**Příloha č. 1 - Specifikace předmětu plnění**

# Obsah

[Obsah 1](#_Toc38461518)

[1 Předmět plnění 2](#_Toc38461519)

[1.1 Cíl 2](#_Toc38461520)

[1.2 Předmět plnění 2](#_Toc38461521)

[2 Specifikace požadavků na SWAD 2](#_Toc38461522)

[2.1 Základní funkční požadavky 2](#_Toc38461523)

[2.1.1 Základní funkce anonymizace 2](#_Toc38461524)

[2.1.2 Další požadavky na aplikaci pro anonymizaci dokumentů 3](#_Toc38461525)

[2.1.3 Požadavky na technologii OCR (Optické rozpoznávání znaků) 3](#_Toc38461526)

[2.1.4 Další požadované funkce 3](#_Toc38461527)

[2.1.5 Integrace 4](#_Toc38461528)

[2.2 Legislativní požadavky 4](#_Toc38461529)

[2.3 Nefunkční a systémové požadavky 4](#_Toc38461530)

[2.4 Bezpečnostní požadavky 4](#_Toc38461531)

[3 Specifikace požadavků na dodávku systému 6](#_Toc38461532)

[4 Infrastruktura a součinnost Objednatele 6](#_Toc38461533)

[5 Seznam pojmů a zkratek 6](#_Toc38461534)

# Předmět plnění

## Cíl

Projekt má za cíl dodávku a implementaci centrálního nástroje pro ad-hoc zpracování (anonymizaci údajů) dokumentů za účelem ochrany osobních údajů a obchodního tajemství, včetně úpravy povinně zveřejňovaných smluv. Nástroj bude určen všem zaměstnancům Objednatele, kteří pracují s dokumenty, které je potřeba anonymizovat z výše uvedených důvodů. Základní očekávanou funkcionalitou je umožnění anonymizace údajů v dokumentech, která bude technicky plně v souladu s legislativními a metodickými požadavky Objednatele. Odpovědnost za věcně správné provedení anonymizace zůstává na Objednateli, nicméně se očekává, že dodaný nástroj tento proces výrazně metodicky i technicky zefektivní, a to například využitím předpřipravených šablon nebo předvýběrem citlivých pasáží v dokumentech.

## Předmět plnění

**Předmětem plnění je:**

* **Dodávka a implementace** Softwaru pro anonymizaci dokumentů (SWAD);
* Uvedení dodaného řešení do provozu a jeho **integrace** adresářovými službami Objednatele;
* Licenční zajištění - neomezená **multilicence**;
* Poskytnutí základního **seznámení s produktem**, metodických materiálů a systémové dokumentace správcům a klíčovým uživatelům Objednatele;
* Služba podpory uživatelům SWAD v rozsahu 15 MD.

**Předmětem Smlouvy není** Dodání související potřebné infrastruktury (HW a systémový SW). Pro provoz aplikace bude využita stávající infrastruktura Objednatele.

**Aplikace pro anonymizaci dokumentů (SWAD)** musí podporovat formou automatizace dva základní procesy:

* Anonymizace dat v dokumentu
* Uveřejnění dokumentu s anonymizovanýmy daty v Registru smluv (dále RS) dle zákona č. 340/2015 Sb., o zvláštních podmínkách účinnosti některých smluv, uveřejňování těchto smluv a o registru smluv (zákon o registru smluv), ve znění pozdějších předpisů.

# Specifikace požadavků na SWAD

V této kapitole jsou definovány požadavky Objednatele na funkcionality požadovaného řešení.

