

# Testplan VERA (Chatbot)

## Doel van het testplan:

Sinds 2015 is er in het Spaarne Gasthuis, na faillissement van het MC Slotevaart, een nieuwe afdeling met obesitas patiënten bij gekomen. Dit is een database van 7000 patiënten waar jaarlijks 1700 patiënten bijkomen. Ondanks dat het Spaarne Gasthuis er naar streeft om deze patiënten binnen 5 jaar op hun eigen voeten te laten staan, kiezer er veel patiënten voor om bij het ziekenhuis ingeschreven te blijven. De rede dat deze patiënten bij het ziekenhuis blijven is dat het Spaarne goede zorg aanbiedt en de band tussen arts en patiënt goed is. Dit allemaal resulteert op een steeds groter wordende afdeling binnen het Spaarne Gasthuis.

Uit onderzoek is gebleken dat er op de obesitas afdeling veel vragen binnenkomen van patiënten. Dit zijn veel voorkomende vragen zoals: “Mag ik ... eten?” of “Geneest mijn wond op de juiste manier?” Deze vragen komen op twee manieren binnen, via de telefoon en via MijnSpaarne (een platform voor patiënten). Doordat er zoveel patiënten zijn die een vraag hebben, kan er op de telefonie soms wel een wachtrij ontstaan van 30 minuten. Daarnaast beantwoorden de werknemers vaak dezelfde vragen, waardoor ze minder tijd hebben om vragen te beantwoorden die meer tijd en aandacht verdienen. Tot slot wordt het systeem MijnSpaarne nauwelijks tot niet gebruikt. Wanneer het systeem wel gebruikt wordt, duurt het vaak erg lang voordat er tot een conclusie is gekomen van de gestelde vraag.

Naar aanleiding van dit vooronderzoek hebben wij Vera ontworpen. Vera helpt een patiënt bij alle vragen over voeding. In combinatie met MijnSpaarne kan ze gericht antwoord geven op vragen zoals: “Mag ik een frikandel in de blender stoppen?”. Doordat Vera in jouw medisch dossier kan kijken, weet zij hoe ver jij bent in je herstelproces.

Het prototype van Vera is nog verre weg van af. De link met MijnSpaarne is nog niet gelegd, dus dit zullen wij moeten faken. Dit is verder geen probleem buiten het feit dat wij de mensen die aan het testen zijn zo nu en dan moeten storen.

Om erachter te komen of onze prototype werkt hebben wij een paar onderzoeksvragen opgezet waar wij achter willen komen:

Hoe ervaart de gebruiker de dialoog van Vera in haar huidige status?

## **Methode:**

Om antwoord te krijgen op onze vraag, zullen we een drietal aan scenario's opzetten. We vragen een x-aantal mensen of ze door het scenario heen willen lopen. Door middel van observatie en vragen achteraf kunnen we onze onderzoeksvraag beantwoorden.

## **Scenario 1:**

Twee dagen geleden heb je een maagverkleining gehad. Het is voor jou nog een beetje onduidelijk wat je wel en niet mag eten. Voor de operatie at je rond een uur of 8 een bakje chips en je vraagt je af of dit na de operatie ook mag eten. Vraag raad aan Vera.

## **Scenario 2:**

Het is inmiddels al 16 dagen geleden dat je een maagverkleining hebt gehad. Je begint je af te vragen of je inmiddels al weer een boterham mag eten. Vraag raad aan Vera

## **Scenario 3:**

Je bent inmiddels 92 dagen in je herstelproces en je bent bijna jarig. Nu wil je voor je gasten koffie en gebak serveren en gaan er plakken leverworst rond. Dit is natuurlijk super verleidelijk en vraag jij je af of je dit wel tot je mag nemen? Vraag raad aan Vera

## **Vragen:**

- Heb je antwoord op je vraag gekregen?
- Hoe vind jij de chatbot momenteel functioneren?
- Wat vond je van de lengte van de antwoorden
- Hoe kwam de bot over, menselijk of meer robot-like
- Heb je verbeterpunten voor de chatbot?

## Ronde 1:

### Ruud Vissers:

1. Nee
2. Niet zo goed, begreep zijn vraag niet terwijl eigenlijk heel simpel is
3. Prima lengte antwoord
4. Robot-like
5. Geen verbeterpunten

### Extra:

6. Wat was prima? Hij reageert lekker snel.
7. Ruud had niet idee dat het een robot was maar ene mens (op basis van de chat). Ruud kiest er wel liever om met de chat te gebruiken dan 30 minuten wachten.

### Arne Reijntjes:

1. Ja, hij denkt van wel.
2. Het was maar een vraag, dus ja, hij functioneert goed. Hij heet maar op een manier beantwoord. "Boterham" "eten" alleen doet ie niet. Wat mag ik dan wel eten werkt goed.
3. De tekst is wel te lang als je zou moeten luisteren, met een chat is het wel handiger. Het is geen super probleem. Als ie korter zou zijn zou het beter zijn, losse punten
4. Bot kwam menselijk over. Combinatie van menselijk en robot kan beter zijn.
5. Beknopte antwoord (overzichtelijk niet te veel tekst bullet points). Een link in de chatbot met het dieet schema kan handig zijn.

### Ruud de Ronde:

1. Hij heeft antwoord gekregen
2. Reageert snel, standaard dingen zoals producten is basis.
3. De antwoorden waren lang maar er staat niet te veel informatie.
4. Het is duidelijk een bot maar het komt menselijk over, wilt meer een menselijke bot.
5. Ga vooral door.

