application that watches HVAC manufacturer product pages for changes. It checks each page, saves a copy, and flags anything that's different from last time. The dashboard shows the status of every monitored source at a glance.

Dashboard

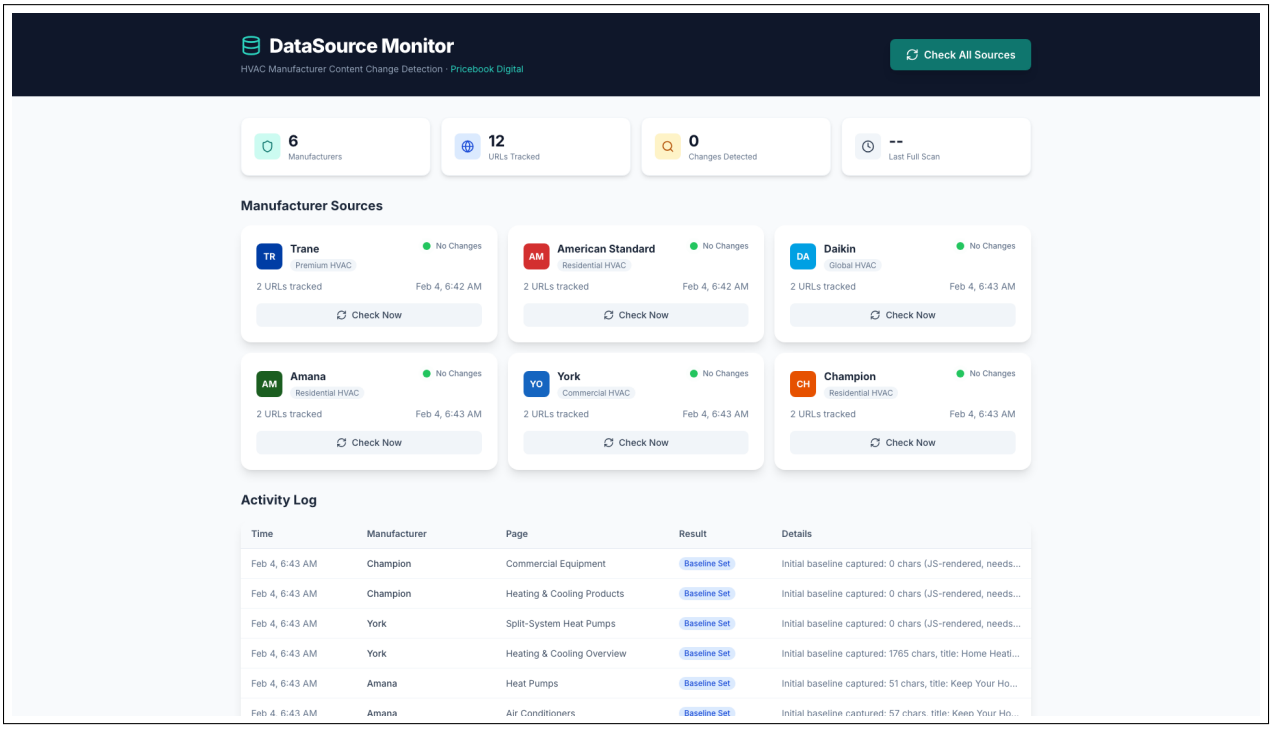

Figure 1: Live dashboard showing all six manufacturers with status indicators. The top row displays aggregate metrics. Each manufacturer card shows its category, tracked URLs, last check time, and current status.

At a glance, the dashboard shows which manufacturers have been checked, when, and whether any content has changed since the last scan. A single button triggers a scan of all sources simultaneously.

