



Řešení pro pokročilou analýzu dat

dāterā

→ Úvod

LISSA je inovativní software navržený pro profesionální hloubkovou analýzu dat, vyvinutý společností DATERA.

V současné digitální éře se každodenně zvyšuje množství dostupných dat. Vyšetřovatelé, analytici a datoví specialisté se často potýkají s potřebou včas a efektivně zpracovat velké objemy informací. Naskenované dokumenty, výpisy z databází, fotografie dokumentů, přepisy hlasových záznamů, příspěvky na sociálních sítích, profily atd. je třeba shromáždit, připravit a analyzovat. Při jakémkoli vyšetřování je doba zpracování a kvalita důkazů zásadní.

LISSA je řešení, které pomáhá jednoduše a rychle získat cenné informace z datového chaosu. Poskytuje uživatelům nástroje potřebné ke zpracování zdrojových dokumentů a datových souborů na vstupu, vyhledání informací pomocí klíčových slov nebo metadat, objevování skrytých vzorců, odhalování souvislostí a usnadnění rychlých a přesných rozhodnutí založených na datech.

→ Informace z produkčního nasazení systému

Celkový počet dokumentů:
více než **1.200.000.000**.

Informačních zdrojů:
více než **250**.

Rychlost vyhledávání:
jednotky vteřin.

Integrace na **i2** repozitář.

Rekurzivní rozbalování kompozitních
a archivních formátů. Kompozitní formáty
se indexují s vazbou potomků a rodičů.

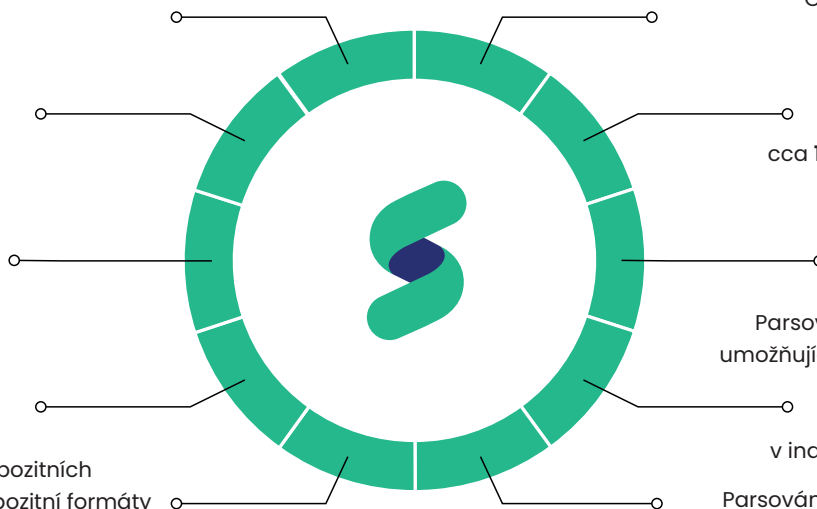
Objem vstupních dokumentů:
50 TB.

Rychlost zpracování
příchozích dokumentů:
cca **10.000.000 dokumentů/hod.**

Podporovaných formátů:
přes **300**.

Parsování databázových „dumpů“
umožňující rozpad na úroveň definice
tabulek a řádků, co řádek,
to samostatný dokument
v indexu a uživatelském rozhraní.

Parsování a zpracování CSV souborů,
včetně velkých souborů s velikostí přes **150 GB**.



→ Klíčové vlastnosti a funkce

LISSA je produkt, který byl postaven na mnohaletých zkušenostech analytiků. Má za úkol poskytnout kvalitní a efektivní správu a analýzu dat. Následující kapitoly popisují hlavní charakteristiky platformy, které ji odlišují od konkurence a přinášejí uživatelům jedinečné výhody.

→ Zpracování dat

Platforma LISSA automaticky detekuje nově vložená data. Archivy, přílohy nebo embedované soubory (např. obrázky ve Wordu) systém na pozadí rozbílí, následně ze všech dat extrahuje veškerý obsah a metadata. Na takto připravených datech provede případné transformace a úpravy. Data poté zaindexuje, aby v nich mohl uživatel snadno vyhledávat.

