

big hub

AI-ready data 2027

Jak připravit firemní data pro
AI a firemní automatizaci

Jasný plán a priority:

Co udělat nejdřív a proč

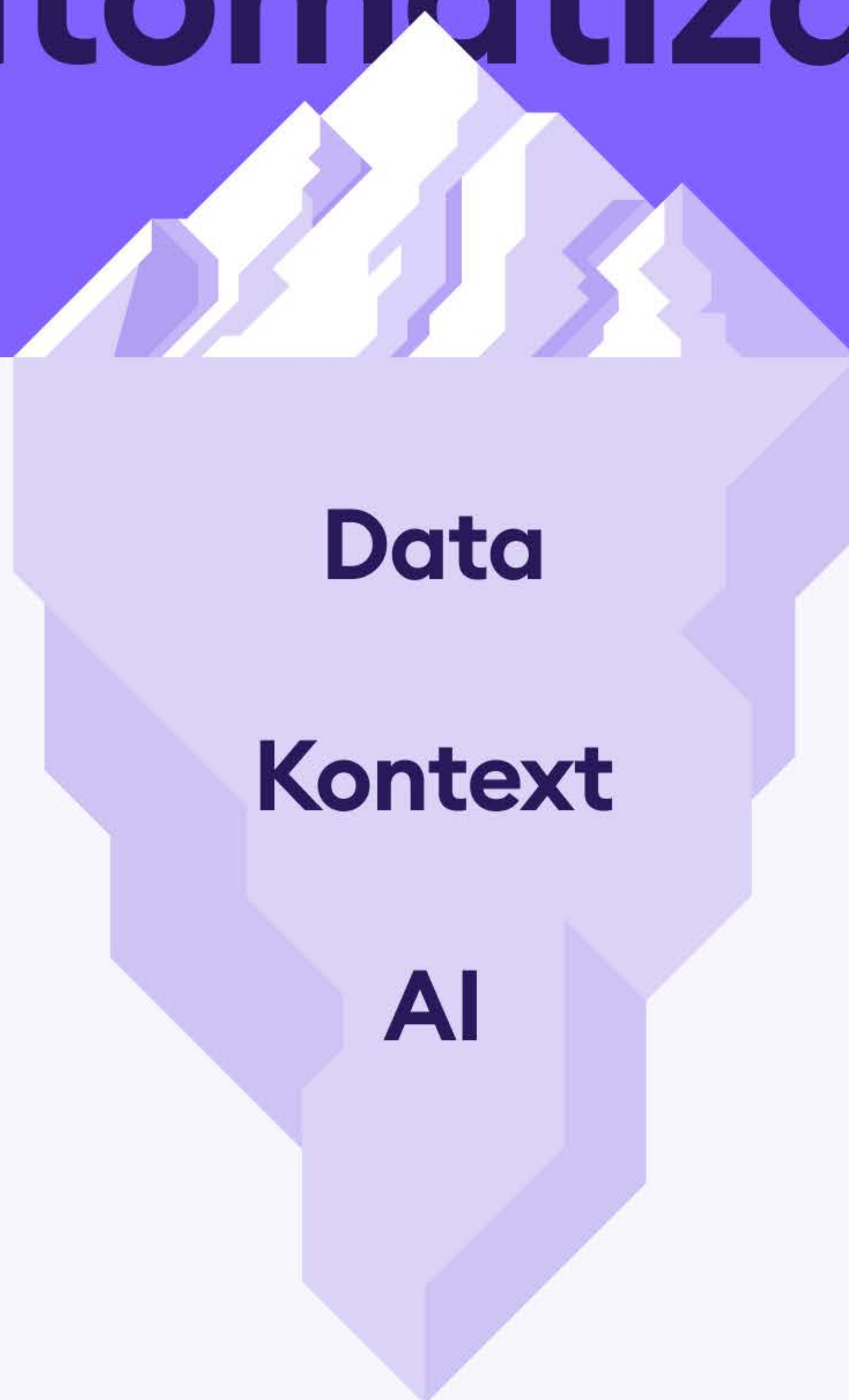
Vzorová architektura:

Architektura pro škálování

Checklist

a 90-denní roadmapa

Automatizace



Bez kvalitních a dostupných **dat** není co chápat ani z čeho rozhodovat. Když data doplníte o **kontext** – definice, dohledatelnost původu a bezpečné přístupy – vznikne důvěryhodný podklad pro řízení firmy. Na něm pak může **AI** spolehlivě hledat souvislosti a doporučovat další kroky. A nakonec **automatizace** převede tato doporučení do bezpečné, přesné a auditovatelné akce.

Obsah

Shrnutí pro vedení	4
Proč BigHub	5
Co jsou AI-ready data	6
Hlavní pilíře: Co zavést jako minimum	7
Proč už nestačí monolitický datový sklad?	8
ODS a hot data: aktuální stav bez rizika pro core systémy	9
Real-time data: aby AI nereagovala se zpožděním	10
API jako produkt: stabilní rozhraní pro AI a nové konzumenty	11
Vzorová AI-ready data architektura	13
Nezbytnosti ke škálování	15
Jak AI bezpečně pracuje s firemními podklady	16
Tok zpracování podkladů	17
Jak měřit kvalitu AI	18
Governance a rizika	19
Směrnice a compliance	20
Řiďte data jako produkt	21
Jak varování nastavit	22
Roadmapa 90 dní + 12 měsíců	23
Orientační skóre datové maturity	25
AI-ready Data Assessment Workshop	26
 Jak e-book používat	
High-level shrnutí (10 min)	4, 23-24
Do hloubky (30–40 min)	5–16

Shrnutí pro vedení

Proč to řešit teď

AI a automatizace zvyšují nároky na **důvěru v data, dohledatelnost a bezpečné přístupy**. Bez těchto základů je AI těžké škálovat a roste riziko chyb.

Co si odnesete

- ✓ Jednoduchou definici AI-ready dat
- ✓ Co musí existovat, aby se AI dala bezpečně škálovat
- ✓ Minimum pravidel a odpovědností (bez nadbytečné byrokracie)
- ✓ 90-denní roadmapa a 12-měsíční výhled

Rychlá kontrola (10 otázek)

1. Máme 1 definici klíčových KPI napříč firmou?
2. Umíme dohledat původ KPI (odkud se číslo vzalo)?
3. Máme jasné vlastníky datových domén?
4. Máme pravidla přístupů podle citlivosti?
5. Máme testy a hlídání kvality u kritických dat?
6. Máme katalog/metadata, kde data najdeme a pochopíme?
7. Víme, které dokumenty jsou „zdroj pravdy“?
8. Umíme AI omezit na vybrané zdroje?
9. Máme systém pro průběžné měření kvality AI odpovědí a výstupů?
10. Máme roadmapu na dalších 90 dní, rozdělené role a termíny?





Proč BigHub

Jsme data & AI engineering a consulting firma, která pomáhá organizacím proměnit data v měřitelnou byznysovou hodnotu – ať už jde o budování datových platforem, přípravu AI-ready dat, návrh a nasazení AI řešení nebo automatizaci klíčových procesů.

Dodáváme end-to-end službu: konzultaci a nastavení strategie, implementaci i následnou údržbu a dlouhodobou podporu. Spolupracujeme se startupy, enterprise firmami i institucemi a nadnárodními organizacemi. Stavíme řešení se standardy pro bezpečnost, odpovědnost, měřitelnost a škálování.

Náš příběh v číslech

2016
založeno

75+
expertů

500+
projektů

100+
enterprise
klientů



Co vidíme v praxi

Firmy mají data, ale chybí společný jazyk KPI, odpovědnosti, dohledatelnost a řízené přístupy. AI pak jen zvýrazní slabiny – a důvěra se ztrácí rychle.

Mini-slovníček

Stopa původu dat: odkud se číslo vzalo.

Vrstva společných definic (KPI): jedna definice metrik napříč firmou.

Datový produkt: data s vlastníkem, standardem kvality a jasným využitím.

