



Northern Periphery and
Arctic Programme
2014-2020

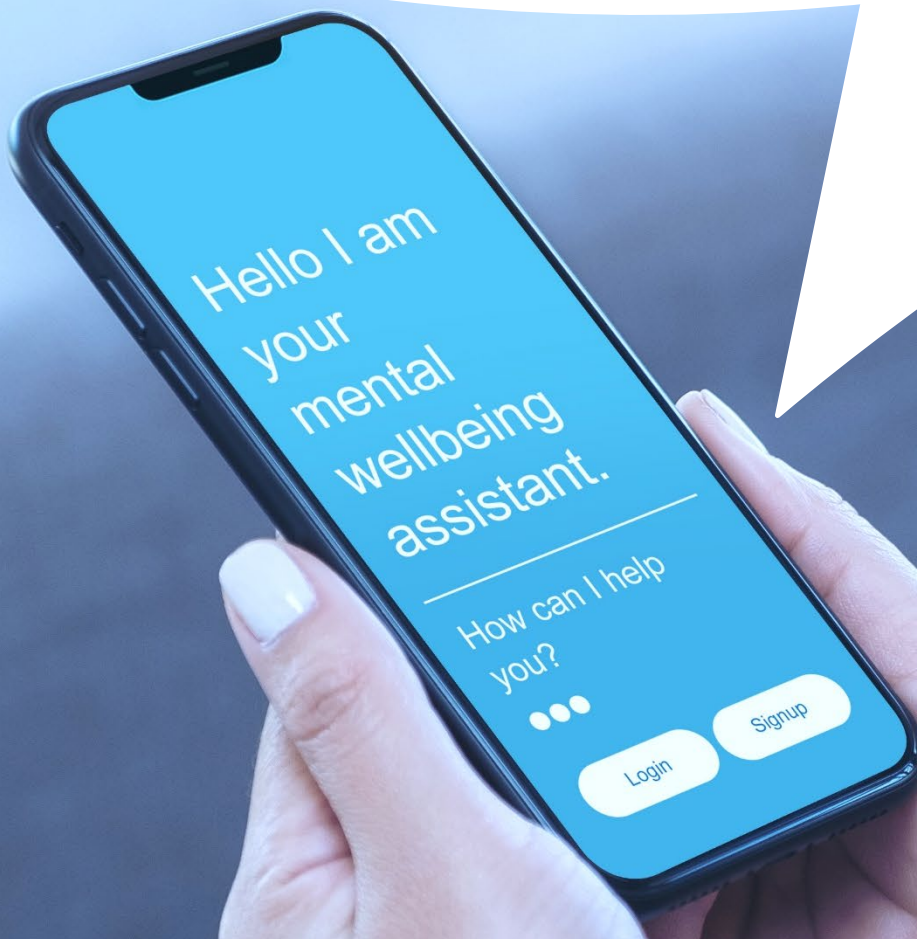


EUROPEAN UNION
Investing in your future
European Regional Development Fund

ChatPal



Digitala samtal för ökat psykiskt välbefinnande



Verktyg för samskapande av chattbotar som ska användas i hälso- och sjukvården

Courtney Potts, Raymond Bond, Maurice Mulvenna, Michael McTear, Fred Booth, Frida Lindström, Catrine Kostenius, Indika Dhanapala, Alex Vakaloudis, Brian Cahill, Kyle Boyd, Thomas Broderick, Andrea Bickerdike, Con Burns, Edward Coughlan, Heidi Nieminen, Anna-Kaisa Vartiainen, Lauri Kuosmanen.



Författare: Courtney Potts¹, Raymond Bond¹, Maurice Mulvenna¹, Michael McTear¹, Fred Booth¹, Frida Lindström², Catrine Kostenius², Indika Dhanapala³, Alex Vakaloudis³, Brian Cahill³, Kyle Boyd⁴, Thomas Broderick⁵, Andrea Bickerdike⁵, Con Burns⁵, Edward Coughlan⁵, Heidi Nieminen⁶, Anna-Kaisa Vartiainen⁶, Lauri Kuosmanen⁶

¹ Ulster University, School of Computing, Belfast, Storbritannien

² Luleå Tekniska Universitet, Institutionen För Hälsovetenskap, Luleå

³ Munster Technological University, Nimbus Research Centre, Cork, Irland

⁴ Ulster University, School of Computing, Belfast, Storbritannien

⁵ Munster Technological University, Department of Sport, Leisure and Childhood Studies, Cork, Ireland

⁶ University of Eastern Finland, Department of Nursing Science, Kuopio, Finland

© 2022 University of Ulster

Alla rättigheter förbehålls. Ingen del av denna publikation får reproduceras eller överföras i någon form eller media på något sätt, elektroniskt, mekaniskt, via fotokopiering, inspelning eller på annat sätt, utan skriftligt medgivande av universitetet eller författarna.