Activity Log

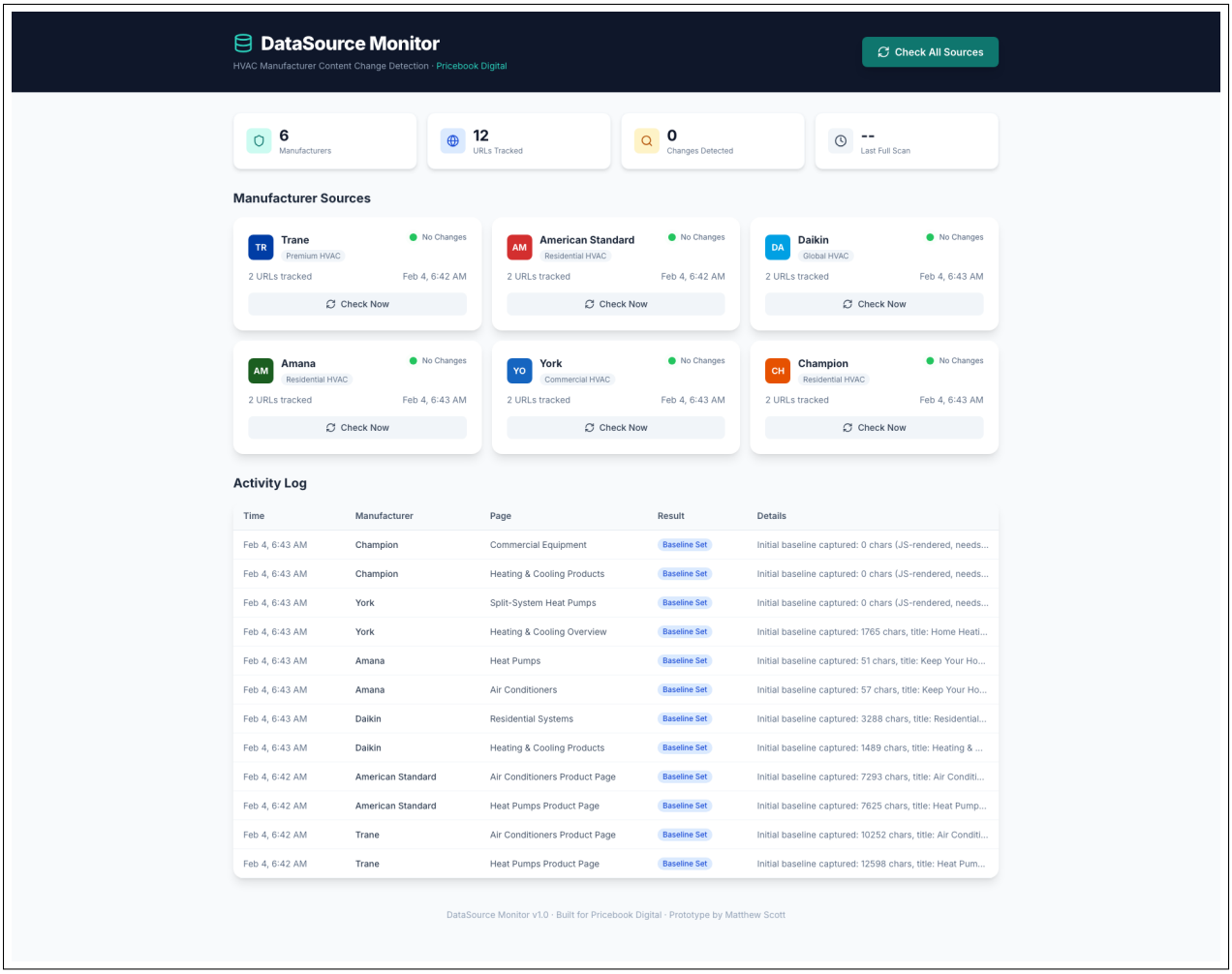

Figure 2: Full view including the Activity Log. Each scan generates a record: manufacturer, page, result, and content details. This is the audit trail that manual checking lacks.

The activity log records every scan — what was checked, when, and what happened. It answers questions like: “When did Trane’s heat pump page last change?” or “Has any Daikin content changed since the last review?”

