Interactie met adviesrobots
Vergroot het vertrouwen tot handelen bij het uiteindelijke advies

Voorlopige uitkomsten van een gebruikersstudie

DEFINITIEF
7 juni 2022

Jeroen van Grondelle
Marcel Stalenhoef
Marlies van Steenbergen
Evelien Besseling

Dit rapport deelt voorlopige uitkomsten uit een design thinking cyclus naar adviesrobots die interactieve exploratie van het gegeven advies mogelijk maken en zelfs aanmoedigen. Dat lijkt bij een flink aantal gebruikers te leiden tot vertrouwen tot handelen bij het uiteindelijke advies. En ze ervaren over het algemeen voldoende controle en invloed op het tot stand komen van het advies.
De gepresenteerde observaties komen uit een eerste analyse van een kwalitatieve gebruikersstudie met 11 reizigers van openbaar vervoer aanbieder RET. Zij zijn geobserveerd bij het gebruiken van aantal varianten van een adviesrobot bij het kiezen van het best passende abonnement en zijn daarna geïnterviewd. De uitkomsten zijn voorlopig, in de zin dat de degelijke annotatie en analyse en daaruit eventueel volgende wetenschappelijke conclusies nog niet beschikbaar zijn.

Vraagstuk

Veel organisaties bieden hun klanten en burgers advies modules bij het maken van complexe keuzes. Deze adviesrobots vragen gebruikers meestal naar hun eigen doelen, situatie en voorkeuren, om daarna met behulp van een algoritme te bepalen waar de burger recht op heeft, welk product of regeling het best bij hen past, etc.
In de praktijk blijkt dat veel burgers na het gebruik van zo'n adviesrobot niet het benodigde vertrouwen tot handelen hebben om het advies ook echt op te volgen. In veel gevallen wordt dan alsnog contact opgenomen met een medewerker. Waar vaak blijkt dat het originele advies op zich klopte.
Zou een interactiever proces om tot het advies te komen leiden tot een hoger vertrouwen in het advies en een hoger vertrouwen om ernaar te handelen? Bijvoorbeeld omdat meer begrip van het onderliggend vraagstuk ontstaat? Of begrip van hoe de robot tot zijn advies kwam? Of omdat niet kloppende aannamen/overtuigingen kunnen worden uitgeprobeerd en duidelijk wordt waarom een andere optie/advies beter passend is?

De interventie: Interactieve exploratie met een adviesrobot

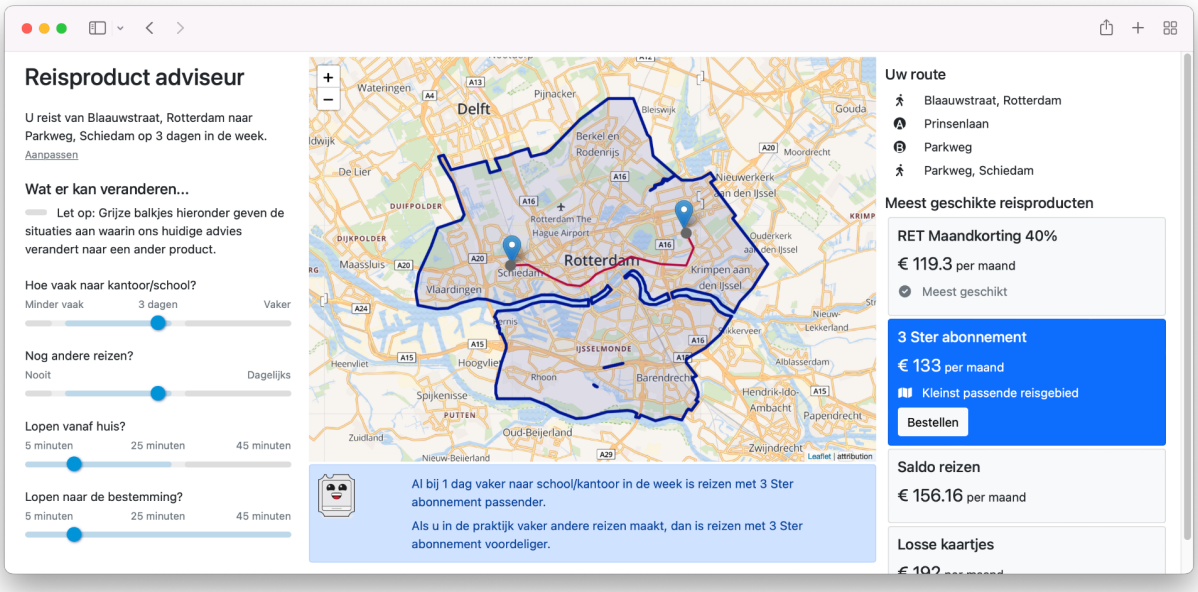
In een design thinking cyclus is een aantal varianten uitgewerkt van manieren waarop een adviesrobot interactieve exploratie mogelijk kan maken.