Seznam hlavních výhod systému LISSA, které jsou důležitým předpokladem pro efektivní práci:

- **Spolehlivá:** LISSA kvalitně zpracovává nepřeberné množství formátů (více než 300). Při vložení nových dat se proces zpracování aktivuje automaticky a uživatel má pak možnost sledovat jeho průběh. Rozsáhlé soubory systém rozdělí na menší části a v průběhu extrakce se dostává až k nehlouběji vnořeným datům, přičemž zachovává informaci o hierarchii. Poskytuje přehled o souborech, které kvůli poškození nemohly být zpracovány.
- **Výkonná:** LISSA zaručuje vysokou rychlost indexace nově přidaného obsahu a připravenost platformy na změnu obsahu a struktury nebo na její rozšiřování bez odstávky systému. Nároky na správu platformy jsou minimální.
- **Otevřená:** LISSA je postavena především na technologiích open-source, je soustavně optimalizována a doplňována funkční výbavou dle požadavků uživatelů. Zároveň dovoluje prostřednictvím otevřených rozhraní integrovat do procesu zpracování komerční produkty určené pro specifické úkony.
- **Modulární:** LISSA může být rozšířena o další moduly. To organizaci umožňuje nasazení platformy v různých konfiguracích dle specifických požadavků uživatelů.
- **Bezpečná:** LISSA poskytuje zabezpečení přístupu na základě definovaných oprávnění, odděleně k primární bázi dat a k výsledkům zpracování. Umožňuje monitorování a auditování uživatelských přístupů i zpracování dat. Dovoluje nasazení výhradně v režimu on-premise.
- **Uživatelsky přívětivá:** LISSA nabízí uživatelské prostředí se širokou škálou vizualizačních nástrojů. Ty usnadňují zadávání dotazů, dodatečných kritérií, zobrazení výsledků s náhledy a jejich export. K dosažení optimálních výsledků analýzy systém poskytuje rozsáhlé možnosti filtrování dle atributů generovaných v průběhu zpracování. Ovládání je intuitivní a přizpůsobené individuálním potřebám.

LISSA

Absolute Intelligence

Legenda

-  Hlavní komponenty LISSA
-  Moduly rozšíření
-  Datové zdroje

Uživatelské rozhraní

Analytická aplikace

Administrátorská aplikace

Monitorovací aplikace

Vlastní UI aplikace



Distribuovaný běh aplikace

Rízení zpráv do fronty

Konfigurace serveru

Bezpečnost

Plánovač

Telemetrie

Vyhledávání

Vlastní moduly

REST API



Příjem dat (zdroje)

Zdroj souborového systému

Zdroj webové služby

REST API

Vlastní zdroj



Zpracování dat (procesory)

Hlavní detektor

Hlavní vytěžovač

Hlavní transformátor

Hlavní překladač

Hlavní emitör

Vlastní procesor

Systém ukládání dat

Úložiště souborů

Úložiště AWS S3

Vlastní úložiště



Vyhledávač

OpenSearch

Vektorová DB

Další vyhledávač



→ Integrace s nástrojem i2

Kombinací nástrojů LISSA a i2 uživatel získá neobvyčejný pohled na svá data z různých perspektiv. Pomocí systému LISSA je schopen procházet zaindexovaná data a nacházet stěžejní informace, zároveň v nástroji i2 data zkoumá přes společné objekty v diagramu a hledá skryté vztahy.

LISSA na pozadí extrahuje detekované objekty, jež vkládá do výměnného formátu pro nástroj ETL (Extract-Transform-Load), který je zpracuje a automatizovaně vloží do i2 repozitáře. Data jsou připravena k prozkoumávání přes rozhraní analytického nástroje i2.



→ Výhody produktu

Řešení LISSA přináší řadu výhod, které uživatel ocení při každodenním používání. Produkt je navržen tak, aby poskytoval přesné a užitečné informace pro další rozhodování.

