



Anonymisering av data med Presidio och LLM



Användning av AI idag

Alla medarbetare har tillgång till AI-plattform för eget bruk och uppmuntras att utforska användandet och skapa verksamhetsnyttiga assistenter.

Förberedda assistenter som underlättar kunskapsspridning i organisationen har framtagits och publicerats.

- Aida:
 - Intranätets sidor & dokument,
 - Kontaktinformation
 - Lagtexter (Webscraping)
- Kommun-Kai
 - Samtliga sidor på externweb
 - Kommunbolagens websidor
 - E-tjänster

Intric



- En plattform som tillgängliggör AI-assistenter för medarbetare
- Erbjuder ett utbud av olika modeller för vektorisering och svar
 - Vi använder främst OpenAI GPT-4o för svar, och tillfället.
- Ett gränssnitt för konfigurering och användning av assistenter.
- RAG databas som möjliggör användning av verksamhetsdata för assistenterna.
- API som möjliggör smidig interaktion med plattformen via script / automatiseringar



Varför utöka anonymisering av data?

Utvöka användningen av värdefull kunskap på nya sätt i organisationen.

- Sammanfatta
- Jämföra
- Sprida

Samtidigt uppfylla kraven från lagar och förordningar:

- Dataskyddsförordningen (GDPR)
- Dataskyddslagen
- Offentlighets- och sekretesslagen



Python och Fast API

Fast API

- Möjliggör snabb och asynkron behandling av data
- Tydlig och enkel utveckling av API



Python

- Lätthanterad och smidig utveckling
- Bra utbud av Open Source bibliotek





Presidio



- Open-source – Startat av MS
- Identifiering och anonymisering av en känslig information (Personally Identifiable Information- PII)
- Använder sig av flera tidigare öppna Natural Language Processing (NLP) projekt



Microsoft



Vår lösning

1. Konfigurering av Presidio-Analyzer & Anonymizer
 2. Analys och Anonymisering (Presidio)
 3. Sammanfattning / Bearbetning (Konfigurerad AI-Assistent)
 4. Uppladdning till Intric
- Allt genom Python-Script

Konfigurering av Presidio-Analyzer & Anonymizer

Skapar en instans av anonymizeraren

Väljer en **NLP-motor** som ska styra hanteringen. (spaCey som körs helt lokalt)

- Applicerar ett antal **recognizers** som styr identifiering.

