PAY-FAIR

# Technická a bezpečnostní specifikace řešení

## Průzkum Pay-Fair 2024

# 1 JAKÉ ŘEŠENÍ VYUŽÍVÁME

Pro realizaci průzkumu využíváme datovou platformu **Rubik** od naší partnerské společnosti Vissto s.r.o. Na této stránce naleznete popis této platformy, technickou a bezpečnostní specifikaci celého řešení.

Rubik představuje unikátní řešení, které kombinuje **MS Excel** jako nástroj pro snadnou a rychlou úpravu dat **s cloudovým databázovým úložištěm**, které dokáže data ihned synchronizovat mezi libovolný počet uživatelů na straně klienta a zabezpečit je pomocí uživatelských přístupů a rolí. Díky tomu umožňuje tato platforma využívat veškeré výhody nejrozšířenější aplikace pro práci s daty MS Excel (funkce, analytické možnosti, jednoduchost, přehlednost), ale zároveň odstraňuje všechna jeho omezení (sdílení a synchronizace dat mezi více uživateli, zabezpečení konkrétních záznamů na úrovni řádků či tabulek, výpočetní výkon v případě rozsáhlých kalkulací). Žádná data nebudete v průběhu tohoto průzkumu posílat mailem, vše budete nahrávat prostřednictvím tohoto zabezpečeného nástroje.

# 2 TECHNICKÉ POŽADAVKY NA PROVOZ SYSTÉMU

Pro úspěšný a bezproblémový chod tohoto systému je potřeba zajistit na straně klienta:

- **OS Windows s MS Excel verze MS Excel 365** (jiné verze MS Excel k ověření)
- **Povolená makra** (lze navázat na Code-Sign certificate)
- V případě, že nelze zajistit, můžeme pro Vás připravit virtuální stanici (PC), ke které se lze připojit vzdáleně. V případě zájmu nás neváhejte kontaktovat

Více podrobností je rozepsáno v odstavcích níže.

# 3 ARCHITEKTURA A ZABEZPEČENÍ RUBIK

Z pohledu softwarové architektury se Rubik dělí na frontend (soubor MS Excel na zařízení klienta) a backend (kolekce propojených cloudových služeb v rámci jednoho poskytovatele).

## 3.1 Backend - databázové rozhraní

- Infrastruktura je kompletně hostována v **certifikovaném cloudovém prostředí společnosti Microsoft Azure**.
    - Toto prostředí je vybaveno robustní ochranou proti různým typům kybernetických útoků a útočných vzorů chování.

    - Veškeré naše infrastrukturní a aplikační komponenty jsou situovány v rámci EU v nizozemském datacentru (West Europe), abychom zajistili soulad s GDPR směrnicí. Z bezpečnostních důvodů společnost Azure nezveřejňuje přesné geografické lokace svých datových center.
- Klientská data jsou ukládána v **zabezpečené a spravované databázi** ("managed database").
    - Toto řešení nabízí automatizované zálohování a funkci Point-in-Time Restore (PITR), která umožňuje obnovu dat v časovém horizontu 7 dnů.
    - Veškerá manipulace s daty formou čtení, zápisu, aktualizace či mazání (CRUD) probíhá výhradně uvnitř samotné databáze, bez nutnosti volání externích služeb či aplikací třetích stran.
- Komunikace mezi frontendem a backendem probíhá přes **šifrované HTTPS spojení**.
    - Přímý přístup k datům v DB není nikdy pro klienty možný, pouze pomocí  frontend Rubik, který data stahuje na bázi přístupových práv skrze API.
- Komunikace mezi API endpointem a DB probíhá pouze přes **privátní virtuální síť (VPC)**.
    - V rámci autentizace využíváme validační přístupové tokeny (JWT), které uživatel dostává při každém přihlášení do aplikace.
    - Z důvodu bezpečnosti mají vydané tokeny nastavenou časovou expiraci.
- Hesla uživatelů ukládáme pouze v **šifrované podobě** - uložení v nešifrované podobě není možné.
- O řízení přístupů uživatelů se starají ověřené interní **RBAC mechanismy**.
    - Přístupy řídíme na bázi přístupových práv a skupin na úrovni jednotlivých business objektů/tabulek i konkrétních řádků pomocí RLS.
- Z důvodu bezpečnosti jsou klientská data vždy ukládána do **striktně separovaných schémat** v rámci databázového systému.
- Cloudová infrastruktura poskytuje **možnost škálování výpočetních zdrojů** při výpočetní či datové zátěži.
- V rámci průzkumu Pay-Fair může klient do systému nahrát i **pseudonymizovaná data**.
    - Nahrávání osobních informací (jméno, příjmení) je čistě volitelné a slouží pouze pro snazší orientaci v připravených reportech.

