Metodika hodnocení nominovaných pracovních procesů pro pilotní projekt AI

Pracovní podklad pro výběr kandidátů pilotního projektu AI v pracovním procesu ČTÚ

1. Účel hodnocení

Cílem hodnocení je z 19 nominovaných pracovních procesů jednotným a transparentním způsobem identifikovat nejsilnější kandidáty pro pilotní projekt AI v pracovním procesu úřadu.

Hodnocení má pracovní skupině pomoci porovnat velmi rozdílné procesy podle stejných hledisek, nikoli je vybírat pouze podle prvního dojmu, osobní preference nebo jediného parametru.

Každý proces je posuzován podle pěti samostatných kritérií. Každé kritérium hodnotí jiný aspekt procesu a získává 1–3 hvězdičky:

- ★ nízké / omezené hodnocení
- ★★ střední / významné hodnocení
- ★★★ vysoké / velmi významné hodnocení

Hvězdičky slouží jako pomůcka pro vzájemné porovnání nominací. Nenahrazují odbornou diskusi a finální rozhodnutí pracovní skupiny.

Vedle hodnocení se u jednotlivých procesů sledují také:

- četnost využití, pokud je z nominace známa,
- charakter dat – veřejná / interní / diskrétní / kombinace.

Tyto informace nejsou součástí hvězdičkového hodnocení. Mohou být zohledněny při finálním rozhodování mezi nejsilnějšími kandidáty.

2. Základní princip hodnocení

Každé kritérium hodnotí jednu konkrétní věc. Silná nebo slabá stránka procesu v jednom kritériu se proto nepřenáší do jiného.

Proces například může využívat pouze několik specialistů → ★ za rozsah využití, ale těmto specialistům přinést zásadní pomoc → ★★★ za přínos, být výborným AI use casem → ★★★ za vhodnost pro AI a současně být příliš široký pro jeden pilot → ★ za vhodnost pro pilot.

To je v pořádku a je to právě smyslem více kritérií.

Pět otázek, které si při hodnocení klademe

1. Jak moc to pomůže ČTÚ?

2. Kolika lidem to může pomoci?

3. Je to opravdu práce pro AI – a jak moc? Jakou roli tady AI hraje?

4. Neukousli jsme si pro jeden pilot příliš velké sousto – rozsahem, komplexností nebo náročností přípravy?

5. Jak po pilotu poznáme, že to opravdu pomohlo?

3. Kritérium 1 – Význam očekávaného přínosu?

Kritérium 1: Jak moc to pomůže ČTÚ?

Co hodnotíme

Hodnotíme, jak významně by využití AI mohlo zlepšit současnou práci nebo poskytovanou službu.

Přínosem nemusí být pouze úspora času. Může jít například o:

- snížení manuální práce,
- rychlejší zpracování,
- snazší získávání potřebných informací,
- zvýšení kvality nebo úplnosti výstupu,
- podporu odborné práce,
- snížení rutinní zátěže,
- zvýšení množství práce, které lze zvládnout za stejnou dobu,
- zlepšení služby poskytované veřejnosti.

Nehodnotíme zde, zda už dnes dokážeme přínos přesně změřit. Měřitelnost řeší samostatné kritérium 5.

★ Omezený přínos

AI přinese spíše dílčí zlepšení současného způsobu práce nebo služby. Pomoc je užitečná, ale netýká se podstatné části pracovní potřeby.

Příklad: AI pomáhá upravovat nebo shrnovat jednotlivé texty, které pracovník již dnes zpracovává bez významnějších obtíží.

Proč: AI práci usnadní, ale celkový dopad na způsob práce je omezený.

★★ Významný přínos

AI přinese znatelné zlepšení současného způsobu práce nebo služby. Může významně snížit manuální práci, urychlit část procesu, usnadnit práci s informacemi nebo zvýšit kvalitu výstupu.

Příklad: AI pravidelně připravuje strukturovaná shrnutí rozsáhlejších podkladů, které dnes pracovníci musí pročítat a zpracovávat ručně.

Proč: Přínos je pro pracovníka prakticky významný a týká se podstatné části jeho práce.

★★★ Velmi významný přínos

AI může přinést výrazné zlepšení současného způsobu práce nebo poskytované služby. Může například zásadně zkrátit časově náročnou činnost, odstranit významnou část manuální práce, výrazně usnadnit obtížně zvládnutelnou práci s informacemi nebo významně zlepšit kvalitu či dostupnost služby.

