JAKUB BŘEZÍK

VÝVOJÁŘ · DATOVÁ A API INTEGRACE

Praha, Česká republika
+420 775 611 501
jakub@brezik.net
brezik.net
linkedin.com/in/jakub-brezik



PROFIL

> Vývojář s 20 lety v IT, z toho posledních 10 v datových skladech a datové integraci mezi systémy v bankovnictví a pojišťovnictví. V letech 2018–2026 jsem v KBC řešil integraci číselníků ze zdrojového systému do datového skladu: návrh, vývoj, orchestraci, testování a každodenní provoz napříč čtyřmi prostředími a deseti entitami skupiny KBC. Navrhuji asynchronní, souborové integrační toky s validací a fallback mechanismy a vystavuji kolem nich služby. Z nepřehledného dělám přehledné, v datech, nástrojích i dokumentaci.

DOVEDNOSTI

Integrace

REST API · OpenAPI · Informatica BDM · Informatica MDM/RDM · Informatica PowerCenter · Stonebranch UAC/UDM

Programování

Java · Spring Boot · Perl · Python · PL/SQL

Nástroje

BTEQ · Apache Velocity · Maven

Cloud

AWS · EC2 · S3 · ALB/WAF · Cognito · Lambda

Databáze

Teradata · Oracle · PostgreSQL

DevOps a CI/CD

Jenkins Pipeline · Docker · AWS CloudFormation · Git · Unix

AI

Claude Code · AI agenti · Vývoj řízený specifikací

Spolupráce

Jira · Confluence

Modelování

PowerDesigner

ZKUŠENOSTI

INTEGRAČNÍ A ETL VÝVOJÁŘ, ČÍSELNÍKY

KBC (přes Adastru) | 2018 – 2026 · Praha

Projekt Rainbow: centrální datový sklad skupiny KBC na Teradata (model FSDM) pro interní a regulatorní reporting včetně reportingu do ECB. Obsluhuje deset lokálních a skupinových entit. End-to-end vlastník toku číselníků z RDM do DWH na Teradata. Číselníky jsou enablerem: bez aktuálních a správných hodnot se zbytek skladu nedá spolehlivě načíst.

Architektura a návrh

- Implementoval jsem exportní mappingy v Informatica BDM nad existujícím kanonickým formátem – view, které složí business objekty RDM na jednu tabulku, takže jeden mapping obsluhuje všechny číselníky a dobře škáluje.
- Navrhl jsem obrácení toku dat mezi prostředími (z produkce do integrace, z integrace do vývoje) pro interní testování dat RDM, které vyžadovalo živá data. Obhájil jsem ho před technickým výborem, získal schválení a vyjednal routing s platformovým a bezpečnostním týmem.
- Zavedl jsem OpenAPI specifikace a swagger-ui pro REST služby data testingu; aplikační tým RDM převzal stejné nastavení pro vlastní rozhraní.
- Udržoval jsem integrační dokumentaci a katalogy obsahu a vlastnictví domén v Confluence, které používá tým, analytici a provoz DWH.

Integrační vývoj

- Vlastnil jsem exportní řetězec: denormalizační view nad databází RDM, mapping v Informatica BDM, export do souboru, přenos souborů na unixové servery, ingest a load přes databázové vrstvy na Teradata.
- Standardní load jeden job na tabulku (400+ jobů) jsem nahradil smyčkovým jobem, který každou tabulku validuje, nahraje ty, které projdou, a zastaví se na null hodnotách nebo nesouladu struktury.
- Napsal jsem moduly pro housekeeping: archivaci a retenci exportů, housekeeping RDM přes jeho REST API a kontroly testování dat, které volají vlastní testovací službu.
- Udržoval jsem XML definice zdrojových souborů, normalizační ETL a loadovací transformace uložené v PostgreSQL.

JAZYKY

Čeština (rodilý mluvčí)

Angličtina (pokročilá)

CERTIFIKACE

AWS Certified Cloud Practitioner
2021 – 2024

Informatica Developer Tools for Big Data Developers; Big Data for Developers
2018

Informatica PowerCenter 9.x Developer, Level I a II
2016

Teradata Advanced SQL
2015

VZDĚLÁNÍ

Masarykova univerzita, Fakulta informatiky, Aplikovaná informatika (nedokončeno)
2005 – 2006

SOŠ a SOU Spojů, Tábor — elektrotechnika, telekomunikace a výpočetní technika
2001 – 2005

HOMELAB

Proxmox a Docker doma: zkouším nové technologie, provozuji vývojové a testovací prostředí pro ukázky a oddělené herní servery, na kterých si moje děti mohou hrát bezpečně.

Proxmox | Docker

Služby a nástroje (Java, Spring Boot)