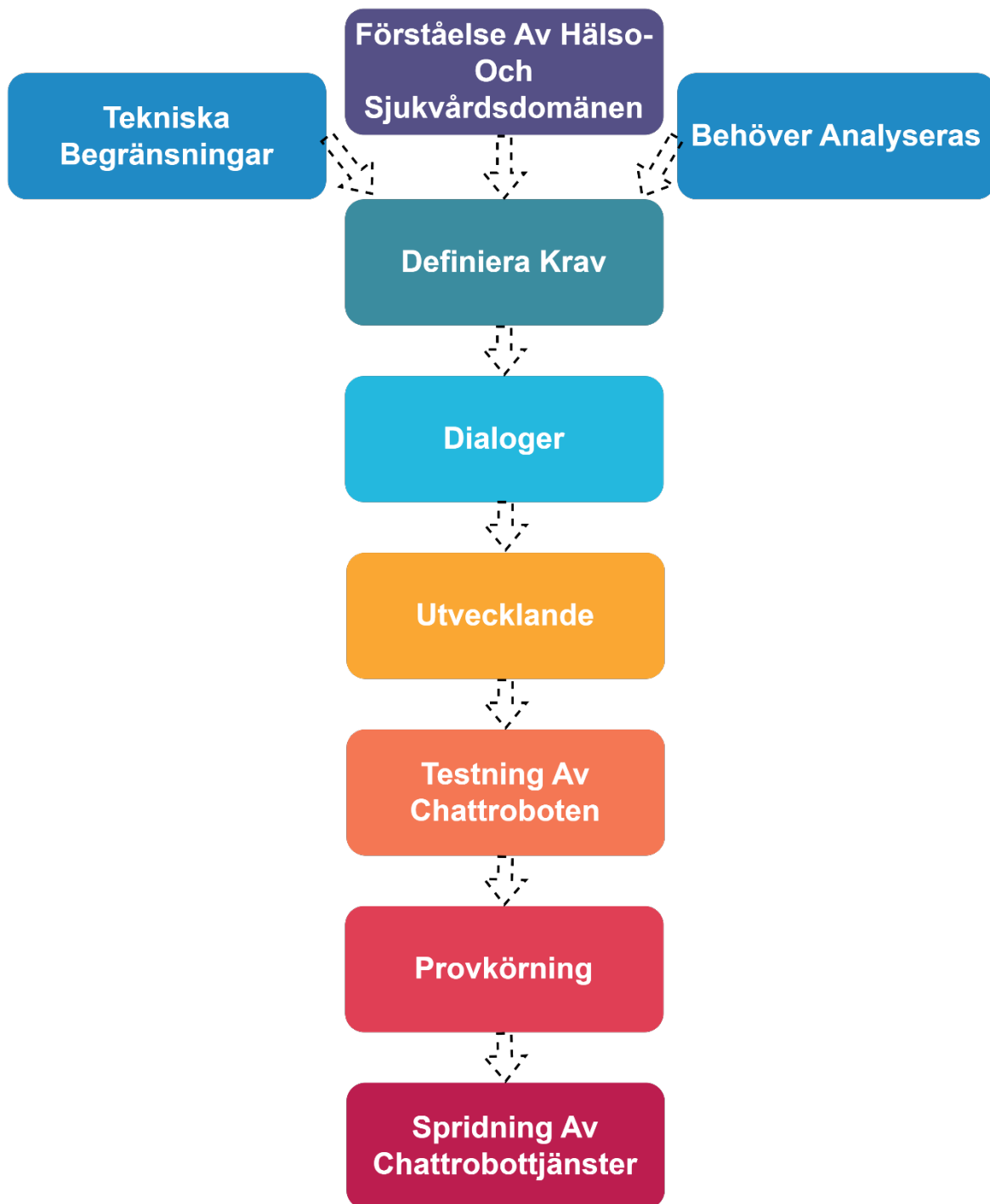
ISBN-13: 978-1-85923-290-3

Innehållsförteckning

Arbetsgång Vid Utformningen Av En Chattbot För Hälsovård	4
Intressentcenterad Design	6
Förståelse Av Hälso- Och Sjukvårdsdomänen	8
Tekniska Begränsningar	10
Behöver Analyseras	12
Definiera Krav	13
Dialoger	14
Utvecklande	18
Testning Av Chattroboten	20
Provkörning	21
Spridning Av Chattrobotjänster	28
Referenser	31
Checklista Inför Utformningen Av En Chattbot För Hälsovård	32

ARBETSGÅNG VID UTFORMNINGEN AV EN CHATTBOT FÖR HÄLSOVÅRD

INTRESSENTCENTERAD DESIGN



Denna verktygssats ger riktlinjer för att utforma och utveckla en chattbot för hälso- och sjukvård.