Deployment

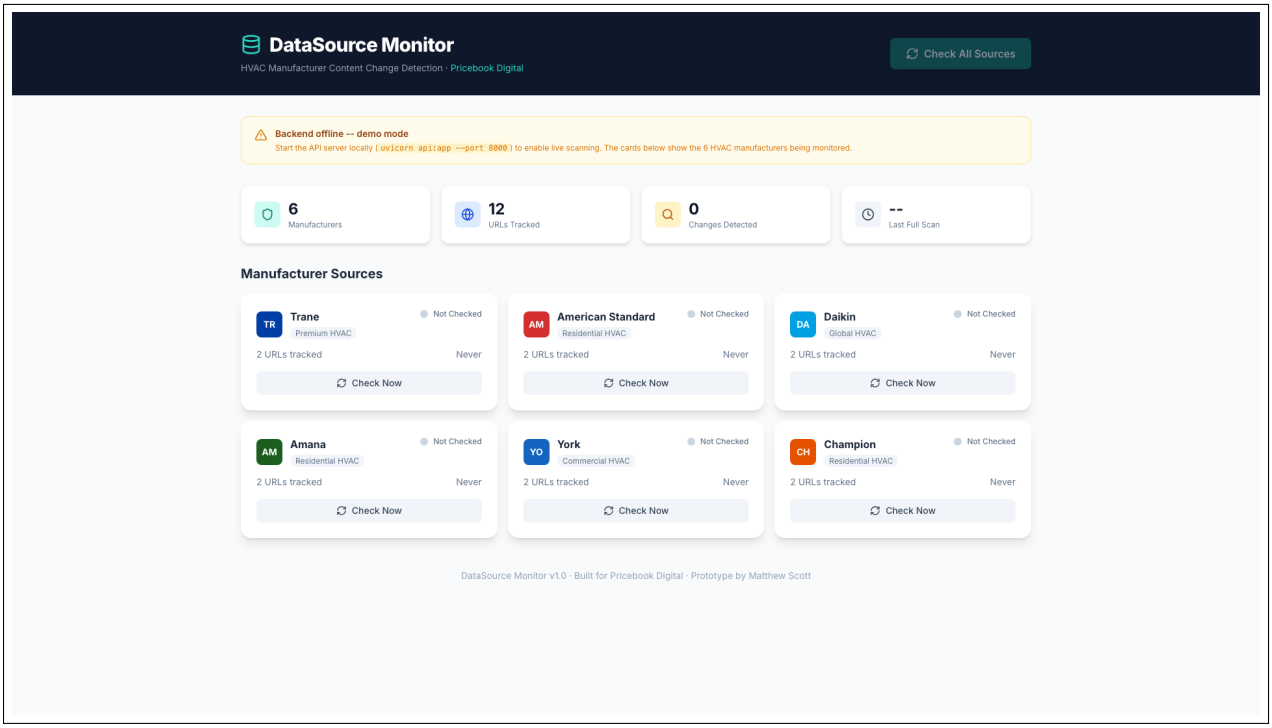

Figure 3: Deployed at `datasource-monitor.vercel.app`. The interface loads in demo mode when the backend is offline, showing the monitoring structure. It connects automatically for live scanning when the API server is running.

The prototype is accessible at datasource-monitor.vercel.app.

How It Works

Scanning Process

When a scan is triggered, the system:

1. **Visits the page** the same way a person would in a web browser.
2. **Reads the product content** — ignores menus, footers, and other website furniture. Only looks at the actual product information.
3. **Checks for changes** — compares what's on the page now to what was there last time.
4. **Shows what changed** — highlights exactly what was added, removed, or modified.
5. **Keeps a record** — logs the date, the page, and the result every time.

Manufacturers Monitored

Manufacturer	Category	Pages Monitored	Status
Trane	Premium HVAC	Heat Pumps, Air Conditioners	OK
American Standard	Residential	Heat Pumps, Air Conditioners	OK
Daikin	Global HVAC	Heating & Cooling, Residential Systems	OK
Amana	Residential	Heat Pumps, Air Conditioners	Minimal*
York	Commercial	Heating & Cooling, Split-System Heat Pumps	OK
Champion	Residential	Heating & Cooling, Commercial Equipment	OK

*These pages returned very little content. The product details may be loaded differently on Amana's site. The scanner flags this automatically so nothing slips through.