Odpovědi s oporou o interní dokumenty: AI odpovídá podle dohledaných firemních zdrojů a uvádí citace.

Včasné varování kvality: systém hlídá, že se data nekaží potichu.

Co jsou AI-ready data

Definice

AI-ready data jsou data připravená pro AI:

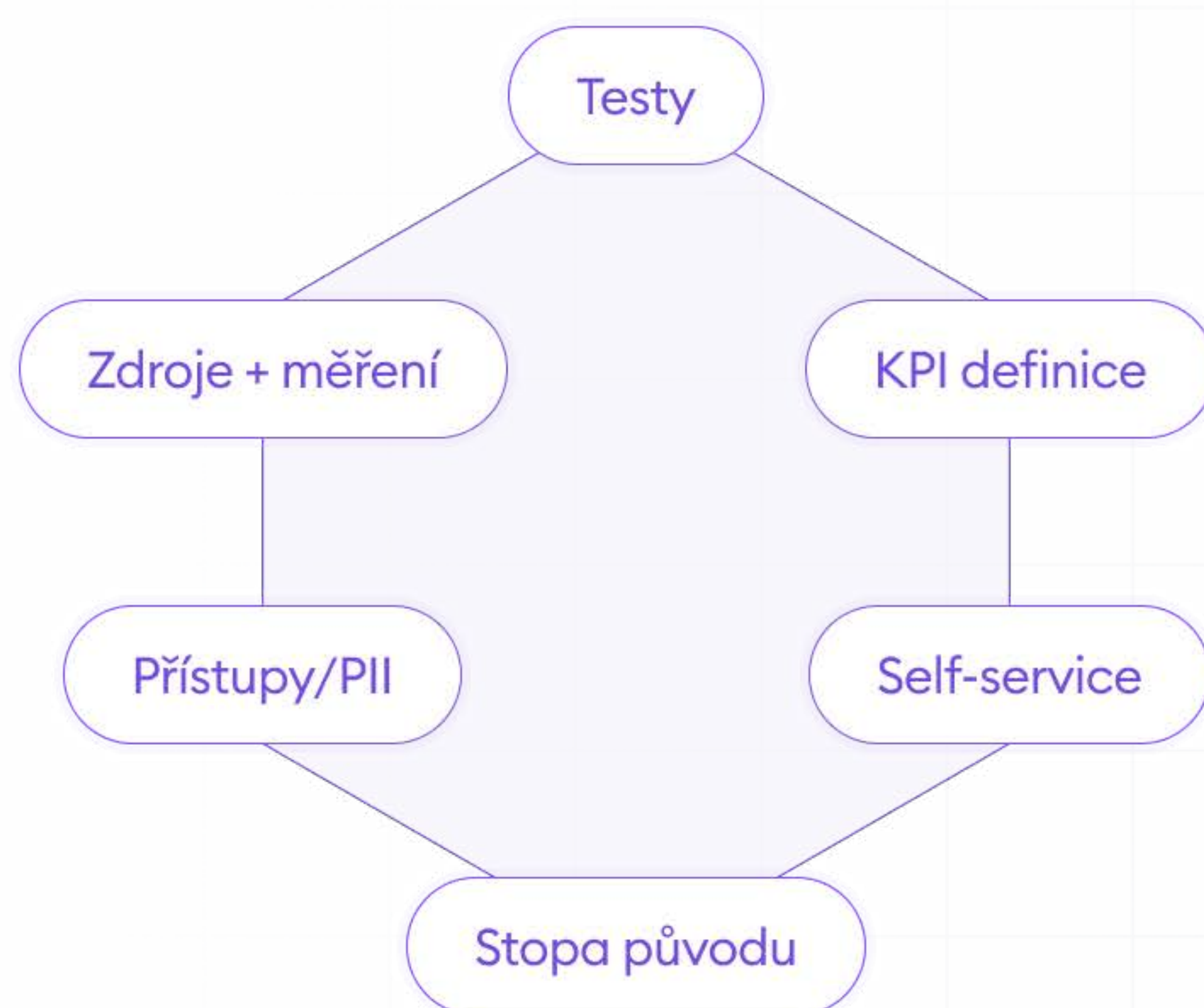
důvěryhodná, s jednotnými definicemi, dohledatelným původem, bezpečně přístupná a použitelná pro rozhodování, automatizaci i práci s dokumenty.

6 pilířů AI-ready dat

1. **Kvalita a spolehlivost** (data se “nedrolí”)
2. **Společný jazyk KPI** (1 definice napříč firmou)
3. **Rychlá dostupnost** (bez zbytečných úzkých míst)
4. **Stopa původu** (odkud se číslo vzalo)
5. **Bezpečné přístupy** (podle citlivosti)
6. **Připravenost pro AI** (zdroje, omezení, měření kvality)

99 Key takeaways

“AI-ready není nákup nástroje. Je to schopnost organizace opakovaně používat data bezpečně a konzistentně.”



Hlavní pilíře: Co zavést jako minimum

Jak to poznáte v praxi

AI-ready data nevzniknou jedním projektem ani nákupem nástroje. V této kapitole ukazujeme šest klíčových pilířů a jejich **minimum**, které stačí zavést, aby data byla důvěryhodná, bezpečně použitelná a připravená pro škálování AI i automatizace.

 Pilíř	 Kritický signál	 Minimum
 Kvalita	Report se mění bez vysvětlení (týden od týdne jiné číslo bez dohledatelné příčiny)	Testy + hlídání čerstvosti dat
 Jednotné KPI	5 definic tržeb (každé oddělení počítá KPI jinak)	Společná definice + jednoduchý proces změn
 Dostupnost	Data chodí pozdě (čekání na data/analýzu v řádu dnů až týdnů)	SLA + vrstva pro rychlé dotazy (ODS/API)
 Stopa původu	Nevíme odkud je KPI (neexistuje jasná stopa: zdroj → úpravy → výpočet)	Dohledatelnost pro top KPI
 Přístupy	Nevíme, kdo co smí vidět (chybí role podle citlivosti, přístupy jsou rozvolněné)	Role podle citlivosti + audit
 AI výstupy	AI občas halucinuje (odpovědi bez opory v interních zdrojích, bez citací)	Schválené zdroje + citace + kontrolní sada otázek

-  **Pilíř** = oblast, kterou musíte mít zvládnutou, aby data byla připravená pro AI
-  **Kritický signál** = konkrétní, snadno rozpoznatelný symptom, že pilíř u vás není v pořádku.
-  **Minimum** = nejmenší smysluplný balíček opatření, který stačí zavést jako základ, aby se situace výrazně zlepšila



V dalších kapitolách ukážeme, jak se tyto pilíře promítají do architektury: **jak dostat správná data ve správný čas (ODS, real-time) a jak je bezpečně zpřístupnit přes stabilní rozhraní (API).opakovaně používat data bezpečně a konzistentně.**

Proč už nestačí monolitický datový sklad?