Här hittar du rekommenderade aktiviteter och metoder som användes för att anta ett intressentcentrerat tillvägagångssätt, vilket är viktigt för ansvarsfull utformning av digital hälsoteknik.

För varje avsnitt nedan kan du hitta ytterligare användbar information, föreslagna steg att följa och förväntade resultat om du följer våra rekommendationer. Dessa rekommendationer baseras på våra erfarenheter genom hela ChatPal-projektet.

Du kan läsa mer om projektet här:

<https://www.youtube.com/channel/UCeYBbGKE0CGx08L311YZgRw>

eller genom att skanna QR-koden.



INTRESSENTCENTERAD DESIGN

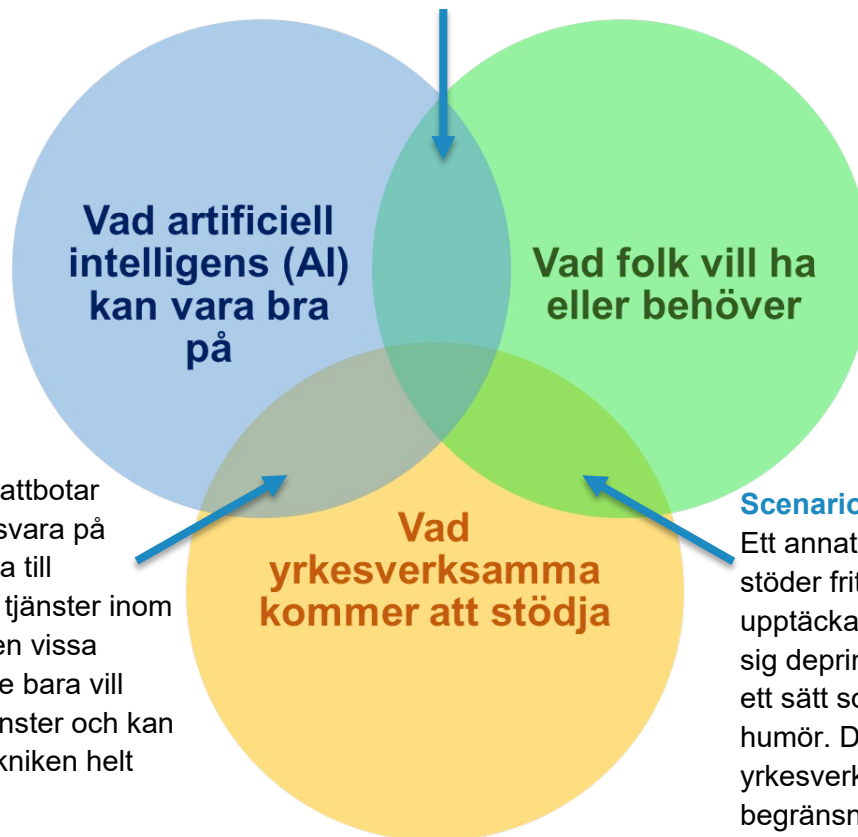
Vid utformningen av en chattbot för hälso- och sjukvård bör man överväga tre saker:

1. Vad artificiell intelligens (AI) kan vara bra på
2. Vad folk vill ha eller behöver
3. Vad yrkesverksamma kommer att stödja

Vi föreslår aktiviteter och metoder som balanserar dessa tre saker:

Scenario 1

Chattbotar kan utan problem hantera skrivna dialoger med fördefinierade svar eller begränsade fritextsvar. Om användarna ville ha en chattbot för att självdiagnostisera eller undersöka, kunde den samla in symtom och använda ett beslutsflöde för att föreslå en diagnos. Det kan dock hända att yrkesverksamma inte stöder detta eftersom det kan resultera i en falsk diagnos, vilket kan begränsa dess trovärdighet och viljan att använda den.



Scenario 2

Alternativt kan chattbotar användas för att svara på frågor och hänvisa till kostnadsbelagda tjänster inom psykisk hälsa, men vissa användare kanske bara vill använda gratistjänster och kan därför undvika tekniken helt och hållet.