- Namn

- Platser

- Personnummer

- Telefonnummer

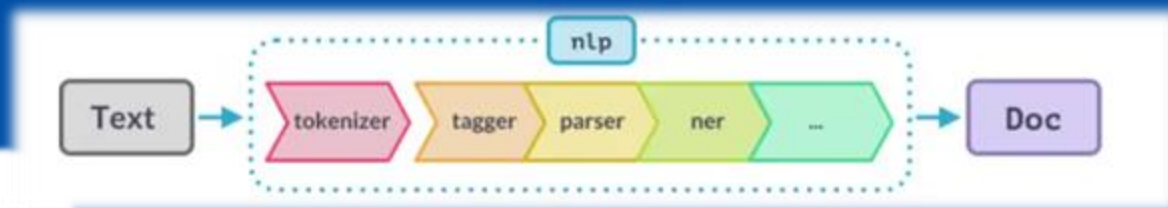
- E-postadresser

- AD-användarnamn

- Postorter och nummer

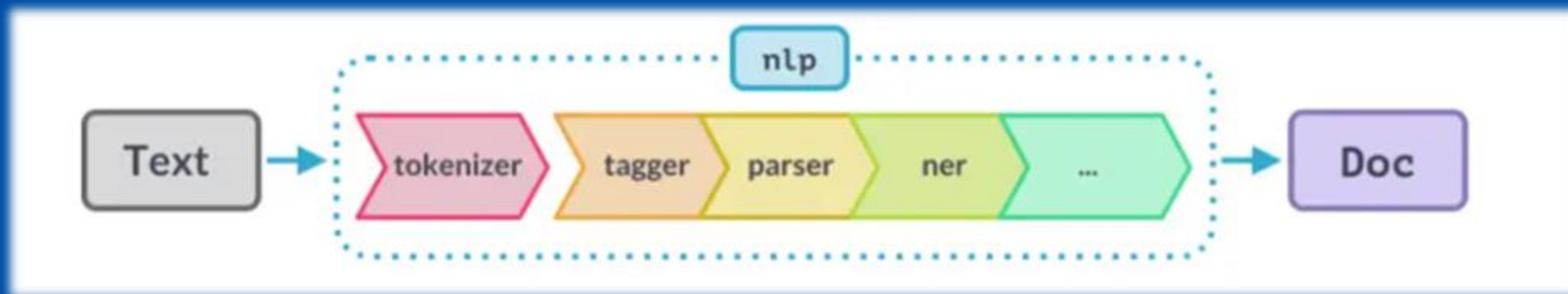


Egenskriva regler
och standarder



Analys & Anonymisering

1. Texten bearbetas och delas in i tokens,
2. Tokens markeras med ett antal tags utifrån NLP-motorns modell och konfiguration.
3. Tags processeras sedan och undersöks för att med så stor säkerhet som möjligt peka ut om de ska anonymiseras.
4. Entiteter som identifierats för anonymisering byts ut mot specificerade operator som <PERSON>, <TELEFONNUMMER>, <EMAIL> eller <PLATS>



Sammanfattning / Bearbetning med LLM



En assistent förbereds:

- Lämpliga kunskaper för att behandla specifika ämnen i dess Knowledges / RAG
- Passande uppdragsbeskrivning för dess uppgift i bearbetningen

Den anonymiserade datan ges till assistenten:

- Via API anrop mot Intric

Svar granskas av användaren:

- Vid behov genereras nya resultat

Uppladdning till Intric



- Det färdiga textmaterialet kan nu laddas upp till valfri assistents RAG för användning i framtida generering av material eller för att svara på verksamhetsfrågor gällande innehållet.
- Kunskap som använts vid skapandet av sammanfattningen/tillsnyggningen stärker kvaliteten på svar utifrån materialet.

Exempel på anonymiseringen



Text är ett anonymiserat mail
som fyllts med fingerad data
gällande namn, nummer
adresser etc.



Ursprungsdata

Från: Maria Andersson maria.andersson@kommun.se

Till: Johan Svensson johan.svensson@kommun.se

Datum: 2024-02-15 11:53

Ämne: VB: Ungdomscoacher saknas i urvalet "Fritid"

Hej!

Jag behöver hjälp med att styra om Anna Nilsson enligt nedan. Jag fick inte eller förstod inte direktiven att skriva KST Fritidsgård Centrum gällande ungdomscoacher. Det är därför det inte fungerat med schemaläggningen osv. 😊

Kan den på lön som hanterar Kultur- och fritid ändra om konteringen på Anna från Kommunens kultur- och fritidskontor 42600?

Med vänlig hälsning,

Maria Andersson





Anonymiserat

Från: <NAMN><INTERN EMAIL>

Till: <NAMN> <INTERN EMAIL>

Datum: <DATUM> 11:53

Ämne: VB: Ungdomscoacher saknas i urvalet "Fritid"

Hej!

Jag behöver hjälp med att styra om <NAMN> enligt nedan. Jag fick inte eller förstod inte direktiven att skriva KST Fritidsgård Centrum gällande ungdomscoacher. Det är därför det inte fungerat med schemaläggningen osv. 😊

Kan den på lön som hanterar Kultur- och fritid ändra om konteringen på <NAMN> från Kommunens kultur- och fritidskontor 42600?