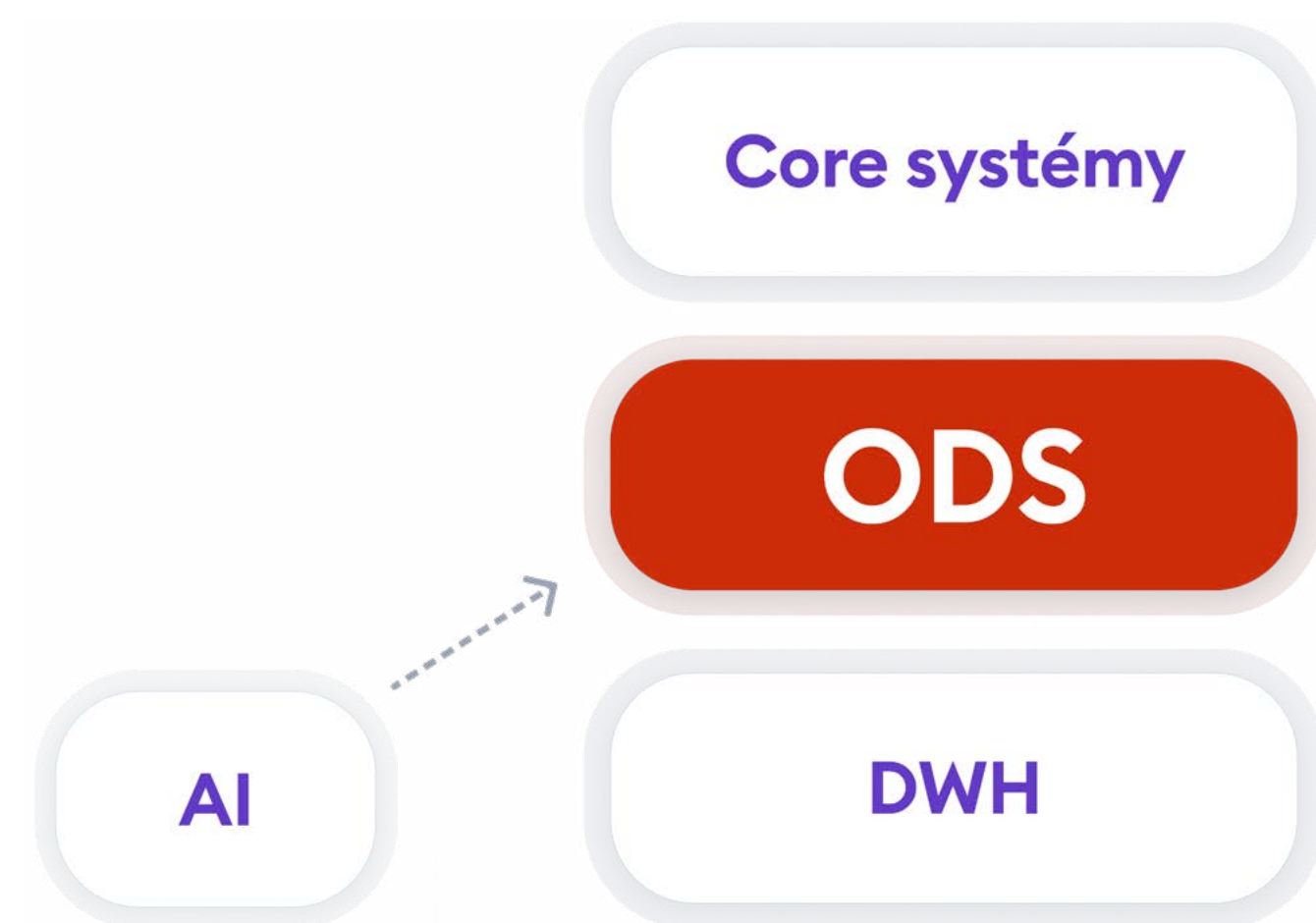
Jeden DWH = jeden kompromis pro všechny use-casy.

Datový sklad a reporting jsou skvělý základ pro analýzu a zpětný pohled. Jenže

AI a automatizace kladou na data vyšší nároky: potřebují **jednotné definice, důvěru, dohledatelnost původu, řízené přístupy a průběžné hlídání kvality.**

- ⊖ **Drahý storage:** Monolitický DWH často drží všechna data „na stejné úrovni výkonu“, i když velká část z nich je jen historická. Náklady pak rostou nejen objemem, ale hlavně tím, že výkonný storage a výpočet škáluješ kvůli pár náročným use-casům.
- ⊖ **Nedostupnost technologií:** DWH je typicky postavené pro SQL analytiku a BI, takže moderní potřeby kolem AI (nízká latence, práce s eventy, nestrukturovaná data, rychlé operativní dotazy) do něj nezapadají „nativně“. Výsledkem jsou dodatečné vrstvy a integrace mimo DWH, které zvyšují komplexitu.
- ⊖ **Optimalizováno na reporting:** Architektura DWH počítá s dávkovým zpracováním, agregacemi a historickými pohledy, ne s „teď hned“ dotazy. Pokud přes něj začneš obsluhovat AI/agentní scénáře, typicky narazíš na latenci a neefektivní dotazy, protože systém není optimalizovaný na operativní workloady.
- ⊖ **Složitý change management:** V monolitickém DWH bývá všechno provázané – úprava schématu nebo pipeline může rozbít reporty, downstream transformace i závislosti dalších týmů. To zpomaluje iterace a zvyšuje riziko změn, což je přesný opak toho, co AI projekty potřebují.
- ⊖ **Nepodpora asynchronních rychlostí DWH:** Monolitický přístup často znamená „jedna rychlost pro všechny“ (typicky D-1 nebo micro-batch), takže real-time a low-latency scénáře se musí přizpůsobit nejpomalejší vrstvě. AI use-casy pak běží na starých datech, nebo se řeší obcházením DWH a vzniká další technický dluh.“

ODS a hot data: aktuální stav bez rizika pro core systémy



ODS a hot data: aktuální stav bez rizika pro core systémy

Mnoho organizací nechce pouštět AI přímo do core systémů (ERP, CRM, SAP,...) – kvůli výkonu, bezpečnosti a riziku dopadu na provoz. Zároveň je datový sklad často navržený pro reporting, ne pro real-time dotazy. U AI use-casů navíc platí jednoduché pravidlo: čím vyšší přínos pro byznys, tím vyšší nárok na aktuálnost.

Co je ODS a proč je praktické

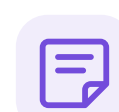
ODS (Operational Data Store) je operativní vrstva pro „hot data“ – tedy pro aktuální stav. Funguje jako rychlá a kontrolovaná mezivrstva mezi core systémy a konzumenty (AI, aplikace, automatizace).

ODS umožňuje:

- ✓ Rychlé odpovědi nad aktuálním stavem (bez čekání na dávkové zpracování),
- ✓ Snížení zátěže core systémů,
- ✓ Nastavení přístupů a auditní stopy,
- ✓ Jasně oddělení: aktuální stav (ODS) vs historie a analýzy (DWH/BI).

Typické use-casy „hot dat“

- Stav objednávky teď hned
- Aktuální cena a dostupnost
- Stav reklamace / incidentu / požadavku
- Aktuální limit, stav smlouvy, stav dodávky



Enterprise realita (např. SAP)

U velkých organizací bývá přístup k datům ze SAPu mimo SAP často drahý a pomalý (změny, schvalování, náklady). ODS pomáhá zrychlit využití dat pro AI tím, že poskytne stabilní operativní rozhraní pro nejčastější dotazy nad aktuálním stavem.

Minimum, které stačí zavést

- Vyberte 3–5 hot use-casů s nejvyšší hodnotou.
- Definujte pro každý use-case „zdroj pravdy“ a pravidla přístupů.
- V ODS držte jen to, co potřebujete pro aktuální stav (ne veškerou historii).
- Nastavte základní dohled: čerstvost, objem, změna struktury.

Real-time data: aby AI nereagovala se zpožděním

Co se nejčastěji pokazí

- Události dorazí pozdě nebo vůbec, ale nikdo si toho nevšimne včas.
- Data nemají základní čištění (duplicitní, neúplná, nelogická).
- Neexistuje jasná odpovědnost a reakční čas, když se tok rozbije.

Není real-time jako real-time.

Je potřeba si nejdříve sjednotit očekávání v real-time.

= kopie core systémů v rámci minut

= analytika na stream datech (hodinová agregace)

= kopie zdrojových systémů na úrovni eventů s SLA na úrovni několika vteřin

= transakční kopie

Ve firmách často existuje real-time vrstva jen na oko: události někde tečou, ale bez jasného toku, základní validace a dohledu. Pro AI to má přímý dopad: systém reaguje se zpožděním – a i „správná“ odpověď může být v okamžiku doručení mimo realitu.

Minimum real-time vrstvy pro AI (bez přehnané složitosti)

1. Jasný tok: odkud → kam → kdo je konzument.
2. Základní validace: povinná pole, duplicita, extrémy.
3. Dohled nad doručením: víme, že data přišla včas a v očekávaném objemu.
4. Varování a reakce: vlastník + reakční čas + jednoduchý postup nápravy.