What the Scan Revealed

Running the scanner against all six manufacturers produced findings that matter for the accuracy of Pricebook's data:

> FINDING 1: CONTENT IS CHANGING
>
> **3 of the 6 manufacturers had content changes within a single afternoon.** York's heat pump page and both Champion pages showed different content between scans run hours apart. If this is happening in one day, how many changes go unnoticed in a month? Without this tool, the only way to catch these changes is for someone to manually check the page at the right time.

FINDING 2: SOME PAGES ARE UNRELIABLE

York and Champion pages don't always return the same content. One scan returned full product listings. The next scan, hours later, returned nothing. This means someone checking manually might see a page that looks fine, while the actual data being served is inconsistent.
This is the hardest kind of problem to catch by hand — because it looks fine when you look at it.

FINDING 3: NOT ALL SOURCES ARE EQUAL

Trane and American Standard returned full, consistent product data every time. These are reliable sources. **Amana returned almost nothing** — the scanner flagged this automatically. Knowing which sources are solid and which ones need attention is the first step to keeping the DataSource accurate.

Methodology

How the Data Was Collected

The scanner visits each manufacturer's public product page the same way a web browser would — it sends a request and reads the page that comes back. This is the same method Google and every search engine uses to index websites.

- Requests are spaced 1–3 seconds apart to avoid putting load on the manufacturer's servers.
- Only **public product pages** are accessed. No logins, no private data, no Pricebook systems.
- The scanner reads what the page contains at that moment — product names, specs, descriptions — and compares it to what was there last time.

What This Prototype Does Not Do

- **No Pricebook data was accessed.** This prototype only looks at manufacturer websites. It does not connect to, read from, or modify any Pricebook systems or databases.
- **No data is collected beyond what's publicly visible.** The same content is available to anyone with a web browser.
- **The dashboard contains no sensitive information.** The live link (`datasource-monitor.vercel.app`) shows manufacturer names, page titles, and change status — all derived from public pages. When the backend is offline, it runs in demo mode showing only the structure.

Why Results May Vary

Some manufacturer websites deliver their content differently depending on when you visit. York and Champion, for example, sometimes return full product listings and sometimes return an empty page that expects a web browser to load the content separately. This is a known behavior of modern websites and is documented in the findings.

If you run the scanner yourself and see different results than what's in this report, that is expected — and it illustrates the inconsistency that makes manual checking unreliable. A production version would use a full browser engine to get consistent results every time.

Research Context

Building this prototype required understanding more than just how to scan web pages. The following research shaped what the prototype does, what it deliberately does not do, and where the real opportunities are.

The HVAC Data Supply Chain

HVAC equipment data doesn't start at a manufacturer's website. It moves through a chain, and each layer adds complexity:

Layer	What Happens Here
Raw Materials	Copper, aluminum, and steel prices set the floor for manufacturing costs. When commodity prices move, equipment prices eventually follow.
Manufacturing	Manufacturers build equipment and publish specs. This is what the prototype monitors today.
AHRI Certification	The Air-Conditioning, Heating, and Refrigeration Institute independently certifies equipment performance (SEER, EER, HSPF ratings). This is the authoritative source for specs.
Distribution	Distributors like Ferguson HVAC and Winsupply set actual contractor pricing. This data sits behind login-protected dealer portals.
Contractor Software	Platforms like ServiceTitan and Housecall Pro consume equipment data to power proposals and invoicing.
The Homeowner	Gets a quote. The accuracy of that quote depends on every layer above being correct.

A product like Pricebook's DataSource sits between distribution and contractor software — it's the data layer that keeps everything downstream accurate.

What the Research Uncovered

> AHRI: A BETTER DATA SOURCE EXISTS
>
> The AHRI Directory of Certified Product Performance is the industry-standard database of verified equipment specifications. It offers a paid Data Subscription Program with API access and bulk data downloads. This means structured, authoritative equipment data is available without scraping individual manufacturer websites — which are inconsistent, change without notice, and sometimes don't return content at all (as this prototype demonstrated with York and Champion).