## 3.2 Frontend - Excelový nástroj

- Uživatelské rozhraní pro práci s daty je integrováno přímo v aplikaci MS Excel.
    - Kompatibilita je optimalizována pro **OS Windows a verze MS Excel 365**.
    - Kompatibilitu s jinými verzemi MS Office lze ověřit na vyžádání.
- V nástroji je připraveno kompletní „klikatelné“ uživatelské rozhraní.

- **Uživatel nepotřebuje mít žádné dovednosti s prací v Excelu** - vše je připraveno a plně automatizováno.
- Uživatel vše ovládá pomocí uživatelských tlačítek a rozhraní.

- **Pro práci s nástrojem je nutné zajistit povolení VBA maker.**
  - Povolení maker lze navázat mj. na náš Code-Sign certifikát.
- Data jsou kompletně uložená v cloudu a stahují se na bázi **přístupových práv** (systém oěřovacích tokenů).
  - Žádný uživatel nemá přímý přístup do databáze.
  - Data lze stahovat z databáze až na základě úspěšného loginu.
- Komunikace s backendem probíhá pomocí API requestů.
  - Komunikace s databází probíhá skrze **zabezpečené šifrované spojení na bázi HTTPS protokolu**.
- V případě, že klient nebude mít možnost zajistit odpovídající OS, verzi Excelu nebo povolení maker, **můžeme pro něj zřídit virtuální stanici (PC)**, ke které se může připojit pomocí služby vzdáleného připojení (Remote Desktop - app/browser).
  - Tuto možnost nabízíme za příplatek a je nutné se nejprve domluvit na přesném rozsahu a možnostech. V případě zájmu nás neváhejte kontaktovat

## 3.3 Volitelný AI modul

V rámci snahy o maximální automatizaci a úsporu času na straně klienta nabízíme k využití náš volitelný, automatizační AI modul integrovaný přímo v Excelovém nástroji Rubik. **Tento modul pomáhá zejména se zařazením pozic klienta do připraveného katalogu**.

V rámci tohoto modulu využíváme zabezpečené služby třetích stran, které uvádíme níže. Do těchto služeb **nikdy neposíláme žádná citlivá data**, pracujeme pouze se základními informacemi, které nemají povahu osobních informací.

- Využíváme jazykové modely (ChatGPT, GPT4,...) od **OpenAI** ve firemní API verzi (nevyužívá se pro trénování a analýzu dat: https://trust.openai.com/).
  - Do tohoto modelu nikdy neposíláme žádná citlivá data – pracujeme pouze se základními informacemi viz níže a dále již jen s výstupy, které tímto způsobem vytváříme.
  - Všechna data, která používáme pro potřeby AI dokážeme jednoznačně identifikovat – abyste měli přehled, jaké informace odchází.
    - Název pozice
    - Název organizačních jednotek
    - Základní popisem organizace
- Pro automatizované překlady dat využíváme **DeepL** API Pro, splňující nejvyšší bezpečnostní normy (TLS encryption a okamžité mazání přeložených dat: https://www.deepl.com/cs/pro-data-security/)