Jak poznáte, že to funguje

- ✓ Zpoždění se zachytí dřív, než ovlivní zákazníka nebo proces.
- ✓ Incidenty mají jasnou příčinu a prevenci.
- ✓ AI odpovědi nad hot daty jsou stabilní a předvídatelné.

API jako produkt: stabilní rozhraní pro AI a nové konzumenty

Proč API často brzdí škálování AI

API v mnoha firmách vzniklo jako vedlejší produkt jedné aplikace (web, mobil, konkrétní stack). Když pak chcete připojit nové konzumenty – AI agenty, partnery, další týmy nebo kanály – ukáže se, že rozhraní je křehké, bez verzí a změny jsou drahé.

Co se děje v praxi

- Každá změna aplikace rozbije další konzumenty.
- Chybí verzování a jasná smlouva rozhraní.
- Připojování nových konzumentů je pomalé a zatěžuje IT i byznys

API jako produkt (minimum, které stačí)

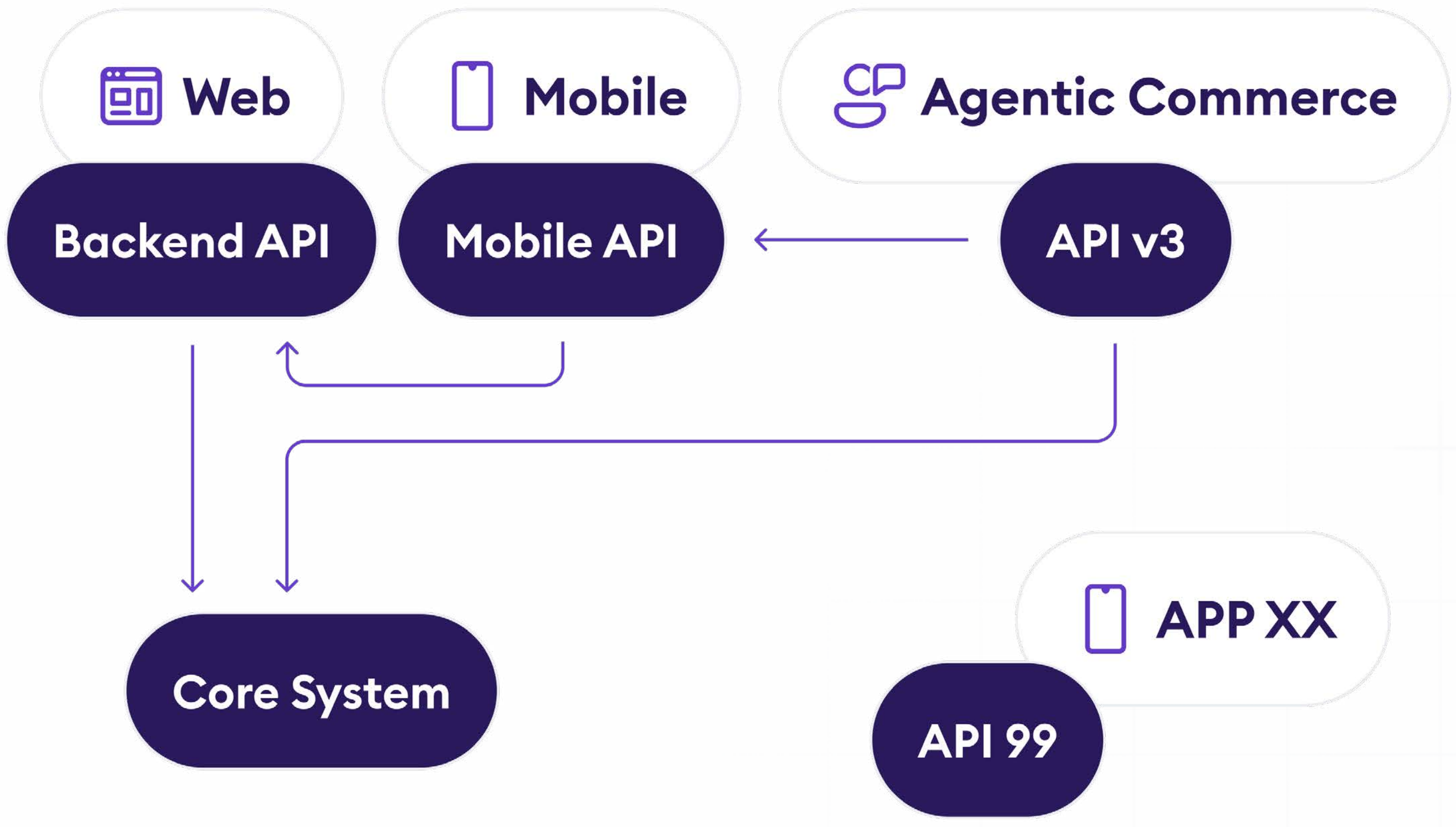
- Stabilita a verzování: změny jsou řízené a dopředu komunikované.
- Jasná definice: co API vrací, v jaké kvalitě, jak často a pro koho.
- Oddělení od aplikace: rozhraní slouží firmě, ne jednomu kanálu.
- Bezpečnost a audit: přístupy podle role, dohledatelnost.

Přínos pro AI

Když má AI stabilní rozhraní pro aktuální stav a ověřené definice, nevznikají obcházení systémů, improvizace a “ruční lepení”. To je rozdíl mezi pilotem a škálováním.