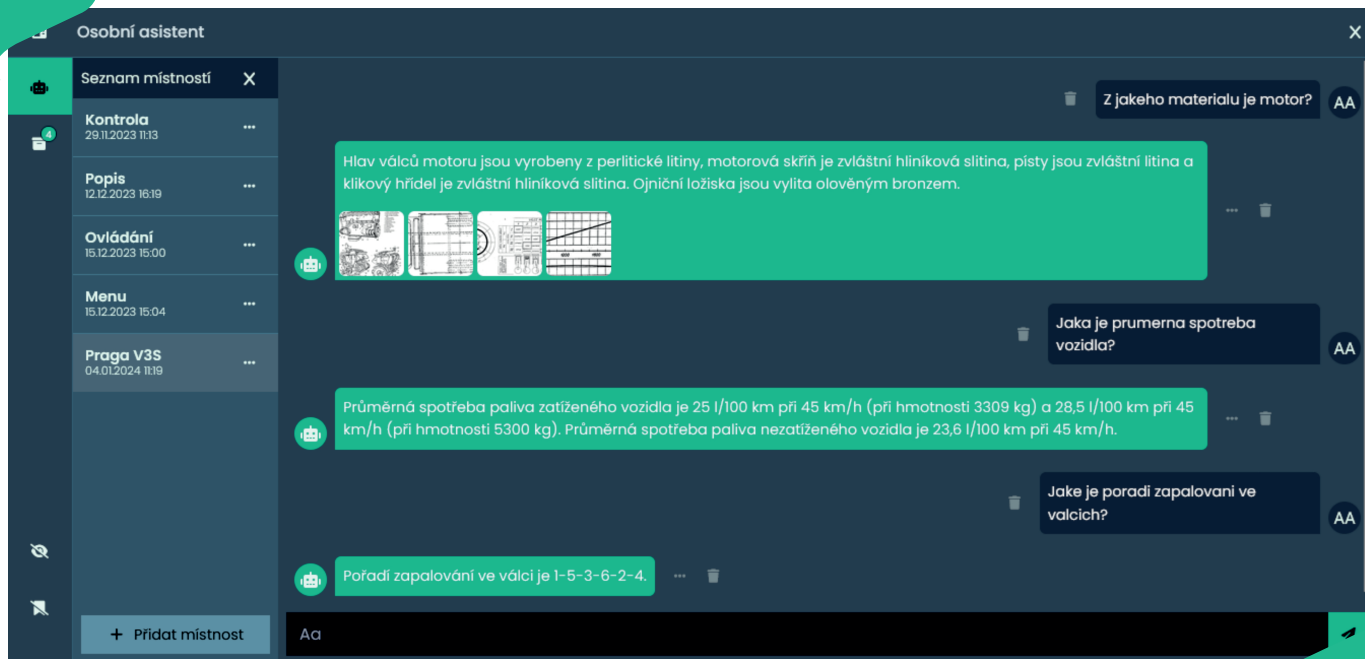
Klíčové výhody produktu LISSA:

- **Podpora pro strategická rozhodování:** Výsledky analýzy poskytují důležité informace pro rozhodovací a plánovací procesy.
- **Zrychlení analýzy:** Podstatné informace lze vyhledat rychle a s důvěrou v úplnost vstupních dat, což optimalizuje celý analytický proces a přináší nezkrácené výsledky.
- **Snížení rizika lidských chyb:** Díky automatizovanému zpracování dat a pokročilým funkcím omezování výsledků vyhledávání je minimalizována pravděpodobnost výskytu chyb způsobených lidským faktorem.
- **Využití moderních technologií:** Aplikace modulů založených na technologických trendech umožňuje uživatelům přistupovat k analýze dat z více směrů, a tím poskytovat hlubší vhled do řešené problematiky.
- **Intuitivnost:** Jednoduché a uživatelsky přívětivé rozhraní umožňuje uživatelům okamžitě využívat jeho funkce.