## Základní funkční požadavky

### Základní funkce anonymizace

1. Anonymizace dokumentů ve formátech PDF, DOC, DOCX, RTF, TXT, XLS, XLSX, ODS, ODT, GIF, JPG, JPEG, TIF, TIFF, PNG, BMP.
2. Definice typu anonymizovaných dat pro různé typy dokumentů.
3. Manuální oprava.
4. Manuální označení libovolné plochy.
5. Sada vzorů pro redigování.
6. Implementace plně automatizovaného redigování (anonymizace).
7. Různé možnosti anonymizace citlivých informací pomocí výsledného zobrazení (černé, šum, …), bez možnosti zpětného vytěžení.
8. Anonymizace a změny metadat uvedených ve vlastnostech dokumentu.
9. Optimalizace velikosti obrázků v dokumentech.
10. Optimalizace výsledných anonymizovaných dokumentů.
11. Označení celých slov jedním kliknutím při manuálním redigování.
12. Označování celých bloků textu při manuálním redigování.
13. Označování plochy stránky dokumentu při manuálním redigování.
14. Mazání nesprávně označených informací (možnost mazat jednotlivá označení, všechna označení na straně, nebo v celém dokumentu).
15. Možnost na zvolených stranách dokumentu (od strany po stranu) vyhledávat citlivé informace.
16. Možnost dávkového zpracování více dokumentů.

### Další požadavky na aplikaci pro anonymizaci dokumentů

Aplikace

1. musí obsahovat sadu přednastavených vyhledávacích objektů pro redigování (rozpoznání podpisů, čísel účtů, IBAN a dalších osobních údajů) včetně definování a ukládání vlastních typů objektů,
2. musí umožňovat redigování na základě barevně zvýrazněného textu (podpora min.: funkce Marker v MS OFFICE a funkce Anotator/Zvýrazněný Text v Adobe aplikacích),
3. poskytne aktualizace přednastavených objektů pro vyhledávání za účelem průběžné zpřesňování objektů pro vyhledání citlivých informací,
4. umožní zápis metadat o smluvní straně s možnosti automatického doplnění údajů o smluvní straně podle IČO nebo datové schránky,
5. umožní fulltextové vyhledávání v dokumentu včetně vyhledávání označených objektů určených k zakrytí,
6. poskytne používání šablon pro zakrývání citlivých informací,
7. pomocí přednastavených objektů a typů dokumentu SWAD umožní automatické redigování dokumentů,
8. umožní manuální opravu automaticky označovaných redigovaných informací (ploch, textu, elektronických podpisů, …) v dokumentu,
9. umožní řízení automatického vyhledávání citlivých informací dokumentu (např. možnost zastavení vyhledávání, přeskočení vyhledaných informací,…),
10. zobrazuje výsledný redigovaný dokument v pdf reader nebo v prohlížeči (MS Internet Explorer, Google Chrome, Mozilla Firefox).

### Požadavky na technologii OCR (Optické rozpoznávání znaků)

Technologie OCR

1. musí být součásti SWAD,
2. má plnou podporu českého jazyka,
3. zahrnuje automatickou korekci šikmo naskenovaných dokumentů.

### Další požadované funkce

1. výstup PDF s textovou vrstvou (OCR), s tím, že anonymizované informace se nedají vytěžit,
2. fultextové vyhledávání,
3. uživatelské rozhraní v českém jazyce,
4. evidence informací o zaměstnanci, který vykonal anonymizaci,
5. zobrazení statistik redigovaného dokumentu (počet redigování, % redigovaných informaci, rychlost vytvoření dokumentu….),
6. logování a statistika procesu konverze a procesu OCR dokumentů.

### Integrace

1. autentizace pomocí přihlašovacích údajů k DS,
2. integrace s RS – smlouvu je po redigování možno odeslat do RS,
3. automatické odesílaní anonymizovaných dokumentů dle nastavení do RS pomocí DS.