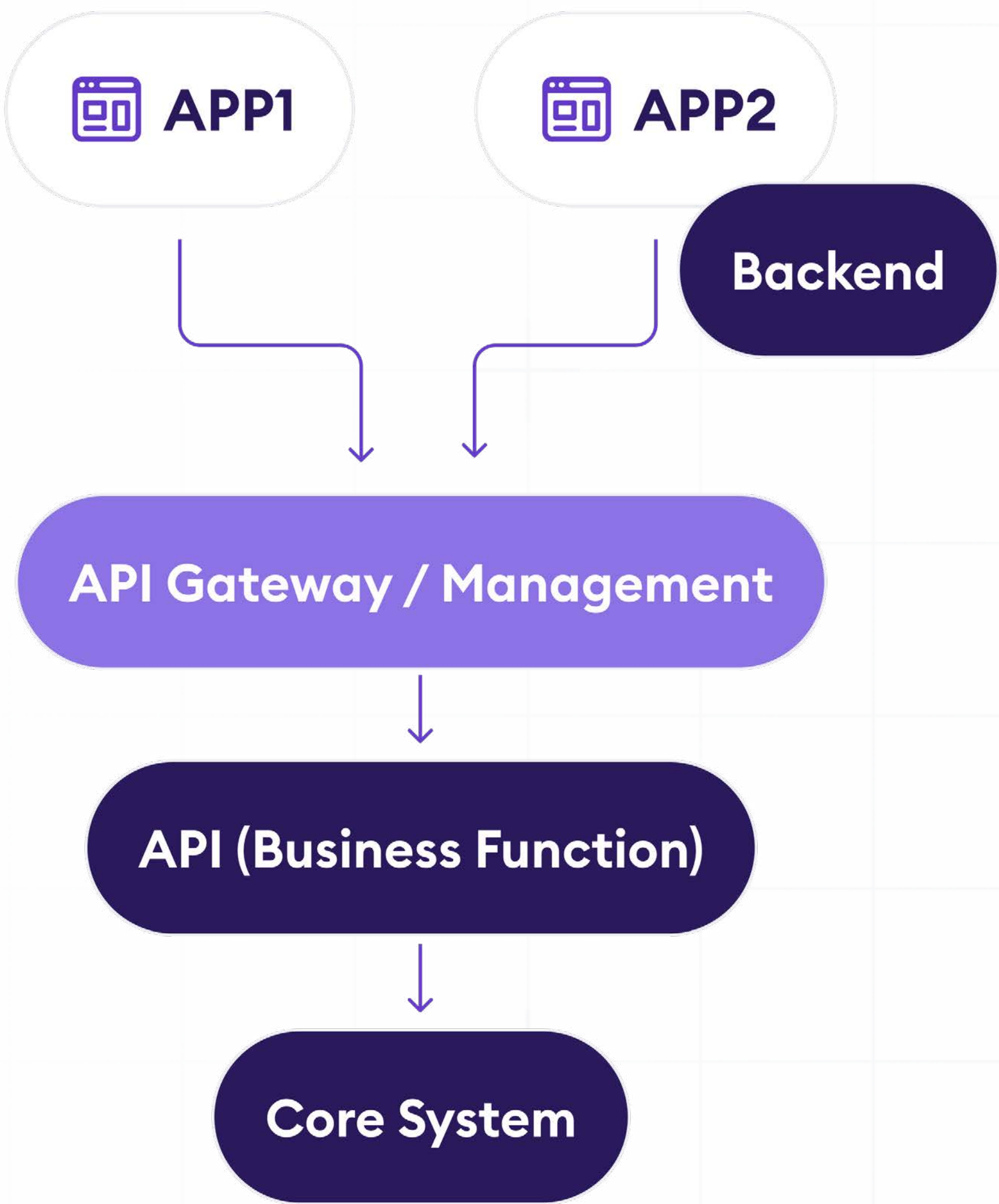
Špatně

API pevně svázané s produktem

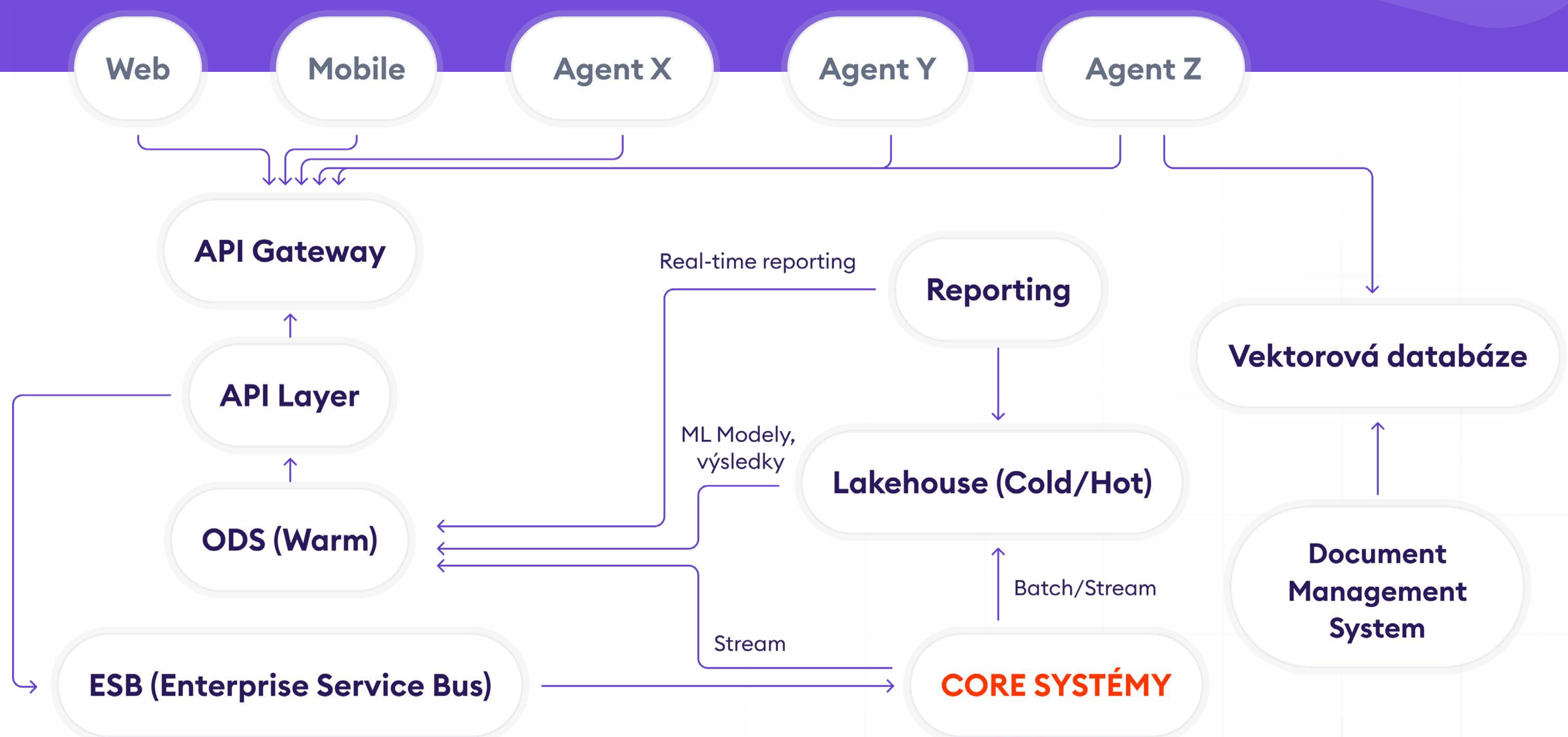


Dobře

API jako samostatný kontrakt



Vzorová AI-ready data architektura



Referenční architektura odděluje provozní systémy od vrstev pro rychlé dotazy (ODS), analytiku a historii (lakehouse) a bezpečné zpřístupnění dat a dokumentů přes API a vyhledávání v podkladech.

Jak architekturu číst

1. Core systémy jsou zdroj provozních dat (objednávky, zákazníci, ceny...).
2. Data se dostávají do datové platformy dvěma režimy: dávkově (batch) nebo průběžně (stream).
3. Vznikají dvě hlavní vrstvy pro použití dat:
 - Lakehouse (Cold/Hot) pro analytiku, historii, ML výpočty a širší datové zpracování.
 - ODS (Warm) pro rychlejší „téměř aktuální“ dotazy nad provozním stavem bez zátěže core systémů.
4. API vrstva zpřístupní data stabilně a bezpečně pro různé konzumenty (web, mobil, agenti).
5. Pro práci s dokumenty se používá Document Management System a pro vyhledávání relevantních pasáží Vektorová databáze.
6. Reporting čerpá z lakehouse (a případně z real-time vrstvy) a poskytuje pohledy pro řízení firmy.

Vysvětlení jednotlivých bloků (co je co)

- **Core systémy:** provozní aplikace (ERP/CRM/SAP apod.), kde vznikají primární data.
- **ESB (Enterprise Service Bus):** integrační vrstva pro přenos dat a událostí mezi systémy (propojování a orchestrace toků).
- **ODS (Warm):** operativní mezivrstva pro rychlé dotazy nad „téměř aktuálním“ stavem (kompromis mezi rychlostí a jednoduchostí). Typicky pro dotazy typu „co se děje teď“ bez čekání na dávky.
- **Lakehouse (Cold/Hot):** centrální datová platforma pro analýzy, historizaci, konsolidaci dat, trénování a vyhodnocování modelů. V praxi v sobě kombinuje „analytický sklad“ i datové jezero.
- **Reporting:** dashboardy, KPI, operativní přehledy. U vybraných scénářů může pracovat i s „rychlejší“ vrstvou (real-time reporting).
- **API Layer + API Gateway:** způsob, jak data a funkce poskytovat konzumentům (aplikace, agenti) stabilně, verzovaně a s kontrolou přístupů. Gateway typicky řeší autentizaci, limity, audit, řízení přístupu.
- **Document Management System:** zdroj interních podkladů (směrnice, smlouvy, postupy).
- **Vektorová databáze:** zajišťuje rychlé vyhledání relevantních částí dokumentů pro AI (např. pro odpovědi s citacemi).
- **Web / Mobile / Agent X/Y/Z:** různí konzumenti – uživatelské aplikace i AI agenti, kteří využívají data a dokumenty přes řízené rozhraní.