Scenario 3

Ett annat exempel är en chattbot som stöder fritext, en som försöker upptäcka när en användare känner sig deprimerad och försöker svara på ett sätt som förbättrar personens humör. Detta kan stödjas av yrkesverksamma, men med tanke på begränsningarna hos AI kan svaren vara olämpliga om chattboten inte förstår vad användaren säger eller om den ger olämpliga råd.

Därför kan en framgångsrik digital intervention ses som skärningspunkten mellan vad användarna vill ha och säger att de behöver, vad de yrkesverksamma förespråkar och vad AI gör bra.

FÖRSTÅELSE AV HÄLSO- OCH SJUKVÅRDSDOMÄNEN

Det första steget i planeringen av en chattbot för hälso-och sjukvård bör alltid vara att förstå domänen.

Det kan vara bra att göra en litteraturgenomgång inom det aktuella området för att identifiera vad som redan har gjorts och vilka luckor som finns. I det här skedet är det bra om teknikerna får samtala med vårdpersonal, först och främst för att ta reda på vilka hälso- och sjukvårdstjänster som finns tillgängliga.

Detta gör det också möjligt för vårdpersonalen att dela med sig av sin kunskap med teknikerna, samtidigt som de får en bild av hur tekniken kan användas inom deras domän. Detta är avgörande i samproduktionsprocessen, eftersom tekniken bör anpassas till den domän där den ska användas.

I det här skedet kan det vara användbart att skicka ut enkäter bland vårdpersonal för att samla in åsikter om vad de skulle vilja se i en chattbot eller hur de ser att chattbotar kompletterar hälso- och sjukvårdstjänsterna.

Dessa aktiviteter ger en bra utgångspunkt till utformningen av en chattbot, eftersom det kan hjälpa en att bestämma syftet med den.

Våra rekommendationer

- Gör en litteraturgenomgång
 - Vad har redan gjorts på detta område?
 - Vilka luckor finns i den aktuella litteraturen?
- Rundabordssamtal/gruppdiskussioner med tekniker och vårdpersonal för att dela kunskap
- Bedöm attityder, preferenser och rekommendationer från vårdpersonal, till exempel med hjälp av en enkät (Sweeney et al. 2021)
- Fastställ chattbotens syfte
 - Vad ska chattboten användas till?
 - Vilka är slutanvändarna?
 - Vilka språk kommer att användas i chattboten?
 - Hur ska den komplettera och integreras i nuvarande tjänster?

Förväntade resultat: Litteraturgenomgång, åsikter från vårdpersonal om vilken digital hälsoteknik de skulle stödja.

TEKNISKA BEGRÄNSNINGAR

Nästa steg är att identifiera vilka tekniker som ska användas.

Det vill säga vilken plattform (t.ex. mobil, webbapplikation osv.), modalitet (skriven text eller röst) och ramverk som ska användas för att bygga chattboten. Det finns många bra ramverk med öppen källkod för utveckling av chattbotar, men vilket alternativ som passar bäst beror på vilka funktioner eller scenarier som hälsovårdschattboten ska hantera.

För att bestämma vilken teknik man ska gå vidare med är det bra att konsultera en expert på området. Det är viktigt att överväga för- och nackdelarna med artificiell intelligens (AI) och naturlig språkbehandling (NLP). Det finns två huvudfrågor som rör NLP: *Natural Language Understanding* (NLU), som handlar om att förstå vad användaren säger, och *Response Generation*, det vill säga att ge ett lämpligt svar.

Ett sätt att komma åt dessa svårigheter kan vara att begränsa möjligheterna för användarna att skriva fritext och att istället låta dem välja bland fördefinierade svar. Nackdelen med det är att man inte får veta användarnas egna frågor och svar när de måste välja från förprogrammerade repliker.

Ett sätt att kringgå detta skulle vara att tillåta fri text men att försöka begränsa dialogen med noggrant utformade uppmaningar. Det är en svår fråga och i säkerhetskritiska områden är det bättre att vara försiktig. Det är också bra att planera hur chattboten ska reagera om användaren skulle säga något oväntat.

Detta är särskilt viktigt inom området psykisk hälsa. Om till exempel någon skulle avslöja att det finns risk att hen kommer att skada sig själv så bör chattboten ha ett lämpligt svar, som att uppmuntra användaren att söka extern hjälp och hänvisa till krishjälp. Det har till exempel dokumenterats tidigare under 2018, när populära chattbotar för psykisk hälsa inte kunde hantera rapporter om sexuella övergrepp mot barn (<https://www.bbc.co.uk/news/technology-46507900>).

Nyckelordsutlösare kan användas här för att informera användarna om var de kan få rådgivning och stöd.

