Het hoe en waarom bij het gebruik van A.I.

Civinc faciliteert chat gesprekken tussen deelnemers met verschillende meningen. Er kan gebruik worden gemaakt van AI om algemene, niet-herleidbare inzichten te vergaren over het sentiment. Hier leggen we dat uit.

Context & doel

Tijdens een Civinc sessie worden deelnemers gekoppeld in een anonieme een-op-een chat met een partner met een ander perspectief. Dat gebeurt aan de hand van stellingen waar de partners een andere mening op hebben. Als A.I. staat ingesteld voor een sessie worden alle berichten per stelling geaggregeerd en met A.I. geanalyseerd, waarna de chats automatisch en permanent worden verwijderd. Het doel van het gebruik van AI is uitsluitend om organisator en deelnemers een beter algemeen begrip te geven t.a.v. hetgeen er leeft binnen de groep en de basis te leggen voor verdieping en (gezamenlijke) reflectie naderhand.

AI in de praktijk & principes

De analyse van de geaggregeerde chat berichten vindt plaats op basis van prompts, specifieke instructies aan het Large Language Model (LLM) voor bepaalde analyses. Zo kunnen we bijvoorbeeld de instructie geven om een samenvatting van alle gesprekken over een bepaalde stelling te maken. Andere prompts hebben betrekking op bijvoorbeeld het identificeren van de meest voorkomende algemene ideeën, argumenten of twijfels die zijn geuit in de gesprekken. Er zijn een aantal principes die ten alle tijden gelden:

- AI analyse is altijd optioneel. Als een klant of admin besluit geen AI te willen gebruiken dan kan het simpel worden uitgeschakeld per sessie.
- De AI analyse wordt alleen gerund op een stelling als er een minimum aantal berichten is verstuurd in meerdere verschillende gesprekken, om een kritische massa te waarborgen.
- De output is altijd generiek van aard, bevat nooit directe quotes of persoonlijke informatie en kan nooit worden herleid tot individuele gebruikers.
- De chats zelf zijn altijd alleen zichtbaar voor de deelnemers en worden permanent verwijderd na de sessie / AI analyse.

LLM verwerking en privacy waarborgen

Civinc gebruikt de Mistral AI API als voornaamstes aanbieder van frontier AI-analyse uit Europa, met specifieke waarborgen voor privacy en gegevensbescherming:

- De data wordt uitsluitend verwerkt op servers in de EU
- De data is niet beschikbaar voor andere klanten, wordt niet gebruikt om AI-modellen te verbeteren en wordt niet gebruikt om producten of diensten van derden te verbeteren
- De data wordt uitsluitend verwerkt voor de doeleinden zoals door Civinc bepaald en wordt binnen 30 dagen permanent verwijderd van de servers van Mistral
- Aangezien Civinc in beginsel geen gegevens verzamelt die herleidbaar zijn tot een individuele gebruiker, wordt dergelijke data ook niet gedeeld met Mistral
- Voor meer over verwerking en privacybeleid: Mistral AI.
- Voor meer over duurzaamheidsbeleid: Mistral AI
- Voor meer over Responsible A.I.-maatregelen: Mistral AI

Kwaliteitsmonitoring, beperkingen en stellingen formuleren

Omdat Civinc gebruik maakt van een extern Large Langage Model (LLM), voeren we testen uit om de kwaliteit van de data output te waarborgen. Tegelijkertijd kan de output onjuistheden bevatten. Data rapportages van Civinc bevatten naast AI inzichten ook andere data inzichten, zoals de geaggregeerde stellingnames en reacties op survey vragen. De AI analyse is daarmee altijd onderdeel van een groter geheel en dient ook zo beschouwd te worden.

Tot slot kunnen Civinc admins de hulp van AI inschakelen om stellingen te formuleren voor Civinc sessies. Hiervoor wordt dezelfde LLM gebruikt. Alleen de admin input / het verzoek wordt gedeeld met de verwerker (geen persoonlijke of andere informatie).