Co si z toho vzít

- Ne všechna data mají stejnou „teplotu“ a stejné nároky na rychlost.
- Proto diagram odděluje vrstvu pro rychlé operativní dotazy (ODS) a vrstvu pro historii/analytiku/ML (lakehouse).
- AI agenti by neměli „lézt“ přímo do core systémů.
- Datová platforma je primárně pro data (sjednocení, kvalita, historie). U složitějších scénářů je běh agentů a orchestrace vhodnější v aplikační vrstvě – typicky lépe nákladově i provozně.
- Bezpečný a škálovatelný přístup jde přes schválené zdroje (ODS/ lakehouse/dokumenty) a přes API vrstvu.
- Dokumenty jsou stejně důležité jako strukturovaná data.
- Aby AI odpovídala správně, musí čerpat z aktuálních a schválených podkladů – a umět je dohledat (vektorová DB + citace).

Nezbytnosti ke škálování

Co musí existovat, aby se to dalo škálovat

- ✓ Jednotné definice top KPI
- ✓ Katalog a popisy dat (aby šla najít a správně použít)
- ✓ Jasní vlastníci domén a datových produktů
- ✓ Přístupy podle citlivosti + audit
- ✓ Hlídání kvality u kritických dat
- ✓ Měření kvality AI odpovědí (kontrolní otázky + průběžné testy)

99

Key takeaways

“Nejdřív důvěra v čísla. Potom automatizace a AI ve velkém.”

Artefakt / Owner	Byznys	Data	IT	Bezpečnost
KPI definice	●	○		
Katalog	○	●	○	○
SLA	○	●	○	
Přístupy		○	○	●
Stopa původu (Lineage)		●	○	○
Měření AI kvality	○	●	○	○

● = primární odpovědnost (Accountable) ○ = spoluodpovědnost (Consulted/Contributing)
Cílem je, aby každá klíčová část měla jasného vlastníka – a nebyla to „zodpovědnost všech a nikoho“.

Proč je to důležité

Bez této mapy odpovědností se děje typicky:

- ⊖ KPI definuje každý jinak (byznys nedrží definice),
- ⊖ data tým je úzké hrdlo (všechno “spadne” na něj),
- ⊖ bezpečnost se řeší pozdě (a pak blokuje škálování),
- ⊖ AI piloty jsou křehké, protože neví, na co se spolehnout.

Jak AI bezpečně pracuje s firemními podklady

Typické use-casy ve firmě

- Interní směrnice a postupy:** HR, onboarding, interní pravidla, procesní manuály
- Smlouvy a právní dokumenty:** vyhledání relevantní části, shrnutí, kontrolní checklist
- Produktové a technické podklady:** FAQ pro obchod a support, návody, incident postupy
- Provozní znalosti:** runbooky, incident management, eskalace

6 pravidel bezpečné AI nad firemními podklady

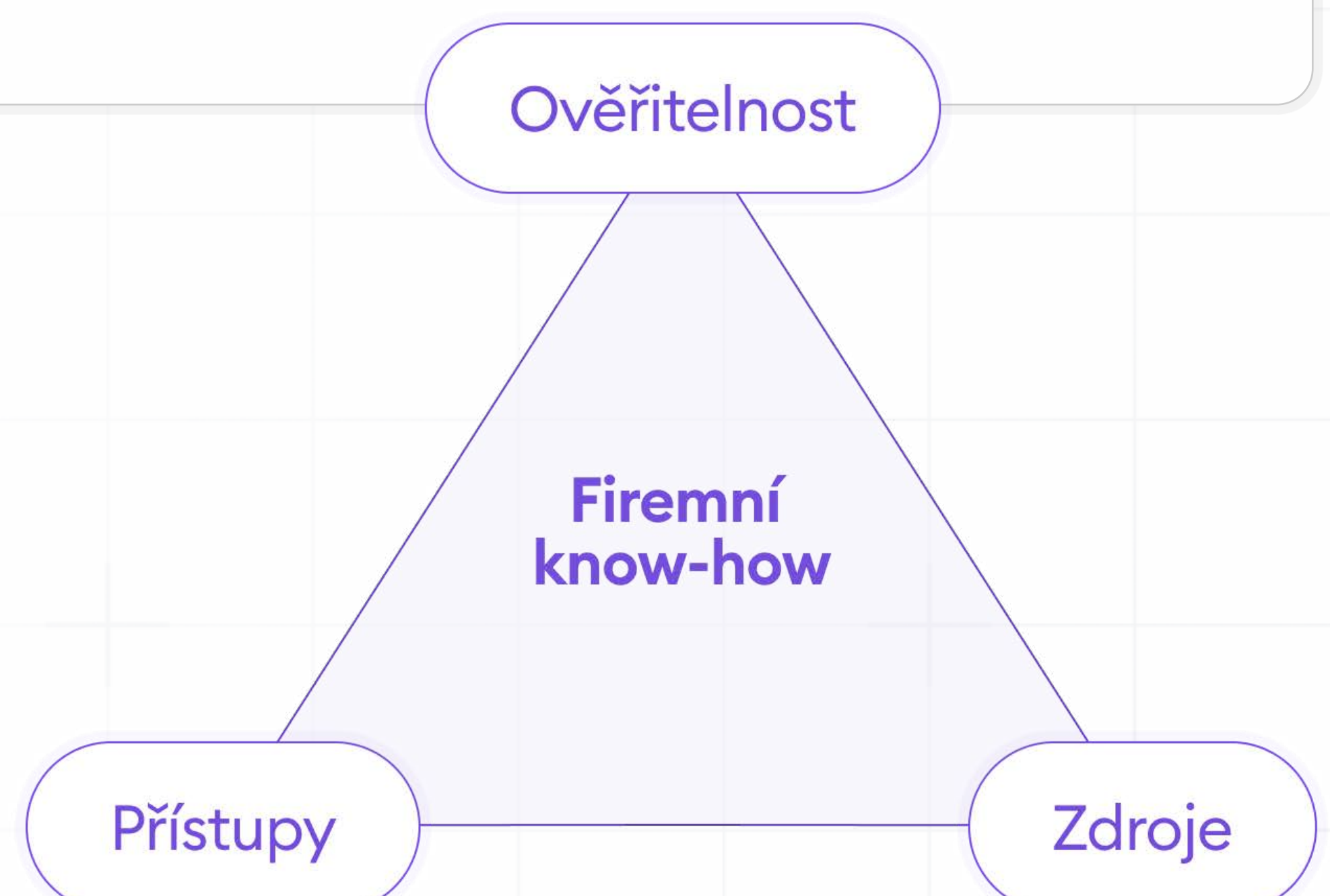
- 1. Pouze schválené zdroje**
(jasně určené “zdroje pravdy”)
- 2. Vlastník obsahu** (kdo odpovídá za správnost a aktualizaci dokumentu)
- 3. Přístupy podle citlivosti**
(AI nesmí ukázat víc než uživatel smí vidět)
- 4. Ověřitelnost odpovědi**
(citace zdrojů u odpovědi)
- 5. Práce s aktuální verzí** (verze, platnost, datum poslední aktualizace)
- 6. Audit a dohledatelnost**
(kdo se ptal na co a z čeho AI čerpala)

Hlavní message

“V praxi není největší riziko samotná AI, ale kvalita podkladů: zastaralé dokumenty, nejasné „zdroje pravdy“ a chybějící vlastník obsahu. Když firma nehlídá verze a platnost, AI může odpovědět správně podle špatného dokumentu – a tím rychle ztratit důvěru uživatelů.”

Pro data/AI tým

Tento přístup se často realizuje metodou RAG (vyhledání relevantních pasáží + odpověď s citacemi)



Tok zpracování podkladů

Tok od podkladů k AI odpovědi s citací

- 1 Vybereme schválené dokumenty
- 2 Připravíme je pro vyhledávání (struktura, popisy)
- 3 Najdeme relevantní pasáže
- 4 AI odpoví a uvede, z čeho čerpala
- 5 Měříme kvalitu a průběžně zlepšujeme



Definition

Citace zdrojů = AI u odpovědi ukáže, které dokumenty/pasáže použila.