En annan studie av Bickmore och kollegor rapporterade om fall där Siri, Alexa och Google Assistant gav dåliga hälsoråd som kunde ha resulterat i en viss grad av skada på patienter. Där ingick även förslag som potentiellt kunde ha lett till dödsfall (Bickmore et al. 2018). Det är viktigt att överväga de etiska frågorna när det gäller samtal mellan människor och datorer och att vara transparent med användarna, så att de inte litar för mycket på informationen som ges av chattboten, samt att hänvisa till extern hjälp eller resurser för att undvika oavsiktliga konsekvenser.

Våra rekommendationer

- Identifiera styrkor och begränsningar för att välja rätt teknik.
- Fokusera på vad tekniken kan göra bra och undvik områden med svagheter (särskilt om chattboten ger hälsorelaterade råd).
- Om noll kostnad är ett alternativ, välj då ett ramverk för utveckling av en chattbot med öppen källkod.
- Överväg att välja en utvecklingsmiljö där åtkomsten till data är privat och fristående (se till att tredje parter som sociala medieplattformar inte har användarnas hälsodata när de interagerar med chattboten. Om detta inte är möjligt se till att användarna får veta det i förväg och samtycka till vilka organisationer som kommer att ha deras uppgifter).
- Överväg att implementera nyckelordsutlösare och planera för meddelanden/chattbotsvar för oväntade kommentarer och känsliga ämnen.

Förväntade resultat: Slutligt beslut om enhet, modalitet och plattform. Riskbedömningsplan för att hantera potentiella etiska problem.

BEHÖVER ANALYSERAS

Det är viktigt att förstå vilka behov och krav de avsedda användargrupperna har när det gäller digitala samtal om hälsan.

Det första steget är att identifiera relevanta intressenter, eftersom detta är en integrerad del av den intressentcentrerade designprocessen.

Gruppdiskussioner eller enkäter med intressenter och målanvändare kan utnyttjas i detta skede för att undersöka vilka behoven är (Potts et al. 2021).

Våra rekommendationer

- Identifiera intressenter.
- Arrangera gruppdiskussioner med användargrupper (eller grupper med liknande demografi som förväntade slutanvändare) och intressenter för att identifiera användarbehov.

Gruppdiskussioner

- Diskussion kring slutanvändarnas vårdbehov
- Demonstration av liknande digital hälsoteknik
- Utveckling av chattbot-persona (kön, ålder, personlighet och karaktärsdrag)
- Samskriva samtalsdialoger som de skulle vilja se i en chattbot och/eller frågor som de skulle vilja att den ska ställa
- Insamling av användarberättelser som en kravspecifikationsmall (dvs. ”Som < typ av användare > vill jag ha < något mål > eftersom < någon anledning >”)
- Alternativt är kartläggning av användargrupper och intressenter ett annat tillvägagångssätt som kan användas för att identifiera användarnas behov.

Förväntade resultat: Användarnas behov och krav.

DEFINIERA KRAV

När de intressentcentrerade designaktiviteterna har slutförts är nästa steg att definiera kraven på chattboten.

Kraven kan härledas från den samlade kunskapen om vad teknik kan göra bra, vad yrkesverksamma stödjer och vad människor behöver eller vill ha. Inledningsvis kan detta vara en lång lista med krav som sedan kan förfinas för de slutliga funktionerna.

Våra rekommendationer

- Använd en samproduktionsmetod för att skapa lösningar som tillgodoser användarnas behov, till exempel
 - tvärvetenskapliga diskussionsgrupper med vårdpersonal och teknikexperter. Om du är värd för online-diskussioner kan virtuella whiteboardtavlor användas för att underlätta samproduktion där deltagarna kan skicka in sina idéer, anonymt (så att alla kan uttrycka sina ärliga åsikter vilket ger mer demokratisk datainsamling).
- Skapa en lista med krav från intressentcentrerade designaktiviteter.
- Prioritera funktioner från kravlistan, till exempel
 - gruppdiskussioner med oberoende röstning och rangordning av krav; Individer från multidisciplinära team kan anonymt rangordna funktioner baserat på upplevd betydelse och hur väl de passar för användarberättelserna.
- Bestäm slutliga chattbot-funktioner baserat på prioriterade krav. Tematisk analys kan användas för att klassificera de slutliga funktionerna i olika tematiska grupper på högre nivå (varje grupp är ett kluster av relaterade funktioner/krav).

Förväntade resultat: En lista över funktioner i hälsovårdschattboten.

DIALOGER

När de sista funktionerna har bestämts är nästa steg att slutföra planer för att utforma och skriva dialogerna.