Med vänlig hälsning,

<NAMN>





Tillsnyggat

- Från: avsändarens interna e-post
Till: intern e-post
Datum: datum 11:53
Ämne: VB: Ungdomscoacher saknas i urvalet "Fritid"
- Hej!
- Jag behöver hjälp med att styra om det enligt nedan. Jag fick inte eller förstod inte direktiven att skriva KST Fritidsgård Centrum gällande ungdomscoacher. Det är därför det inte fungerat med schemaläggningen osv. 😊
- Kan den på lön som hanterar Kultur- och fritid ändra om konteringen från Kommunens kultur- och fritidskontor 42600?
- Med vänlig hälsning,
Avsändaren





Sammanställt (hela konversationen)

- Denna konversation rör en diskussion mellan olika personer inom en kommunal organisation om hur ungdomscoacherna skall schemaläggas och hanteras administrativt. Det framgår att det finns vissa missförstånd kring administrativa direktiv för att skriva in personer i det kommunala systemet och att vissa schemaändringar behövs för att få systemet att fungera korrekt.
- Problemen verkar delvis bero på att ungdomscoacher inte har varit korrekt inlagda i urvalet "Fritid", vilket lett till svårigheter med schemaläggning och löneadministration. Det diskuteras olika lösningar som involverar att justera konteringen (dvs. vilken kostnadsenhet lönekostnaderna ska belasta). Dessutom betonas vikten av rätt utformade anställningsavtal för att undvika liknande problem i framtiden.
- Personerna i dialogen kommer överens om att vissa administrativa förändringar behövs, inklusive kontakt med lönekontoret för att justera lönekonteringar och att vissa justeringar i anställningsavtal kan behövas. De diskuterar även att den närmaste chefen för vissa ungdomscoacher ska vara en specifik person inom organisationen. Det nämns att en del av detta kan behöva göras efter lönekörningar beroende på vad som behöver bokas om.
- Genomgående upprepas vikten av att hålla rätt chef och rätt konteringsinformation i systemen, för att säkerställa korrekt schemaläggning och lön för coacherna. Röster från olika chefer och administratör finns representerade i konversationen, som syftar till att samordna och lösa de uppkomna problemen.



Baserat på en större mängd...



Chat History Insights

Hur bör jag gå tillväga för att få hjälp med att få ut löner till min personal som inte fått någon?

Om du har personal som inte har fått sina löner, är det viktigt att kontakta löneavdelningen direkt för att få hjälp med detta. Beroende på orsaken till problemet kan lösningen variera. Här är några steg du kan ta baserat på tidigare konversationer:

1. **Kommunicera tydligt problemet:** Skicka en e-post till löneavdelningen och beskriv situationen noggrant. Ange vilka medarbetare det gäller och vilka perioder de inte fått lön för. Detta liknar tillvägagångssättet i flera konversationer där man kommunicerar problemet direkt **1**.
2. **Kontrollera med bemanningsenheten:** I vissa fall kan problemet bero på att timmar eller tjänstgöringsrapporter inte har rapporterats eller exporterats korrekt till löneavdelningen. Säkerställ att alla nödvändiga rapporter har skickats **2**.
3. **Direktkontakt:** Om det är brådsakande, kan det vara bra att ringa löneavdelningen för att få snabb återkoppling. Många gånger rekommenderas det att ringa under specifika telefontider för att få direkt hjälp **3**.
4. **Efterlysning av försvunna dokument:** I vissa fall kan dokument som skickats internt ha försvunnit (t.ex. anställningskontrakt), vilket skulle kunna vara orsaken till problemet. Dubbelkolla att alla dokumentationen har kommit fram **4**.
5. **Samverka med ekonomiavdelningen:** I vissa fall kan det behövas att samverka med ekonomiavdelningen om problemet relaterar till banköverföringar och liknande **5**.
6. **Korrigerera schemaläggning och rapportering:** Korrigera eventuella schemafel i tidrapporter om problemet ligger där **6**.





Nästa steg för oss

- Arbete tillsammans med verksamheter för att utveckla både värdefulla användarfall och lösningar
- Utveckling av användargränssnitt
- Fortsatt utveckling av presidiokonfiguration för bättre precision och fler olika syften.



Tack för oss

Viktor Sjögren,
Mönsterås kommun