Pro data/AI tým

- kombinované vyhledávání (klíčová slova + sémantika)
- filtrace podle přístupů a čerstvosti
- přerovnání výsledků pro relevanci
- citace + omezení kontextu

Jak měřit kvalitu AI

Jak měřit kvalitu AI odpovědí (bez dojmů)

- 1. Vytvořte **kontrolní sadu otázek z praxe** (20–50)
- 2. Hodnoťte tři věci:
 - Relevance** (odpověď řeší otázku)
 - Ověřitelnost** (je opřená o zdroj, ne “domýšlení”)
 - Bezpečnost** (je v mezích oprávnění)
- 3. Po každé změně zdrojů/procesu spusťte znovu testy
- 4. Sledujte produkční signály: eskalace, nespokojenost, chybné odpovědi

Problém	Jak poznat	Co upravit
Mimo téma	Nízká relevance	Zlepšit vyhledání, popisy dokumentů
Bez opory	Chybí citace	Posílit citace, omezit zdroje
Citlivé informace	Audit	Zpřísnit přístupy, redakce citlivých údajů
Zastaralé informace	Kontrola čerstvosti	Filtrace podle data, update procesu

Provozní měření: co sledovat po nasazení

 Míra eskalací na člověka	Klesá, ale nesmí být nulová u citlivých témat
 Zpětná vazba uživatelů	Užitečné / neužitečné
 Neúspěšné dotazy	Nenašli jsme podklady
99 Citace zdrojů	Kolik odpovědí je doložených
 Top důvody chyb	Zastaralé dokumenty, chybí definice, chybí přístup



Governance a rizika

Aby šla AI ve firmě škálovat, musí být **bezpečná, předvídatelná a obhajitelná**. Governance není byrokracie. Je to praktický rámec, který drží AI “v mantinelech” a zároveň umožní rychlé rozšiřování do dalších týmů a procesů.

3 otázky, které určují schopnost škálovatelnosti

Kdo nese odpovědnost?

Bez jasného vlastníka use-casu se AI rychle stane “zodpovědností všech a nikoho”.

Co je pro AI zdroj pravdy?

AI smí pracovat jen se schválenými daty a dokumenty – a musí být jasné, kdo ručí za jejich platnost a aktualizaci.

Jak to obhájíme?

V klíčových situacích musíme umět doložit, kdo se ptal, z čeho AI čerpala a jak jsme řešili chybu nebo incident.

Riziko není stejné pro všechny use-casy

Stejná AI může být v jednom scénáři neškodná a v jiném vysoce riziková. Proto škálování začíná jednoduchou klasifikací:

Nízké riziko	Automaticky + evidence	Interní FAQ, shrnutí postupů, orientace v dokumentech.
Střední riziko	Člověk schválí + evidence	Návrhy a doporučení (např. návrh odpovědi zákazníkovi), podpora rozhodování.
Vysoké riziko	Povinná kontrola člověkem + omezení zdrojů + eskalace + evidence	Právní nebo finanční dopad, citlivé údaje, rozhodnutí ovlivňující zákazníka.

Minimální governance:

1. Vlastník AI use-casu (accountable)
2. Schválené zdroje (data + dokumenty)
3. Přístupová pravidla podle citlivosti
4. Evidence a auditní stopa (kdo / kdy / jaké zdroje / jaký výstup)
5. Incident proces (chyba, únik, nesprávná odpověď)

Směrnice a compliance

Při škálování AI roste tlak na auditovatelnost a řízení rizik – ať už ze strany interního auditu, zákazníků, nebo regulace (např. EU AI Act). Cílem je zavést minimum, které umožní AI používat rychle, bezpečně a bez blokad.

Compliance minimum (co musí existovat)

1. **Klasifikace use-casů podle rizika** a přiřazený režim kontroly
2. **Auditní stopa** pro střední a vysoké riziko (účel, uživatel, zdroje, výstup)
3. **Ochrana citlivých údajů** (pravidla práce s osobními/citlivými daty, redakce, omezení citací)
4. **Transparentnost a eskalace** (kdy AI musí předat dotaz člověku; u vybraných scénářů jasně komunikovat použití AI)
5. **Prověření nástrojů a dodavatelů** (kde se data zpracovávají, zabezpečení, smluvní podmínky, incident management)

Praktické nastavení bez zpomalení

- Začněte 3–5 use-casy a nastavte pravidla podle rizika.
- Sepište 1stránkovou interní směrnici: co je povolené, co zakázané, kdy eskalovat.
- Evidenci a audit držte jen tam, kde je to potřeba (střední a vysoké riziko).

Porušení EU AI Act pravidel mohou znamenat sankce v podobě:

- až **7 %** z ročního celosvětového obrátu firmy
- fixně **35 mil. EUR** za zakázané AI aplikace



HAVEL & PARTNERS
ÚSPĚCH SPOJUJE

Ve spolupráci s největší česko-slovenskou advokátní kanceláří HAVELS & PARTNERS zajišťujeme kompletní audit a zavedení praktik v souladu s regulacemi.

[Více zde](#)

Řidte data jako produkt

Data nejsou jednorázový výstup (report), ale služba, která musí být spolehlivá a opakovaně použitelná. Proto se k nim má přistupovat stejně jako k produktu.

Datový produkt = data, která mají vlastníka (ownership) a standardy

4 principy:

- 1. Má uživatele (kdo to používá)
- 2. Má parametry kvality a dostupnosti (SLA)
- 3. Má jasné definice (KPI)
- 4. Má vlastníka a plán zlepšování

Datový produkt v praxi (příklad)

Datový produkt:	Tržby (oficiální)	Doména:	Finance & Sales	Vlastník:	CFO, Controlling, Obchod, CEO dashboard, forecasting
Vlastník:	Finance Data Owner (Controlling lead)	Delivery tým:	Data platform / Analytics engineering		
Co to obsahuje (výstup):	Denní a měsíční tržby, členění podle kanálu, produktu, regionu Základní dimenze: zákazník, produkt, kanál, region				
Definice:	Tržby = fakturované tržby bez DPH Storno/vratky = odečtené v měsíci zaúčtování Pozn.: objednávky ≠ tržby (oddělená metrika)				
Spolehlivost (SLA):	Profil: Cold (dle potřeby) Čerstvost: denně do 08:00 (D-1) Dostupnost: 99,5 % (měsíčně) Kvalita: automatické kontroly (duplicitní doklady, chybějící region, záporné hodnoty mimo vratky)				
Dohledatelnost původu (odkud se to bere):	Zdroj: ERP fakturace + CRM (kanál) Transformace: standardní výpočty v datové vrstvě (link do katalogu) “Zdroj pravdy” dokument: definice KPI (link)				
Přístupy (citlivost):	Standard: CFO/Finance/Controlling + vybrané role Omezení: detail na úroveň zákazníka pouze pro schválené role Audit: log přístupů a změn				
Změny a verze:	Změny definic schvaluje vlastník (Finance Data Owner) Každá změna má verzi + datum + dopad na reporty/AI Komunikace: release note do #data-updates / e-mail stakeholderům				

Jak varování nastavit

Začněte u dat, která mají největší dopad na rozhodování a AI.