Scenario	Beschrijving
Baseline: Formulier met twee pagina's: Uw situatie en vervolgens Ons advies	Veel adviesrobots scheiden het uitvragen van de situatie (en daarmee ook doelen, wensen en voorkeuren) enerzijds, en het daarna tonen van het eindadvies. Meestal staan die twee op eigen pagina's, en kan alleen met een "Terug" knop van het advies worden "teruggekomen" op de gegevens die eerder ingevoerd zijn.
1: Interactieve exploratie mogelijk maken door op dezelfde pagina wijzigen van situatie en tonen van (onmiddellijke) impact op advies	In deze interventie worden de vragen over de situatie en het tonen van het advies dat daarbij past gecombineerd op één pagina. En leidt het bijstellen van de situatie beschrijving tot onmiddellijke bijstelling van het advies.
2: Guided exploratie middels bandbreedtes op sliders	In een variant van de interventie wordt er in de achtergrond bijgehouden door een algoritme wat de bandbreedte van het advies is. De bandbreedte bestaat uit alle andere situaties die tot hetzelfde advies leiden. Die bandbreedtes worden vervolgens visueel getoond op de sliders, zodat gebruikers geholpen worden in de exploratie.
3: Guided exploratie middels chatbot meldingen over grensgevallen	In een andere variant wordt op vergelijkbare wijze bepaald in hoeverre de beschreven situatie een grensgeval betreft. En dus een kleine wijziging al zal leiden tot een ander advies. In die gevallen verschijnt een chatbotje dat in natuurlijke taal dat grensgeval toelicht.

Tabel 1: De 3 ontworpen interventie varianten.

Klantonderzoek met Reizigers

Hoewel in het ontwerpproces interactieve invullingen van adviesrobots in meerdere adviesvraagstukken verkend zijn, is uiteindelijk een casus uitgewerkt rond het kiezen van een OV abonnement. Daarvan bestaan allerlei vormen, van losse kaartjes tot percentuele kortingsproducten tot gebiedsgebonden abonnementen. Daarbij speelt het precies af te leggen traject een grote rol, maar kunnen kleine wijzigingen in de reisfrequentie in werk en vrije tijd en de bereidheid om een stukje te lopen/fietsen leiden tot slimmere opties waarmee geld is te besparen. De complexiteit van de verschillende producten leidt er op dit moment toe dat veel reizigers na gebruik van de online reisproduct adviseur alsnog contact opnemen met de klantenservice.



Figuur 2: Het uitgewerkte prototype voor reisproduct adviseur

In een kwalitatief gebruikersonderzoek is aan elf reizigers gevraagd om met behulp van een werkend prototype het best passende abonnement te kiezen bij een aantal door ons aangereikte situaties. Daarbij zijn, in wisselende volgorde, alle beschreven varianten van de interventie aangeboden. Na het uitvoeren van die taak is de reizigers gevraagd naar hun ervaringen.

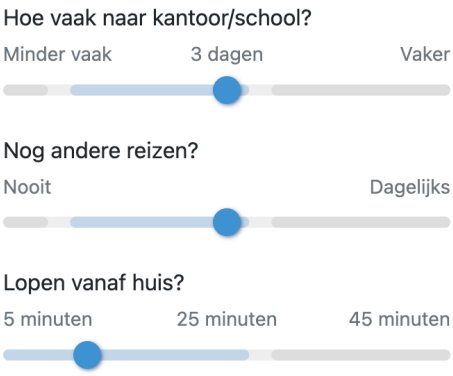
Waargenomen overeenkomsten en verschillen

Gebruikers herkennen in elk van de drie scenario's de functie van de sliders en gebruiken deze vrijwel altijd om het effect van wijzigingen op het advies te verkennen.