Příklad: Pracovník dnes ručně prochází velké množství dokumentů, aby nalezl relevantní předchozí případy. AI může významně urychlit jejich vyhledání a prvotní porovnání.

Proč: Přínos se týká samotné podstaty pracovní potřeby a může významně změnit způsob, jakým je dnes práce vykonávána.

Rychlá orientace

★	Dílčí pomoc
★★	Znatelné zlepšení
★★★	Výrazné zlepšení

Kritérium 2: Rozsah využití
Kolika lidem to může pomoci?

Co hodnotíme

Hodnotíme šíři okruhu potenciálních uživatelů řešení.

Uživateli mohou být:

- jednotliví specialisté,
- pracovníci jednoho útvaru,
- více útvarů,
- větší část úřadu,
- konkrétní skupina externích uživatelů,
- široká veřejnost.

Důležité

Rozsah využití není totéž jako četnost využití.

Řešení určené třem specialistům, kteří jej používají každý den, má stále omezený rozsah uživatelů. Jeho každodenní používání ale zaznamenáme samostatně a jeho význam se může projevit například v kritériu 1.

★ Omezený rozsah

Řešení je určeno malému okruhu uživatelů, například několika specialistům nebo jednomu úzce vymezenému týmu.

Příklad: Řešení využívají tři specialisté jednoho útvaru.

★★ Významný rozsah

Řešení je určeno širší skupině pracovníků, více týmům nebo útvarům, případně vymezené skupině externích uživatelů.

Příklad: Řešení může využívat několik desítek pracovníků napříč více útvary.

★★★ Velmi významný rozsah

Řešení může využívat velmi široký okruh uživatelů nebo veřejnost.

Příklad: AI vyhledávač informací na webu úřadu určený široké veřejnosti.

Rychlá orientace

★	Malý okruh uživatelů
★★	Širší okruh uživatelů
★★★	Velmi široký okruh uživatelů / veřejnost

Četnost využití uvádíme samostatně, například: denně / týdně / měsíčně / několikrát ročně / dle potřeby.

Kritérium 3 – Vhodnost pro využití AI

Je to opravdu práce pro AI – a jak moc? Jakou roli tady AI hraje?

Co hodnotíme

Hodnotíme, nakolik charakter požadované práce odpovídá tomu, v čem může AI přinést specifickou přidanou hodnotu.

Sledujeme zejména práci s obsahem a významem informací, například:

- vyhledávání relevantních informací,
- shrnování,
- získávání informací z dokumentů,
- klasifikaci a kategorizaci,
- porovnávání,
- práci s kontextem,
- hledání souvislostí,
- syntézu informací z více zdrojů,
- přípravu nového obsahového výstupu.

Zásadní pravidlo

Více hvězdiček neznamená technicky složitější řešení.

Hodnotíme roli AI při řešení pracovní potřeby, nikoli technickou proveditelnost, konkrétní model, aplikaci nebo technologii.

★ Omezená vhodnost – AI doplňuje proces

AI může pomoci, ale práce s významem a obsahem informací není hlavní podstatou pracovní potřeby. Významná část procesu je založena na předem stanovených pravidlech, rutinních operacích nebo přesunech informací.

Příklad: Z dokumentu se získá několik předem stanovených údajů a následně se přenesou do jiného systému.

Proč: AI může pomoci s rozpoznáním údajů, ale podstata procesu spočívá především v rutinním zpracování.

★★ Významná vhodnost – AI zpracovává obsah

AI má v procesu jasnou a významnou roli a řeší konkrétní, dobře vymezenou obsahovou úlohu. Typicky může jít o shrnutí dokumentu, klasifikaci obsahu, získání konkrétních informací, práci s jasně vymezeným zdrojem nebo vytvoření návrhu textu na základě zadaných podkladů.

Příklad: AI přečte rozsáhlý dokument a připraví jeho strukturované shrnutí.

Proč: AI musí porozumět významu textu, ale pracuje s jasně daným vstupem a vytváří jednoznačně vymezený výstup.

Další příklad: AI přečte dokument a podle jeho obsahu navrhne zařazení do jedné z předem definovaných kategorií. AI pracuje s významem obsahu, ale požadovaný výsledek je jasně vymezen.