Interaktionsstilen bör övervägas, inklusive typen (fördefinierad, styrd, öppen) och hur den styrs (systemstyrd, användarstyrd, både och). Om chattboten är textbaserad kan tillägget av multimedia (bilder, videor, giffar) användas för att förbättra användarupplevelsen (UX).

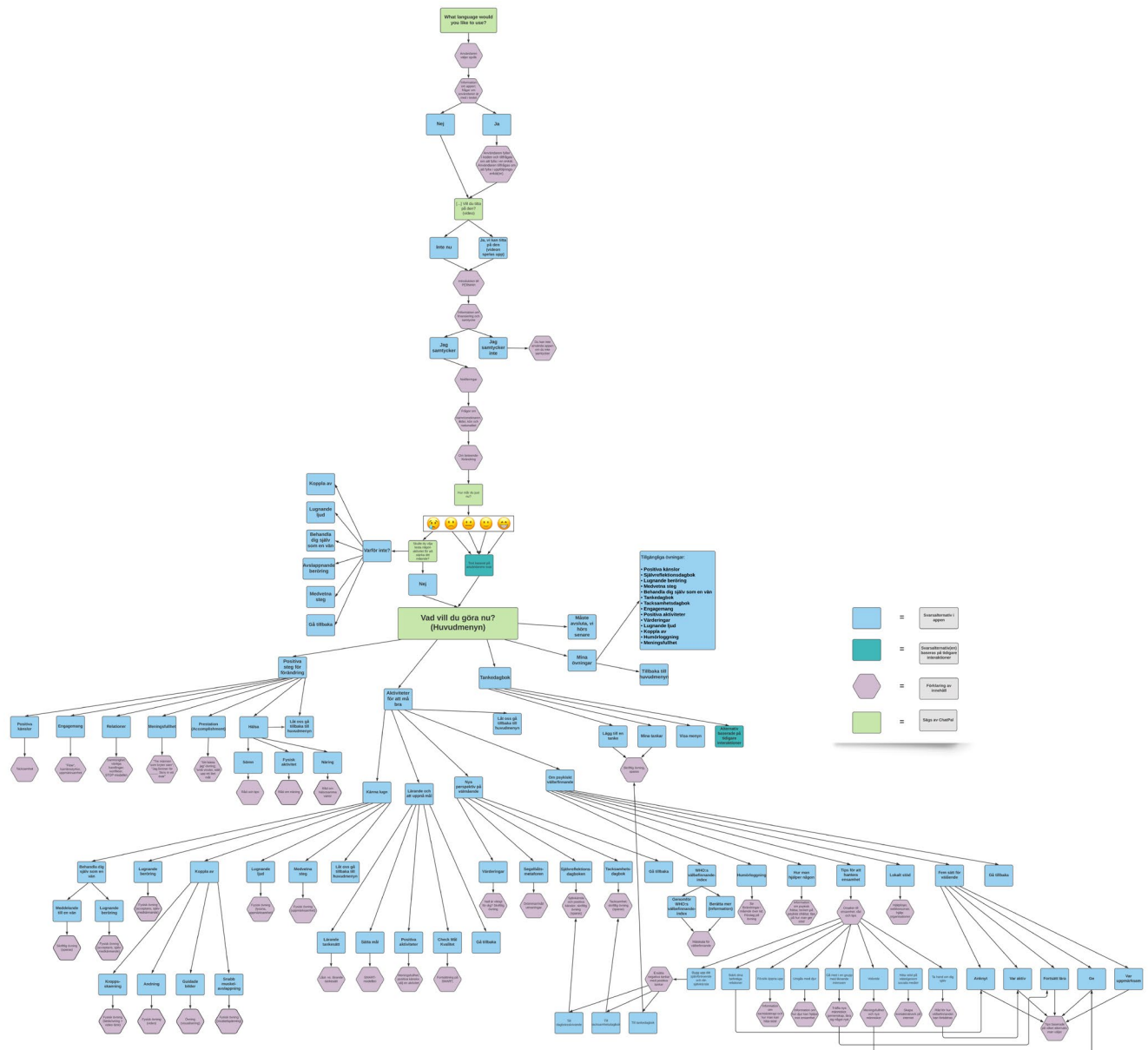
Personifiering kan också förbättra UX så det är viktigt att bestämma om och hur chattboten ska använda personlig information om användaren (namn, ålder, kön, tidigare interaktioner osv). Chattbotar kan också använda extern kunskap, såsom databasregister eller hälsoinformation på nätet vid behov. Det är fördelaktigt att arbeta i ett multidisciplinärt team när dialogerna ska utvecklas, och att även involvera expertis inom samtalsdesign av UX, såväl som vårdpersonal som arbetar direkt med patienter eller klienter och forskande vårdpersonal.