→ Inovativní přístup zapojením AI

Součástí systému LISSA je modul, který pro hledání informací využívá velké jazykové modely (LLM). Tyto modely rozpoznávají a analyzují text, přináší relevantní výsledky ve formě srozumitelné odpovědi na dotazy. Modul zajišťuje, že hledané informace budou snadno dostupné a přehledně uspořádané. To uživateli umožní rychle a efektivně získat informace potřebné pro rozhodování a strategie, může se spolehnout na špičkovou technologii, která usnadní a zrychlí práci.

Jazykové modely využívá LISSA pro dva druhy činností – pro vyhledávání a pro generování zpráv na základě interních vstupů (dokumentů).



The screenshot displays the LISSA AI assistant interface. On the left, a sidebar titled "Osobní asistent" (Personal Assistant) contains a "Seznam místností" (List of rooms) section. This section lists several items: "Kontrola" (29.11.2023 11:13), "Popis" (12.12.2023 16:19), "Ovládání" (15.12.2023 15:00), "Menu" (15.12.2023 15:04), and "Praga V3S" (04.01.2024 11:19). At the bottom of the sidebar is a button labeled "+ Přidat místnost" (Add room). The main area on the right is a chat window. It shows a series of interactions: a user query "Z jakého materiálu je motor?" (What material is the engine made of?) followed by an AI response detailing the materials of engine components (cylinders, valves, pistons, etc.) accompanied by four technical diagrams. Another user query "Jaka je průměrná spotřeba vozidla?" (What is the average fuel consumption of the vehicle?) is followed by an AI response providing fuel consumption data for different vehicle weights. A third user query "Jake je pořadí zapalování ve válcích?" (What is the firing order in the cylinders?) is followed by an AI response stating the firing order as 1-5-3-6-2-4. The interface includes standard chat elements like a trash icon for deleting messages and a text input field at the bottom with a placeholder "Aa".

Osobní asistent

Seznam místností

- Kontrola 29.11.2023 11:13
- Popis 12.12.2023 16:19
- Ovládání 15.12.2023 15:00
- Menu 15.12.2023 15:04
- Praga V3S 04.01.2024 11:19

+ Přidat místnost

Z jakého materiálu je motor?

Hlav válců motoru jsou vyrobeny z perlitické litiny, motorová skříň je zvláštní hliníková slitina, písty jsou zvláštní litina a klikový hřídel je zvláštní hliníková slitina. Ojnicní ložiska jsou vylita olověným bronzem.

Jaka je průměrná spotřeba vozidla?

Průměrná spotřeba paliva zatíženého vozidla je 25 l/100 km při 45 km/h (při hmotnosti 3309 kg) a 28,5 l/100 km při 45 km/h (při hmotnosti 5300 kg). Průměrná spotřeba paliva nezatíženého vozidla je 23,6 l/100 km při 45 km/h.

Jake je pořadí zapalování ve válcích?

Pořadí zapalování ve válci je 1-5-3-6-2-4.

→ Vyhledávání

LISSA je vybavena algoritmem založeným na nejnovějších vývoji v oblasti jazykových modelů. Díky tomu je schopna porozumět textovým datům v nejrůznějších jazycích, analyzovat je, a srozumitelně interpretovat výsledky. **Pokročilá analýza textu** pomocí jazykových modelů dovolu je například hledat souvislosti a vztahy mezi entitami, které běžným hledáním pomocí klíčových slov není možné rozpoznat.

→ Rozpoznávání entit

LLM má schopnost rozpoznat požadované entity v dokumentech, jako jsou jména, místa, termíny a další důležité informace, což usnadňuje vyhledávání a kategorizaci dat.