DISTRIBUTOR PRICING IS NOT ACCESSIBLE

The most valuable data for HVAC contractors — what they actually pay for equipment — lives behind authenticated dealer portals. Each distributor manages their own pricing, often negotiated per customer. Web scraping cannot reach this data without credentials, and using credentials without authorization would be a terms-of-service violation. This is a gap that requires business relationships, not technology.

COMMODITY PRICES DRIVE EQUIPMENT COSTS

HVAC equipment costs are directly tied to copper, aluminum, and steel markets. The ongoing refrigerant transition from R-410A to R-454B is contributing to a 10–20% increase in new equipment costs. Tracking these upstream signals could predict when manufacturers will adjust pricing — before the changes appear on their websites.

CONTRACTOR SOFTWARE DOESN'T SOLVE THIS PROBLEM

ServiceTitan, Housecall Pro, and FieldEdge all rely on manually-maintained internal pricebooks or direct manufacturer partnerships for equipment data. None of them automatically detect when a manufacturer updates their product line. This is the gap Pricebook's DataSource fills — and the gap this prototype helps protect.

How This Shaped the Prototype

This research informed specific decisions about what the prototype does and does not attempt:

- **It monitors manufacturer websites** because that's where Pricebook's data originates today. The prototype validates whether automated detection works at this layer. It does.
- **It does not extract structured data** (individual model numbers, SEER ratings) because the AHRI API is a better source for that. Scraping product pages for structured specs would be reinventing a wheel that already exists.
- **It does not attempt distributor pricing** because that data is behind authentication walls. No amount of scraping solves a business-relationship problem.
- **It does not predict price changes** yet, but the commodity-price research suggests this is feasible — and valuable — as a future capability.

The prototype is scoped to answer one question well: *are manufacturer product pages changing, and can we detect it automatically?* The research confirms that this is the right starting point, and that there are clear paths forward once it's validated.

Next Steps

This prototype validates the approach. The following options build on what exists:

1. **Walk through it together** — Run a scan, review the dashboard, and discuss how it maps to Pricebook's current DataSource maintenance workflow. The tool is at datasource-monitor.vercel.app.
2. **Map the full update workflow** — The prototype handles detection. Understanding how detected changes flow into Pricebook's database — who verifies, who updates, what the turnaround time is — would shape the next version.
3. **Improve Amana coverage** — Two Amana pages are not returning full product details yet. This is fixable.
4. **Define notifications** — When a change is detected, who should know and how? Email, Slack, or a daily digest are straightforward to add.
5. **Schedule automated scans** — Right now scans are triggered manually. A production version would run on a schedule (e.g., daily) and keep a history of every change.

QUESTIONS FOR KEVIN

1. Is keeping the DataSource current a known pain point for the team, or is this surfacing something that hasn't been a focus yet?
2. How many manufacturers does Pricebook track total? This prototype covers 6 — is that a representative sample or a small slice?
3. Does Pricebook currently use AHRI data, or is that a new avenue worth exploring?

> The portfolio analysis identified the DataSource maintenance challenge. This prototype tests whether the solution works. The monitor is running, the scanning logic is validated, and the remaining questions are about how to connect it to Pricebook's internal workflow.

About This Report

This report is a companion to the *Strategic Portfolio Analysis* delivered in February 2026. The prototype was built using publicly available manufacturer product pages. No proprietary Pricebook data was accessed or required.

Technical Summary

Component	Detail
Backend	Python web server that fetches and compares pages
Change Detection	Compares page content against saved copies
Frontend	Modern web dashboard (React, TypeScript)
Storage	File-based (prototype); upgradeable to database
Deployment	Vercel (dashboard), local server (scanning engine)
Size	Small, focused codebase — backend + frontend

Report generated using the Texume document platform (Jinja2 + LaTeX).

Prepared For	Kevin McCarron
Prepared By	Matthew Scott
Date	February 2026
Version	Companion Document