Experter inom hälsa och sjukvård är bäst lämpade att skriva det initiala innehållet för dialoger med chattbotar, men det är viktigt att de följer bra designprinciper (Cameron et al. 2018b). Det är viktigt att informera användarna om begränsningarna hos chattbottekniken i förväg (i början av samtalet), att till exempel göra dem medvetna om att chattbotar bara har begränsad intelligens. När man skriver innehåll bör man ha chattbotens persona i åtanke hela tiden, för att säkerställa konsekvens. Det är också viktigt att tänka på hur användarnas input används. Om en användare till exempel avslöjar information om en händelse och chattboten upprepar detta för användaren många veckor senare kan användaren uppleva att det inte känns bra, eftersom hen kanske har glömt att hen avslöjat det här och/eller därför känner sig övervakad.

När manusen är färdiga ska det slutliga innehållet utvärderas och bekräftas. Då kan det vara användbart att ha ett antal frågor eller kriterier för att kvalitetsbedöma vart och ett av dialogmanusen. Läs mer om samtalsdesign hos Deibel et al 2021. Om chattboten är flerspråkig är det här ett bra tillfälle att börja översätta innehåll när manusen är färdiga.

Ett bra sätt att visualisera allt innehåll i appen är att skapa ett flödesdiagram som visar funktionerna och var man kan komma åt dem. Detta kan delas med slutanvändare eller vårdpersonal så att de tydligt kan se de tillgängliga alternativen och får möjlighet att diskutera dessa med sina klienter.

Flödesdiagrammet som skapats för ChatPal-appen finns nedan som ett exempel. Den här typen av flödesdiagram kan också hjälpa designteamet att identifiera återvändsgränder i chattboten (att det t.ex. finns ett svar om en användare väljer ”Nej” via en snabbsvarsknapp eller om formgivarna råkade missa denna dialogväg).



En pdf-version av diagrammet kan laddas ner från <https://chatpal.interreg-npa.eu/resources/>

Våra rekommendationer

- Fatta de slutliga besluten för utformningen av chattboten när det gäller modalitet och användning av multimedia eller röst, typ av interaktion, personifiering och extern kunskap.
- Genomför multidisciplinär innehållsutveckling och manusarbete under ledning av eller med rådgivning från experter inom samtalsdesign och bästa praxis.
 - Det kan vara gruppdiskussioner där deltagarna utformar dialoger tillsammans som en grupp med hjälp av samarbetsverktyg som Google-dokument.
- Var transparent och informera användarna om tekniska begränsningar i förväg vid introduktionsfasen av chattboten.
- Fundera över hur chattboten använder användardata och om det finns risk för att chattbotens personlighet kan frammana en "kuslig dal" (Mori et al. 2012). Användarna kan känna oro inför artificiella mänskliga representationer i form av chattbotar.
- Översätt det slutliga innehållet till andra språk om det är aktuellt.
- Tillämpa kriterier för att bedöma och signera manusen. Följande frågor kommer från två befintliga appbedömningsverktyg (Enlight - kvalitetsbedömningsavsnitt och användarversion av Mobile App Rating Scale uMARS) och en "bästa praxis"-publikation (Cameron et al. 2018a). Frågan om chattbotens personlighet kommer att vara specifik, beroende på hur den är tänkt att användas och vilken persona som användargrupperna föredrar. Ett enskilt chattbot-samtal bedöms och man tar upp frågor om samtalsflödet, informationens kvalitet, kulturell lämplighet och chattbotens personlighet.

1. Samtalsflöde

- Är innehållet tillräckligt varierat (emojier, giffar osv.)? (Cameron et al. 2018a)
- Ingår det tillräckligt många samtalsinslag som efterliknar verkliga samtal (turtagning, humor/skämt osv.)? (Cameron et al. 2018a)
- Använder chattboten ett språk och formuleringar som är lämpligt för den användning den är avsedd för? (Cameron et al. 2018a)

2. Informationens kvalitet

- Är informationen korrekt? Finns det evidensbaserade tekniker som är relevanta för att uppnå det önskade syftet med programmet? (Baumel et al. 2016)
- Tillhandahålls informationen på ett tydligt och lämpligt sätt för målgruppen? (Baumel et al. 2016)
- Finns det tillräcklig information i samtalet utan några utelämnanden, överförklaringar eller irrelevanta uppgifter? (Baumel et al. 2016)
- Är innehållet i samtalet korrekt, välskrivet och relevant för målet/ämnet för hälsochattboten? (Stoyanov et al. 2016)

3. Kulturella aspekter

- Upplevs samtalet som kulturellt lämpligt? (Cameron et al. 2018a)

4. Persona

Kontrollera att formuleringarna i samtalsutkastet stämmer överens med det önskade:

- Chattbotens ålder
- Chattbotens kön
- Chattbotens personlighetsdrag

Förväntade resultat: Slutligt innehåll i chattboten.