★★★ Velmi významná vhodnost – AI propojuje a vyhodnocuje obsah v kontextu

AI musí interpretovat potřebu, pracovat s významem a kontextem informací, vyhledávat relevantní obsah, porovnávat nebo propojovat informace a na jejich základě připravovat nový užitečný výstup.

Příklad: Pracovník potřebuje zjistit, zda byl v minulosti řešen obdobný spor a jakým způsobem. Vyhledává ve velkém množství předchozích dokumentů podle významu a kontextu, pracuje s více různými zdroji a současně potřebuje kontrolovat připravovaný text ve vztahu k relevantní platné legislativě.

Proč: AI musí pracovat s kontextem konkrétního případu, významově vyhledávat, porovnávat informace z různých zdrojů a pracovat se souvislostmi.

AI zde podporuje práci odborníka; samotné odborné posouzení, rozhodnutí a odpovědnost zůstávají člověku.

Další příklad – AI vyhledávač: Uživatel položí otázku vlastními slovy. AI interpretuje jeho potřebu, významově vyhledá relevantní informace a z nalezeného obsahu připraví odpověď.

Rychlá orientace

★	AI doplňuje proces
★★	AI zpracovává obsah
★★★	AI propojuje a vyhodnocuje obsah v kontextu

Kritérium 4 – Vhodnost procesu pro pilot z hlediska rozsahu a komplexnosti

Neukousli jsme si pro jeden pilot příliš velké sousto – rozsahem, komplexností nebo náročností přípravy?

Co hodnotíme

Hodnotíme rozsah a komplexnost samotného pracovního procesu a také to, zda jej lze v přiměřeném čase připravit k pilotnímu ověření.

Sledujeme například:

- počet různých činností,
- počet variant a rozhodovacích bodů,
- množství a různorodost vstupů a výstupů,
- návaznosti na další části pracovního procesu,
- počet pracovních potřeb zahrnutých do jedné nominace,
- počet zapojených rolí,
- možnost jasně stanovit hranice pilotu,
- množství podmínek a návazností, které je nutné před pilotem vyřešit,
- zda příprava k pilotnímu ověření není vzhledem k charakteru pilotu neúměrně dlouhá.

Zásadní rozlišení

Komplexnost pracovního procesu → ANO, hodnotíme zde.

Komplexnost práce AI s informacemi → NE, patří do kritéria 3.

Technickou komplexnost AI řešení → NEHODNOTÍME.

Velké množství dokumentů, více zdrojů nebo náročná práce s kontextem tedy samy o sobě hodnocení tohoto kritéria nesnižují.

★ Nízká vhodnost – příliš velké sousto

Proces je příliš rozsáhlý nebo komplexní, případně už z dostupných informací vyplývá, že jeho příprava k pilotnímu ověření bude neúměrně dlouhá či náročná. Pro pilot by bylo nutné proces významně zúžit, případně není vhodný pro tento konkrétní pilot.

Příklad: AI má podporovat celý proces případu – od přijetí a klasifikace přes vyhledání podkladů, práci s legislativou a tvorbu rozhodnutí až po následnou komunikaci.

Proč: Ve skutečnosti jde o několik různých pracovních potřeb, činností a výstupů spojených do jednoho návrhu.

Stejné hodnocení může být na místě, pokud je proces sice dobře popsán, ale už víme, že než by bylo možné pilot vůbec zahájit, bylo by nutné dlouhodobě řešit velké množství zásadních podmínek a návazností.

★★ Střední vhodnost – zvládnutelné sousto

Proces je pro pilot zvládnutelný, ale obsahuje více souvisejících činností, variant, výstupů nebo podmínek. Před zahájením pilotu bude potřeba přesněji vymezit jeho hranice nebo některé části omezit.

Příklad: AI má při řešení případu vyhledávat obdobná rozhodnutí, porovnávat je a současně provádět několik různých kontrol připravovaného dokumentu.

Proč: Činnosti spolu věcně souvisejí, ale pilot obsahuje více funkcí a bude potřeba přesně určit jeho hranice.

★★★ Vysoká vhodnost – přiměřené sousto

Proces je jasně ohraničený a svou komplexností přiměřený pilotnímu ověření. Z dostupných informací zároveň nevyplývají významné překážky, které by bránily jeho přípravě k pilotu v přiměřeném čase.

Příklad: Na základě nového případu má AI v předchozích rozhodnutích vyhledat obsahově obdobné případy, porovnat jejich relevantní okolnosti a připravit pracovníkovi strukturovaný přehled.