1 Vyberte “kritická data” (20 % s největším dopadem)

- Top 5–10 KPI a reportů, které používá vedení
- 1–2 datové toky, které napájí nejdůležitější procesy
- (Pokud používáte AI nad podklady) 10–20 klíčových dokumentů, které jsou “zdroj pravdy”

2 Ke každému kritickému toku nastavte 3 základní typy varování

1. **Čerstvost** – data přišla pozdě / nepřišla
 2. **Objem/anomálie** – přišlo výrazně méně/více než obvykle
 3. **Změna struktury** – změnil se pole/formát
- Pro real-time navíc
4. **Zpoždění doručení událostí** – přichází pozdě / nepřichází
 5. **Validace událostí** – neúplná/duplicitní/nereálná data

3 Nastavte prahy tak, aby to nebyl spam

- Začněte jednoduchými pravidly typu: “později než obvykle”, “výrazná odchylka”, “změna struktury”
- V první verzi raději méně upozornění, ale relevantních
- Po 2–3 týdnech prahy doladíte podle reálných falešných poplachů

4 Určete, co se stane po upozornění

Každé varování musí mít:

- vlastníka (kdo řeší)
- cílový čas reakce (např. do 1 hod / do 1 dne)
- jednoduchý postup (ověřit zdroj → ověřit transformaci → informovat uživatele)

5 Ukažte dopad uživatelům

Když dojde k incidentu, uživatel má vědět:

- co je ovlivněno (které KPI/report/AI odpovědi)
- kdy bude opraveno
- zda se změnila definice nebo jen došlo k výpadku

Roadmapa: 90 dní + 12 měsíců

AI-ready data za 90 dní

Týdny 1–2	Fokus + ROI strategie	Vybrat top 3-5 use casů dle ROI
Týdny 3–6	Důvěra v čísla	Definice KPI, vlastníci, katalog
Týdny 7–10	Datové produkty	1–2 produkty se standardy a hlídáním kvality
Týdny 11–13	AI pilot	práce se schválenými zdroji + měření kvality

Roadmapa 90 dní

Týdny 1–2: ROI strategie

Výstupy:

- Seznam **3–5 prioritních use-casů** (kde AI reálně ušetří čas / sníží riziko / zvýší rychlost).
- Posudte jaké use casy má největší potenciál návratnosti (ROI) a těmi začněte
- Vybraná **1 doména** pro start (např. HR, compliance, customer support, produkt).
- Definované **5–10 klíčových KPI** (pokud jde o čísla) nebo „**zdroje pravdy**“

Vedení rozhodne:

Kdo je **vlastník use-casu** (owner) a kdo přiděluje konkrétní zdroje (lidi, čas a priority)

Týdny 3–6: Důvěra v data a společný jazyk

Výstupy:

- Pro vybranou doménu vznikne **jednotná definice top KPI** nebo seznam schválených dokumentů (“zdroje pravdy”).
- Základní **katalog**: kde data/dokumenty jsou, kdo je vlastní, jaká je platnost/verze.
- Zavedeno minimum přístupů **podle citlivosti** (kdo co smí).
- První “stopa původu” pro 3–5 nejdůležitějších výpočtů (odkud se číslo vzalo).

Milník:

Management vidí 1 stránku: co je zdroj pravdy, kdo to vlastní, co je definice KPI.

Týdny 7–10: 1–2 datové produkty a hlídání kvality

Výstupy:

- 1–2 **datové produkty** (ne reporty): mají vlastníka, popis, parametry čerstvosti a kvality.
- Základní **hlídání kvality** pro kritické toky (čerstvost, anomálie, změny struktury).
- Jednoduchý proces změn: když se změní definice KPI nebo zdroj, je jasné **kdo schvaluje** a jak se to komunikuje.

Milník:

“Když se něco rozbije, dozvíme se to včas a víme, kdo to řeší.”

Týdny 11–13: AI pilot s měřením a bezpečným režimem

Výstupy:

- AI pilot pro vybranou doménu (např. práce s interními podklady).
- **Kontrolní sada otázek** (20–50) + jednoduché skórování (relevance, ověřitelnost, bezpečnost, použitelnost).
- Nastavená **eskalace** pro citlivé dotazy (kdy AI předává člověku).
- Výstup: jasné rozhodnutí “škálovat / upravit / stop” na základě dat, ne pocitu.

3–12 měsíců: Škálování napříč firmou

Měsíce 3–6: Rozšíření domén a standardů

- Rozšířit jednotné definice KPI / zdroje pravdy na další domény (typicky 2–3).
- Zpřesnit přístupový model a auditní stopu pro střední a vysoké riziko.
- Zlepšit katalog tak, aby byl reálně používaný (self-service).

Měsíce 6–9: Provozní standardy a opakovatelnost

- Standardizovat datové produkty: šablona, vlastník, parametry kvality.
- Plošnější hlídání kvality u kritických dat (ne vše, ale “top dopad”).
- Zautomatizovat kontrolu, že se kvalita AI po změnách nezhoršuje.

Měsíce 9–12: Škálování AI a automatizace

- Rozšířit AI scénáře do dalších týmů (podle risk klasifikace).
- Napojit vybrané procesy na automatizaci (tam, kde je vysoká jistota a auditovatelnost).
- Průběžný cyklus zlepšování: incidenty → opravy → lepší zdroje → lepší měření.

Orientační skóre datové maturity

Každé zaškrtnutí = 1 bod. Součet = vaše skóre.

Důvěra v data

- ☐ Testy a hlídání kvality pro kritická data
- ☐ Varování na čerstvost/objem/strukturu
- ☐ Incidenty mají jasnou příčinu a prevenci

Jednotná čísla (KPI)

- ☐ Top KPI mají jednotné definice
- ☐ Změny definic jsou verzované a komunikované
- ☐ Existuje vrstva společných definic

Odpovědnosti a dohledatelnost

- ☐ Vlastníci domén jsou určení
- ☐ Původ top KPI je dohledatelný
- ☐ Katalog/popis dat se používá

Bezpečnost

- ☐ Přístupy jsou podle citlivosti
- ☐ Citlivé údaje jsou chráněné
- ☐ Audit přístupů je dohledatelný

AI připravenost

- ☐ AI pracuje jen se schválenými zdroji
- ☐ Odpovědi obsahují citace
- ☐ Existuje kontrolní sada otázek a průběžné testy

Orientační skóre datové maturity

0–5 bodů → Základy chybí (AI bude křehká a riziková)

6–9 bodů → Dá se udělat pilot, ale škálování bude narážet

10–12 bodů → Připraveno škálovat první domény/use-casy

13–15 bodů → Silná připravenost škálovat napříč firmou

AI-ready Data Assessment Workshop

Získejte jasný obraz o tom, jak jste na tom s AI-ready daty dnes, co vás nejvíc brzdí a jaké kroky přinesou nejrychlejší výsledky. Výstupem získáte prioritizovanou roadmapu na míru, kterou lze okamžitě převést do realizace.

Co je součástí workshopu:



Posouzení datové maturity

v klíčových oblastech: důvěra v data, jednotné KPI, dohledatelnost, přístupy, připravenost pro AI



Doporučení rychlých výher pro 1–2 domény nebo use-casy



Mapa rizik pro škálování AI (bezpečnost, audit, nekonzistence, zastaralé podklady)



Prioritizovaná roadmapa na 90 dní + konkrétní backlog kroků (lidé / procesy / technologie)



Doporučený režim řízení (klasifikace rizik a pravidla kontroly podle typu use-casu)

Pro koho je workshop vhodný (Toto jsou orientační kritéria.)

Největší hodnotu typicky přináší organizacím, které:

- mají více týmů a systémů (data nejsou “na jednom místě”),
- řeší jednotnost KPI, dohledatelnost a přístupy napříč firmou,
- chtějí škálovat AI mimo pilotní fázi (do procesů, provozu nebo více oddělení),
- pracují s citlivými daty nebo podléhají interním auditům.

Jak to probíhá:

- strukturovaný workshop se stakeholdery (byznys, data/IT, bezpečnost)
- rychlé zmapování současného stavu, priorit a rizik
- společná shoda na “zdrojích pravdy”, KPI a režimu kontroly
- na konci jasný plán a výstupy připravené k realizaci

Omezená kapacita a výběr účastníků:

Abychom udrželi kvalitu a věnovali se každé organizaci do hloubky, přijímáme **omezený počet** workshopů a zájemce **vybíráme**. Před potvrzením spolupráce si uděláme krátké úvodní posouzení vhodnosti (rychlý call), abychom ověřili, že workshop přinese vysokou hodnotu.

Domluvit úvodní konzultaci



Založeno 2016 | 75+ expertů | 500+ projektů | 100+ enterprise klientů

bighub