- Služba pro testování dat: porovnává exporty se zdrojem v RDM a cílem na Teradata a vystavuje kontroly jako REST API. Nahradila ruční porovnávání v Excelu, kde jeden běh dotazu trval desítky minut; teď trvá sekundy. Výsledky jsou napojené na testovací sady v Jira Zephyr a tabulky v Confluence přes jejich API.
- Generátor prostředí: vytváří integrační, akceptační a produkční varianty zhruba 1 000 šablonových objektů entity včetně toků a orchestrace, takže jedno vývojové prostředí se chová jako tři prostředí. Load sandboxu pro jednu entitu trval kolem 90 minut a sandbox se musel mezi datovými sadami mazat; migrační testy teď nahrávají všechna tři prostředí najednou a analytici porovnávají živě. Doba běhu testovacích sad se zkrátila ze dnů na hodiny, což umožňuje několik loadů denně.
- Provozní dashboard: spouští kontrolní SQL nad PostgreSQL a po každém denním loadu publikuje do Confluence kompaktní stavovou mřížku (tři prostředí × deset entit × vrstvy), plánováno v Jenkinsu. Každá dlaždice se rozbalí na detail defektů a odkazy. Pokrývá místo v databázi, chybějící objekty, posun sloupců, null hodnoty, zpoždění loadů, měkké chyby, nové a zastaralé tabulky, skoky objemů a prázdné loady, čímž zkracuje dobu odhalení a opravy chyb. Předáno provozu DWH.
- Offline prohlížeč číselníků: převádí exporty do jednoho excelového sešitu s rejstříkem a záložkou pro každý číselník včetně metadat o vlastnictví, pro uživatele bez přístupu do RDM. Zavedl jsem doručování generovaných sešitů číselníků na firemní SharePoint přes platformu pro řízený přenos souborů.

Provoz a DevOps

- Každodenní provoz loadů číselníků na všech prostředích. Číselníky jsou enablerem pro zbytek skladu, takže chyby diagnostikoval z logů Unixu, Teradata a PostgreSQL a opravoval je okamžitě.
- Psal jsem vlastní Jenkins pipeline pro buildy, nasazování i provozní úlohy.

Java | Spring Boot | REST | OpenAPI | Informatica BDM | Informatica MDM/RDM | Stonebranch UAC/UDM | Teradata | Oracle | PostgreSQL | Perl | Python | Jenkins | Jira | Confluence | Unix

DWH DEVELOPER

Komerční banka (přes Anywhere) | 2017 – 2018

Interní rozvoj DWH na Teradata: rozšíření modelů HR tabulek a workflow, rozšíření a aktualizace KPI transformací pro další rok.

Teradata | BTEQ | PowerDesigner

BPM DEVELOPER

IBM (kontraktor) | 2017

Refaktoring hlavního obchodního modulu a implementace procesů na platformě IBM BPM.

MS SQL | IBM BPM | JavaScript

DWH DEVELOPER, IBFS

Société Générale (přes Accenture) | 2016 – 2017

Stavba datových skladů na Teradata pro mimofrancouzské země podle šablon: parametrizace a nasazení podle šablony, modelování specifik jednotlivých zemí v PowerDesigneru. Zavedl jsem Apache Velocity pro generování kódu ze šablon.

Teradata | BTEQ | PowerDesigner | Apache Velocity | PowerShell

ETL DEVELOPER, SOLVENCY II

NN (přes Accenture) | 2015 – 2016

ETL transformace v Informatica PowerCenter nad Oracle a datové modelování v PowerDesigneru. Funkční a technický design, analýza defektů, code review, unit testy, testovací scénáře a podpora systémového a UAT testování.

PowerDesigner | Informatica PowerCenter

DŘÍVĚJŠÍ PRAXE

CSC, SkillZone | 2006 – 2014

Zákaznická podpora, koordinace nasazení aplikací a migrací serverů, knowledge management a reporting, správa sítí a systémů.

Remedy | ServiceNow | MS SQL | IBM Cognos

PROJEKTY A AI

AI je můj každodenní pracovní nástroj: rozbor chyb a logů, skripty, testy, revize kódu. Každý výstup ověřují testy a review, takže zrychlení nejde na úkor přesnosti.

ARCANISTA – SPRÁVA SBÍRKY MAGIC: THE GATHERING

Vlastní projekt (rozpracováno)

Webová aplikace pro sběratele TCG: co mám, kde to leží, kolik to stojí a co chybí do kompletní sady nebo balíčku. Produkt, UI/UX, architektura i vývoj sám, s AI agenty jako vývojovým týmem.

Vývoj s AI

- Agenty Claude Code řídím jako tým: schválená specifikace, z ní testy v jedné session, implementace v jiné, která testy nesmí měnit. CI to vynucuje.
- Paralelní agenti v oddělených git worktrees berou issues z GitHubu a otevírají pull requesty; já rozhoduji a schvaluji.
- Za dva měsíce přepis z Django/Python na TypeScript: 1 100+ commitů, 170 pull requestů, 220 testovacích souborů, 37 zaznamenaných architektonických rozhodnutí (ADR).

Čím se liší

- Každá fyzická karta je samostatná kopie se stavem, jazykem, nákupní cenou a umístěním, ne jen počet kusů.
- Skutečné uložení: krabice, alba a složky do libovolné hloubky, takže víte, kde karta leží.
- Skenování mobilem: rozpoznání karet z fotky přes obrazové embeddingy, učí se z oprav; nic se nezapíše bez potvrzení.
- Ceny z českých, evropských a amerických obchodů s historií; zdroje nemíchá a chybějící cenu nenahrazuje nulou.
- Každou úpravu jde vrátit, smazané jde do koše a ovládá se jako desktopová aplikace.

Cloud

- Vývojové prostředí na AWS jako infrastruktura v kódu (CloudFormation): EC2, ECR, ALB s WAF, Cognito, Route 53, S3, SES s Lambdou, rozpočtový alarm. Nasazení přes OIDC, bez uložených klíčů.
- Nasazení je otázkou pár kliknutí; mezi testovacími koly běží aplikace lokálně v Docker Compose kvůli úspoře nákladů.

Claude Code | TypeScript | Node.js | Hono | React | TanStack | Vite | PostgreSQL | pgvector | Drizzle | ONNX Runtime | Docker Compose | Vitest | Playwright | Storybook | AWS CloudFormation