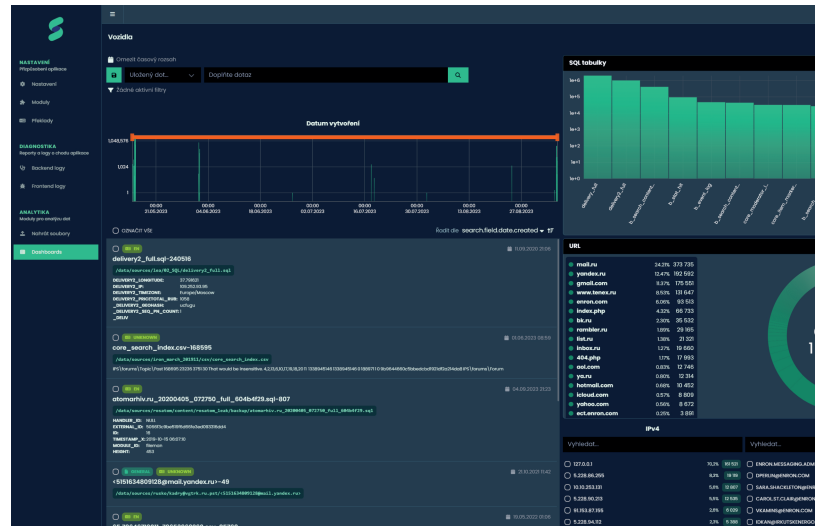
→ Generování obsahu

Na základě interních dokumentů a souborů je uživatel schopen získat automaticky návrhy výstupů, které obsahují klíčové informace ze vstupních dat. Například lze takto generovat souhrnný dokument vycházející z několika vstupů.

➤ **Automatizace procesu:** Modul LLM výrazně snižuje potřebu manuální analýzy a zpracování dokumentů. Výsledkem je uvolnění kapacit a ušetření času a zdrojů pro jiné důležité úkoly. Tvorba výsledného dokumentu je řízena pomocí definovaného workflow. Výstup je možné v průběhu tvorby upravovat a ladit jeho obsah.

→ Závěr

LISSA představuje revoluční krok v oblasti analýzy dat. Její pokročilé funkce a intuitivní rozhraní umožňují uživatelům objevovat nové informace, řešit složité problémy a podporovat strategická rozhodnutí. S důrazem na rychlost, efektivitu a spolehlivost poskytuje LISSA v dnešním digitálním světě velkou konkurenční výhodu.





LISSA

Absolute
Intelligence



Pavel Prokop



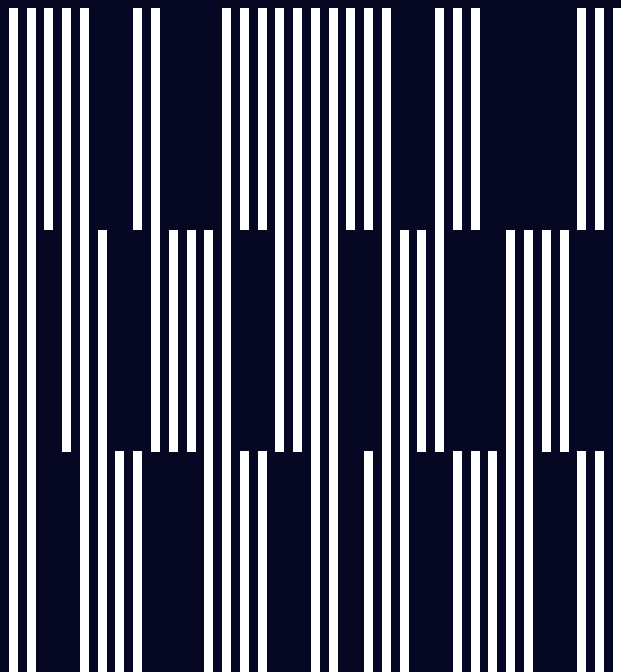
pavel.prokop@datera.cz



linkedin.com/in/prokoppavel



www.datera.cz



DATERA s.r.o. působí na trhu od roku 2012. Náš tým disponuje v součtu desítkami let zkušeností v ICT. Společnost DATERA v současné době zaměstnává přes 30 pracovníků, roční obrát společnosti dosahuje každoročního růstu a pohybuje se kolem 250 milionu Kč.

Pro naše zákazníky jsme spolehlivým dodavatelem a partnerem v oblastech infrastruktury, datových uložišť a analýzy strukturovaných a nestrukturovaných dat.

S vášní pro technologie a inovace se snažíme poskytovat nejlepší řešení na trhu.

datera