Proč: Práce AI může být obsahově velmi náročná, ale samotný předmět pilotu je dobře ohraničený. Víme, co vstupuje, jakou pracovní potřebu řešíme a jaký výstup očekáváme.

Rychlá orientace

★	Příliš velké sousto – výrazně zúžit / nevhodné pro tento pilot
★★	Zvládnutelné sousto – potřebujeme přesněji vymezit
★★★	Přiměřené sousto – dobře ohraničené pro pilot

7. Kritérium 5 – Měřitelnost přínosu

Kritérium 5 – Měřitelnost přínosu
Jak po pilotu poznáme, že to opravdu pomohlo?

Co hodnotíme

Hodnotíme, zda budeme po pilotu schopni poznat a doložit, zda očekávaný přínos skutečně nastal.

Nemusíme předem vědět, kolik přesně času ušetříme nebo jak velké zlepšení nastane. Důležité je, zda dokážeme určit, podle čeho změnu poznáme.

Může jít například o:

- čas potřebný k provedení konkrétní činnosti,
- kolik práce lze zvládnout za stejnou dobu,
- snížení množství manuální práce,
- snížení chybovosti nebo počtu oprav,
- správnost, úplnost nebo kvalitu výstupu,
- rychlost či úspěšnost nalezení potřebné informace,
- snížení počtu rutinních dotazů na pracovníky,
- spokojenost a zpětnou vazbu uživatelů.

Důležité pravidlo

Nehodnotíme, zda nominující ve formuláři navrhl konkrétní KPI.

Pokud konkrétní metrika nebyla součástí nominace, proces za to nesmí být penalizován.

Ptáme se:

Lze z charakteru popsaného přínosu rozumně odvodit způsob jeho ověření?

★ Nízká měřitelnost

Bude obtížné přesvědčivě doložit, zda očekávaný přínos skutečně nastal.

Příklad: Přínos je popsán pouze jako „usnadnění práce“ nebo „lepší orientace v informacích“ a ani z charakteru procesu není zřejmé, jakou konkrétní změnu bychom mohli sledovat.

Proč: Přínos může být reálný, ale bude obtížné jej po pilotu spolehlivě ověřit.

★★ Střední měřitelnost

Přínos dokážeme smysluplně vyhodnotit, ale potřebujeme kombinovat více ukazatelů nebo objektivní údaje se zpětnou vazbou uživatelů.

Příklad: AI pomáhá pracovníkům připravovat kvalitnější odborné podklady. Lze sledovat čas potřebný k jejich přípravě, posoudit úplnost či kvalitu a doplnit hodnocení zpětnou vazbou pracovníků.

Proč: Dokážeme zjistit, zda AI pomohla, ale výsledek nelze dostatečně vystihnout jedním jednoduchým ukazatelem.

★★★ Vysoká měřitelnost

Máme jasný způsob, jak očekávaný přínos ověřit a porovnat se současným stavem.

Příklad 1 – čas: Porovnáme dobu potřebnou k vyhledání a zpracování relevantních podkladů před pilotem a s podporou AI.

Příklad 2 – množství zvládnuté práce: Dnes pracovník za určitou dobu zpracuje například 5 dokumentů. S podporou AI za stejnou dobu zvládne 10.

Příklad 3 – klasifikace: Výsledek navržený AI lze porovnat se správným zařazením potvrzeným pracovníkem.

Příklad 4 – AI vyhledávač: Lze ověřovat, zda uživatel nalezl požadovanou informaci, hodnotit užitečnost odpovědi a případně sledovat změnu počtu souvisejících rutinních dotazů na úřad.

Rychlá orientace

★	Obtížně doložíme, zda to pomohlo
★★	Dokážeme to vyhodnotit, ale potřebujeme více pohledů
★★★	Máme jasný způsob, jak přínos ověřit

8. Co není součástí hvězdičkového hodnocení

Četnost využití

Uvádíme ji samostatně, pokud je z nominace známá. Například: denně / týdně / měsíčně / několikrát ročně / dle potřeby.

Četnost nesmí uměle zvyšovat nebo snižovat rozsah využití.

Charakter dat

Uvádíme samostatně: veřejná / interní / diskrétní / kombinace.

Charakter dat může být důležitý pro finální rozhodnutí o kandidátech, ale nepřidává ani neubírá hvězdičky.