Die gebruikers die het baseline scenario als eerste scenario krijgen aangeboden verkennen in de meeste gevallen geen alternatieven. Gebruikers die het baseline scenario in een later stadium krijgen, gebruiken in veel gevallen de terug-knop om toch ook in dit scenario alternatieven te verkennen zoals ze dat in de eerder aangeboden scenario’s met de sliders deden.

De guided exploration in variant 2 en 3 zijn beide gebaseerd op gevoeligheidsanalyse van het advies. Daarbij wordt door een algoritme automatisch geanalyseerd of kleine wijzigingen in de invoer leiden tot geen, kleine of zelfs grote wijzigingen in de uitkomsten. In het geval van een adviesrobot gaat het om de vraag of een licht veranderde situatie of behoefte tot hetzelfde advies zou leiden of dat het initiële advies dan helemaal niet meer van toepassing is. Zo ontstaat als het ware een bandbreedte van houdbaarheid van het advies.

Variant 2 toont de bandbreedte visueel op de sliders, zoals getoond in figuur 3. Dat wordt door een grote meerderheid niet gezien en na uitvragen slechts door een minderheid begrepen. Die heeft in deze vorm geen toegevoegde waarde.



Figuur 3: De bandbreedte van het advies in de huidige situatie

Variant 3 toont in natuurlijke taal de grensgevallen en wordt goed begrepen. Een voorbeeld hiervan staat in figuur 4. Deze variant heeft in een flink aantal gevallen een activerend en uitnodigend effect. De gebruiker wordt op het idee gebracht iets in de situatie te veranderen, waar men wellicht zelf niet aan zou hebben gedacht.

Figuur 4: Een voorbeeld van een uitleg van een grensgeval

Enige keerzijde die een aantal respondenten noemt is dat de suggestie in de praktijk soms niet kan: “Ik kan helemaal niet een dag minder gaan werken...”. Een flink aantal deelnemers geeft in het interview deze variant de voorkeur.

Aantal uitspraken uit de reflectie

De codering van de interviews is nog niet gedaan. Daarom delen we hier in dit stadium slechts een paar relevante uitspraken die ons opvielen.

Een van de onderwerpen in de interviews was in hoeverre de respondent nu in staat zou zijn om een goede keuze te maken. Een aantal respondenten benoemt daarbij zaken rond het goed begrijpen van wat het beste is, bijvoorbeeld “Ik kan zien wat het me opbrengt of kost als ik er creatief mee om ga. En kan op basis hiervan mijn keuze maken.”. En: “Ik kan mijn eigen keuzes duidelijk maken en de consequenties ervan zien”. En ook: “Ik word op het idee gebracht om iets te veranderen in mijn situatie, waar ik misschien zelf niet aan had gedacht”.

Het lijkt erop dat het gegeven dat de respondent zelf de specifieke situatie kan invullen, hier keuzes in kan maken en aanpassen, het gevoel van controle en daarmee het gevoel van vertrouwen in het advies vergroot. Anderen gebruiken woorden die te maken hebben met vertrouwen in het advies van de RET: “Het wordt geadviseerd, dus het zal wel goed zijn”, “Is er een addertje onder het

gras?”, "Deze tool gebruikt gelukkig NIET de booking.com manier”, of zijdelings bij het invullen: “Ik trap er weer in”.

De respondenten ervaren over het algemeen voldoende controle over de totstandkoming van het advies. “Het botje reikt wel aan, maar voert niet vanzelf door”.

Een aantal respondenten benoemen verschillende aspecten van het feit dat alle producten worden getoond en niet alleen de (volgens RET) “relevante”:
Dat is “mooi onafhankelijk, dat geeft geen sturing vanuit RET”. Maar roept bij anderen ook verwarring op: “Losse kaartjes is toch geen serieuze optie hier?”

De kleine lettertjes

Dit rapport bevat de voorlopige uitkomsten van het beschreven onderzoek. Daarin hebben de onderzoekers hun aantekeningen en observaties naast elkaar gelegd en geprobeerd een aantal wezenlijke overeenkomsten en veelzeggende verschillen die ze waarnamen te benoemen. Dit rapport is NIET gebaseerd op de grondige annotatie, interpretatie en analyse die later gaat worden gedaan om wetenschappelijke conclusies te verbinden aan het onderzoek. Hoewel de onderzoekers verwachten dat de hier gedeelde beelden in grote lijnen wel overeind zullen blijven, kan dat natuurlijk niet worden gegarandeerd.