## Legislativní požadavky

Navržené řešení a výsledný IS musí respektovat povinnost Objednatele splňovat legislativní požadavky, vyplývající z českého právního řádu v platném znění, zejména

1. zákona č. 340/2015 Sb., o zvláštních podmínkách účinnosti některých smluv, uveřejňování těchto smluv a o registru smluv (zákon o registru smluv);
2. zákona č. 106/1999 Sb., o svobodném přístupu k informacím;
3. zákona č. 110/2019 Sb., o zpracování osobních údajů;
4. zákona č. 134/2016 Sb. o zadávání veřejných zakázek, ve znění pozdějších předpisů;
5. zákona č. 300/2008 Sb., o elektronických úkonech a autorizované konverzi dokumentů;
6. zákona č. 111/2009 Sb., o základních registrech;
7. zákona č. 127/2005 Sb., o elektronických komunikacích a o změně některých souvisejících zákonů (zákon o elektronických komunikacích).

## Nefunkční a systémové požadavky

Kapitola definující zákaznické požadavky, které kladou omezení na design a provedení řešení (například požadavky na zálohování, formáty a rozsah evidovaných dat nebo designové omezení):

1. řešení bude provozováno u Objednatele (on-premise),
2. nezávislost na provozní platformě (MS, LINUX, …), univerzální řešení nezávislé na aplikacích třetích stran,
3. podpora internetových prohlížečů Firefox 3.6+, IE 9+, Chrome 11+, Opera 12+,
4. plná funkčnost ve webovém browseru (min. uvedených v bodě 3) bez nutnosti instalace dodatečných modulů,
5. uživatelské rozhraní a dokumentace musí být v českém jazyce,
6. plná podpora otevřených standardů pro integraci a další vývoj, třívrstvá architektura; nativní podpora SOA architektury, podpora protokolů SOAP a WSDL nebo REST,
7. komunikace s externími systémy prostřednictvím webových služeb a emailových zpráv,
8. řešení musí poskytovat webové služby pro možnost volání anonymizačních funkcionalit externími systémy,
9. podpora interního nebo externího OCR serveru,
10. škálovatelnost řešení pomocí fronty vůči více OCR serverům,
11. šifrované ukládaní všech důležitých informací,
12. šifrování komunikace pomocí protokolu TLS 1.2 nebo 1.3,
13. multithreadingové zpracování úkolů serverem.
14. předpokládá se že SWAD bude užívat cca pět tisíc uživatelů.
15. řešení bude podporovat autentizaci a autorizaci uživatelů vůči systému MS Active Direktory (případně obecného řešení LDAP)
16. bezpečnost zálohování a obnovy – Dodavatel předloží Objednateli návrh plánu obnovy pro aplikace a systémy přímo ovlivňující předmět Plnění.

## Bezpečnostní požadavky

Bezpečnostní požadavky vyplývají z povinností kladených zákonem č. č. 181/2014 Sb., o kybernetické bezpečnosti a o změně souvisejících zákonů (zákon o kybernetické bezpečnosti), v platném znění, resp. jejich naplnění službami nového systému, jsou chápány jako základní vlastnosti systému a musí být systémem prosazeny již od samého začátku. Zadavatel neumožní implementaci bezpečnostních mechanismů po implementace aplikace. Funkcionalita bezpečnostních mechanismů bude ověřována již v rámci testovacího provozu a akceptace, a to formou provedení bezpečnostních a penetračních testů, které musejí splňovat bezpečnostní požadavky dle metodiky OWASP. Jejich splnění bude testováno dle OWASP (Testing Guide v 4.0 (https://www.owasp.org/images/1/19/OTGv4.pdf). Zjištění zranitelností ze seznamu OWASP Top 10–2017 (https://www.owasp.org/images/7/72/OWASP\_Top\_10-2017\_%28en%29.pdf) bude s dodavatelem řešeno jako závažné porušení smlouvy.