Proces s veřejnými daty tedy automaticky není lepším kandidátem než proces s interními či diskrétními daty.

Pro základní posouzení pracovního procesu potřebujeme znát klasifikační stupeň informací, se kterými se v něm pracuje.

PŘEHLED KLASIFIKACE INFORMACÍ DLE zp ČTÚ č. 63/2024

Všechna aktiva z pohledu důvěrnosti klasifikována do tří stupňů.

VEŘEJNÉ – nízká úroveň důvěrnosti
Informace je veřejně přístupná nebo byla určena ke zveřejnění. Informace může být dále poskytována a šířena bez omezení. Narušení důvěrnosti informací neohrožuje oprávněné zájmy povinné osoby.

INTERNÍ – střední úroveň důvěrnosti
Informace nejsou veřejně přístupné a tvoří know-how ČTÚ. Ochrana informací není vyžadována žádným právním předpisem nebo smluvním ujednáním. Informace může být sdílena v rámci organizace příjemce a případně také s dalšími partnerskými subjekty příjemce, avšak nikoli skrze veřejně dostupné kanály.

DISKRÉTNÍ – vysoká až kritická úroveň důvěrnosti
Informace nejsou veřejně přístupné a jejich ochrana je vyžadována právními předpisy, jinými předpisy nebo smluvními ujednáními, například obchodní tajemství nebo osobní údaje, případně informace vyžadují nadstandardní míru ochrany. Informace může být poskytnuta pouze konkrétním osobám, které splňují princip need-to-know a jejichž informování je důležité pro vyřešení problému či hrozby uvedené v informaci.

9. Jak postupovat při hodnocení jednoho procesu

Krok 1 – přečíst celou kartu procesu

Neomezovat se pouze na název. Vycházet z popisu procesu, vstupů, výstupů, očekávaného využití, přínosů a dalších dostupných informací.

Krok 2 – hodnotit kritéria postupně a odděleně

U každého si položit příslušnou kontrolní otázku:

Jak moc to pomůže ČTÚ?

Kolika lidem to může pomoci?

Je to opravdu práce pro AI – a jak moc? Jakou roli tady AI hraje?

Neukousli jsme si pro jeden pilot příliš velké sousto?

Jak po pilotu poznáme, že to opravdu pomohlo?

Krok 3 – přidělit ★, ★★ nebo ★★★

Hodnotit podle popisu jednotlivých úrovní, nikoli podle celkového dojmu z nominace.

Krok 4 – nepřenášet jedno kritérium do druhého

„Používají to jen tři lidé.“ → ovlivňuje rozsah, nikoli automaticky přínos.

„AI musí procházet tisíce dokumentů.“ → samo o sobě neznamená nízkou vhodnost pro pilot.

„Jsou tam diskrétní data.“ → zaznamenáme jako charakter dat, automaticky nesnižujeme hvězdičky.

„Nevíme, jak technologicky řešení postavit.“ → není předmětem tohoto hodnocení.

Krok 5 – zaznamenat četnost a charakter dat

Bez hvězdiček.

10. Jak výsledek interpretovat

Hvězdičkové hodnocení má pomoci strukturovat úsudek pracovní skupiny, nikoli vytvořit matematický automat na výběr vítězů.

Vyšší celkové hodnocení znamená, že proces vychází silněji napříč stanovenými kritérii.

U nejsilnějších kandidátů je ale potřeba podívat se také na profil jednotlivých hodnocení.

Například dva procesy mohou získat stejný počet hvězdiček:

Proces A: ★★★ / ★ / ★★★ / ★★★ / ★★★

Proces B: ★★ / ★★★ / ★★ / ★★★ / ★★★

Celkové skóre může být podobné, ale jde o dva velmi odlišné kandidáty.

Finální rozhodnutí proto vzniká diskusí pracovní skupiny nad nejsilnějšími kandidáty, nikoli pouze automatickým výběrem dvou nejvyšších čísel.

Metodika v pěti větách

1. Přínos – Jak moc to pomůže ČTÚ?

2. Rozsah – Kolika lidem to může pomoci?

3. AI – Je to opravdu práce pro AI a jakou roli tam AI hraje?

4. Pilot – Neukousli jsme si příliš velké sousto?

5. Měření – Jak poznáme, že to opravdu pomohlo?

Každé kritérium hodnotíme samostatně podle toho, co skutečně měří – ne podle celkového dojmu z kandidáta.