## **Ronde 2:**

Voor ronde 2 hebben we een paar aanpassingen gemaakt in de dialoog. De trainings phrases zijn nu wat minder specifieke Deze aanpassingen hebben we gebaseerd op de resultaten van de interviews.

### **Roel Snijders:**

**Heb je antwoord op je vraag gekregen?**

Ja ik heb antwoord gekregen.

**Hoe vind jij de chatbot momenteel functioneren?**

Uitstekend, het functioneert goed.

**Wat vond je van de lengte van de antwoorden?**

Er zat alle benodigde informatie in dus is de lengte goed.

**Hoe kwam de bot over, menselijk of meer robot-like?**

Het kwam over als een mens.

**Heb je verbeterpunten voor de chatbot?**

Nee.

### **Evi Snijders:**

**Heb je antwoord op je vraag gekregen?**

Ja.

**Hoe vind jij de chatbot momenteel functioneren?**

Prima. Het lijkt net alsof je een berichtje stuurt naar een andere persoon en zo een 'normaal' gesprek voert.

**Wat vond je van de lengte van de antwoorden?**

Ook prima. Er was een variatie in zinslengte en zinsopbouw.

**Hoe kwam de bot over, menselijk of meer robot-like?**

Menselijk, ik zou het zo geloven als mij verteld werd dat een ander mens deze berichtjes heeft teruggestuurd in plaats van een bot.

**Heb je verbeterpunten voor de chatbot?**

Nee.

**Jacqueline van Gerwen:**

**Heb je antwoord op je vraag gekregen?**

Nee bij maar 1 van de 3 vragen heb ik antwoord gekregen

**Hoe vind jij de chatbot momenteel functioneren?**

Niet goed, doordat ik twee het aantal dagen moest invullen en bij de andere vragen zelfs geen antwoord kreeg.

**Wat vond je van de lengte van de antwoorden**

Duidelijk, houdt persoonlijk ook van informatie. Ooit lang maar niet storend dus.

**Hoe kwam de bot over, menselijk of meer robot-like**

Robot-like, door de mankementen in het systeem. Antwoorden waren wel menselijk.

**Heb je verbeterpunten voor de chatbot?**

Ziet waarde in de chatbot, maar kan het niet goed beoordelen door de mankementen. Vooral storend door het opeens verschijnen van Engelse woorden.

**Melanie van Gerwen:**

**Heb je antwoord op je vraag gekregen?**

Ja, maar onduidelijk bij scenario drie of het nu over het gebak of leverworst ging.

**Hoe vind jij de chatbot momenteel functioneren?**

Goed directe antwoorden.

**Wat vond je van de lengte van de antwoorden**

Prima, doordat er ook meteen uitgelicht wordt waarom iets niet mag.

**Hoe kwam de bot over, menselijk of meer robot-like**

Het was wel duidelijk dat het een bot was, maar de antwoorden hadden ook van een mens kunnen zijn.

**Heb je verbeterpunten voor de chatbot?**

Duidelijk aangeven over welke producten het nu gaat, wanneer er meerdere producten in de vraag worden gesteld.

## Ronde 3 (Klasgenoten/ experts):

### Martijn Fleurkens:

1. Er is deels antwoord gegeven de eerste vraag gebak wel en de tweede leverworst niet
2. In het begin is het engels, toen dacht hij dat het in het engels moest gaan praten (probeer te fixen) Het was vaag dat ie zei of ie in jouw medisch dossier mocht. Knoppen zou wel chill zijn. Of dit product van maken of het product invullen geen [].
3. Het antwoord opdelen in twee delen. Eerst zeggen dat ie het afraad. En vervolgens extra informatie laten zien. Heeft u nog een ander vraag zou je in een nieuw bericht kunnen zetten
4. Human-like, zou een assistent bij een ziekenhuis kunnen zijn. Vind Martijn wel prettig, zolang ie maar weet dat het een chatbot is. (Wordt ook aangegeven in het begin)
5. Fixen dat je twee vragen kan stellen en knoppen is chill.

### Estee Verhorevoort:

1. Ja, maar ze werd wel benieuwd naar wat ze niet mocht hebben, en dat kreeg ze niet te horen. Ze zeggen duidelijk: dit voedsel is geschikt. Wat dan niet?
2. De chatbot functioneert prima, er zijn geen standaard zinnen die je krijg met fallback
3. Het is heftig lang. Het is te lang. Je kan het niet snel opzoeken, je moet het lang overtypen of schrijven
4. De chatbot kwam meer menselijker over. Vind ze wel fijn, omdat het minder overkomt als een voorgeprogrammeerd iets.
5. Het is goed dat er iig niet hoeft in te loggen op MijnSpaarne, dat je ook zelf je aantal dagen kan invoeren. Iets minder tekst, iets beknopter.

## **Conclusie:**

Uit ons onderzoek is gebleken dat ons chatbot verreweg van af is. Het dialoog dat geïmplementeerd is in Dialogflow is erg lineair waardoor onze testpersonen niet goed door het dialoog kwamen. Daarnaast is de chatbot erg beperkt in zijn kunnen, wat betekend dat er maar een aantal opties mogelijk zijn: “Mag ik ... eten”, “Ik heb ... gegeten? Wat nu?” en “Wat mag ik eten?”. De beperkte en lineaire vragen hebben we een beetje kunnen verbeteren, zo kunnen er nu vragen worden gesteld zoals: “Is het goed als ik ... eet” of “Ik heb honger en wil graag een ... eten”.

Tot slot hebben we kunnen concluderen dat de chatbot te onstabiel is om in de eindpresentatie te demonstreren. We doen het onszelf niet ten goede. Daarom hebben wij besloten om in plaats van een demonstratie te geven een video te laten zien hoe VERA zou moeten functioneren.