**Základní bezpečnostní požadavky:**

1. nástroj nesmí využívat služeb třetích stran, být realizován na prostředcích třetích stran, které jsou mimo kontrolu Objednatele (např. služby OCR nebo textové analýzy, provozované v cloud prostředí třetích stran);
2. služba anonymizace smí být dostupná pouze autentizovaným a autorizovaným uživatelům;
3. systém musí být zabezpečen proti přesměrování na podvržené stránky;
4. veškerá komunikace musí probíhat v rámci zabezpečeného komunikačního kanálu (protokolem TSL 1.2 nebo 1.3) se serverem využívajícím důvěryhodný certifikát zajišťující jeho autenticitu.
5. požadavky na použitou kryptografii musí být v souladu s doporučením NCKB v oblasti kryptografických prostředků (<https://www.govcert.cz/cs/doporuceni-v-oblasti-kryptografickych-prostredku/>)
6. bezpečnost zálohování a obnovy – Dodavatel předloží Objednateli plán obnovy pro aplikace a systémy dodaných Dodavatelem a přímo ovlivňující předmět Plnění. Zálohy budou šifrovány a bude průběžně testována obnova z nich, aby byla zajištěna úspěšnost obnovy.
7. Nástroj musí vytvářet a zaznamenávat události (logy) a to minimálně v níže uvedeném rozsahu:

Logování událostí loguje minimálně:

a) zastavení nebo spuštění zpracování

b) změna stavu dávky v rámci celého workflow digitalizace

c) vyřazení dávky ze zpracování v rámci celého workflow digitalizace

d) přesun obrazů a metadat z lokálního do centrálního úložiště

e) uložení vytěžených, validovaných dat do centrálního úložiště

Monitorovací systém sleduje a loguje minimálně:

a) datum a čas včetně specifikace časového pásma,

b) typ činnosti,

c) identifikaci technického aktiva, které činnost zaznamenalo,

d) jednoznačnou identifikaci účtu, pod kterým byla činnost provedena,

e) jednoznačnou síťovou identifikaci zařízení původce a

f) úspěšnost nebo neúspěšnost činnosti,

g) přihlašování a odhlašování ke všem účtům, a to včetně neúspěšných pokusů,

h) úspěšné i neúspěšné manipulace s účty, oprávněními a právy,

i) neprovedení činností v důsledku nedostatku přístupových práv a oprávnění,

j) události vzešlé z činnosti uživatelů, které mohou mít vliv na bezpečnost digitalizace nebo operacemi a manipulacemi s daty,

k) zahájení a ukončení činností technických aktiv,

l) kritická i chybová hlášení technických aktiv,

Předávání událostí (logů) – Události budou zaznamenávány a do SIEM ČP předávány ve strojově čitelném formátu s podporou předání službě typu syslog či JSON.

# Specifikace požadavků na dodávku systému

Projekt bude dodán tak, aby seznámení uživatelů s obsluhou SWAD, požadovaná integrace, uvedení do produktivního provozu bylo řádně provedeno nejpozději do 30 (třiceti) kalendářních dnů od nabytí účinnosti Smlouvy.

# Infrastruktura a součinnost Objednatele

Objednatel je schopen zajistit následující technické komponenty a služby v rámci součinnosti a vlastního prostředí:

* Síťová instalace

V případě, že by Dodavatel využíval jiné databáze, musí do Ceny promítnout veškeré náklady s licencováním a technickou podporou databáze, implementací v prostředí Objednatele, seznámení administrátorů s její správou a poskytnutí plné součinnosti s nastavením zálohování;

* Serverová a virtualizační infrastruktura,
* Úložná kapacita pro zajištění provozního úložiště, zálohování a archivace,
* Prostředí MS SQL Server,
* Koncové stanice MSW v.10 + MS Office.

# Seznam pojmů a zkratek

**Popis použitých zkratek**

|  |  |
| --- | --- |
| **Zkratka** | **Význam** |
| SWAD | Software pro anonymizaci dokumentů |
| RS | Registr smluv |
| DS | Datová schránka |
| OCR | Optické rozpoznávání znaků |
| HW | Hardware |
| SW | Software |