UTVECKLANDE

Utvecklandet av chattboten kan börja när innehållet har slutförts.

Detta steg involverar utveckling och lansering av en så liten prototyp som möjligt och vi rekommenderar ett agilt tillvägagångssätt för detta. Under arbetet med att utveckla chattboten är det en bra idé att registrera allt som händer i en databas för framtida analys. Då kan man utforska användarnas engagemang och vilka funktioner folk använder oftast till exempel.

Man bör överväga att tillåta chattboten att stödja multimedialinnehåll där så är möjligt, och inkludera appviseringar för att uppmuntra användningen av den. Det kan också vara fördelaktigt att använda sig av s.k. ecological momentary assessment (EMA), dvs. frågor som presenteras och besvaras av användare "i nuet".

EMA kan användas för en mängd olika syften som att spåra humör över dygnets timmar (Bond et al. 2019). Ett annat exempel från Goldstein och kollegor är en mobilapp som de utvecklat för att hjälpa till att förebygga dietåterfall hos överviktiga individer genom att be deltagarna att registrera snedsteg i kosten tillsammans med en rad potentiella triggers med hjälp av EMA (Goldstein et al. 2017). Dynamisk beräkningsmodellering användes för att uppskatta risken för återfall i ätbeteendet och interventioner inklusive textmeddelanden för att förhindra detta (Goldstein et al. 2017).

Våra rekommendationer

- Registrera användarhändelser automatiskt så att de kan analyseras.
- Det kan vara hjälpsfullt att använda EMA i chattboten (t.ex. frågor om hur användaren mår just då).
- Tänk på chattbotens förmåga att stödja multimedia och aviseringar för att hålla användarna engagerade och hålla kvar dem.

Förväntade resultat: Chattbot-prototyp.

Skanna QR-koden nedan om du vill veta mer om teorin bakom chattbotar, hur de fungerar och hur vi har skapat ChatPal.

QR-kod som länkar till: <https://www.youtube.com/watch?v=GnPChzAw0CI>



Film där Alex Vakaloudis (Munster Technological University) intervjuas om chattbotutveckling.

TESTNING AV CHATTROBOTEN

När den första prototypen har utvecklats kan chattboten testas.

Då är det användbart att börja med att arbeta i ett multidisciplinärt team och testa systemet för fel, som kan rapporteras till utvecklarna och åtgärdas. När de första problemen har lösts är chattboten redo att testas mer av användare.

Användbarhetstester med ett litet antal deltagare kan vara till hjälp för att få feedback för att optimera användarupplevelsen och utvärdera chattboten tidigt, såväl som under hela utvecklingsprocessen.

När tekniska fel har åtgärdats kan chattboten testas av personer som motsvarar målgruppen.

Våra rekommendationer

- Multidisciplinär testning och buggloggning.
- Användbarhetstestning (n= \sim 10-20 deltagare) med prototyp och efterföljande versioner av appen.
- Att samla iterativ feedback från användare är bra för den fortsatta utvecklingen och framtida versioner av appen.

Förväntade resultat: Optimerad chattbot-prototyp.

PROVKÖRNING

Ett bra sätt att mäta engagemanget är att testa appen genom att provköra den i testgrupper.

Dessutom är det användbart att genomföra någon studie som en pilotstudie, en genomförbarhetsstudie eller en randomiserad studie med slutanvändare. I den här verktygssatsen hänvisar vi till testgrupper där chattboten laddas ner organiskt från appbutiker och används i den verkliga världen. Testanvändarna ger samtycke på distans att tillåta forskare att analysera chattbot-loggdata och alla andra interaktionsdata som samlas in via appen. I dessa fall är de insamlade uppgifterna mer naturligt förekommande, men kanske inte är lika rigorösa som i en mer kontrollerad studie.

Däremot ser vi den här typen av utprovning som mer kontrollerad när det gäller rekrytering och datainsamling, och att deltagarna kanske har mer kontakt med forskaren. Utifrån våra erfarenheter rekommenderar vi att man gör ett kort test först för att identifiera och lösa tekniska eller logistiska problem och problem med datainsamling.

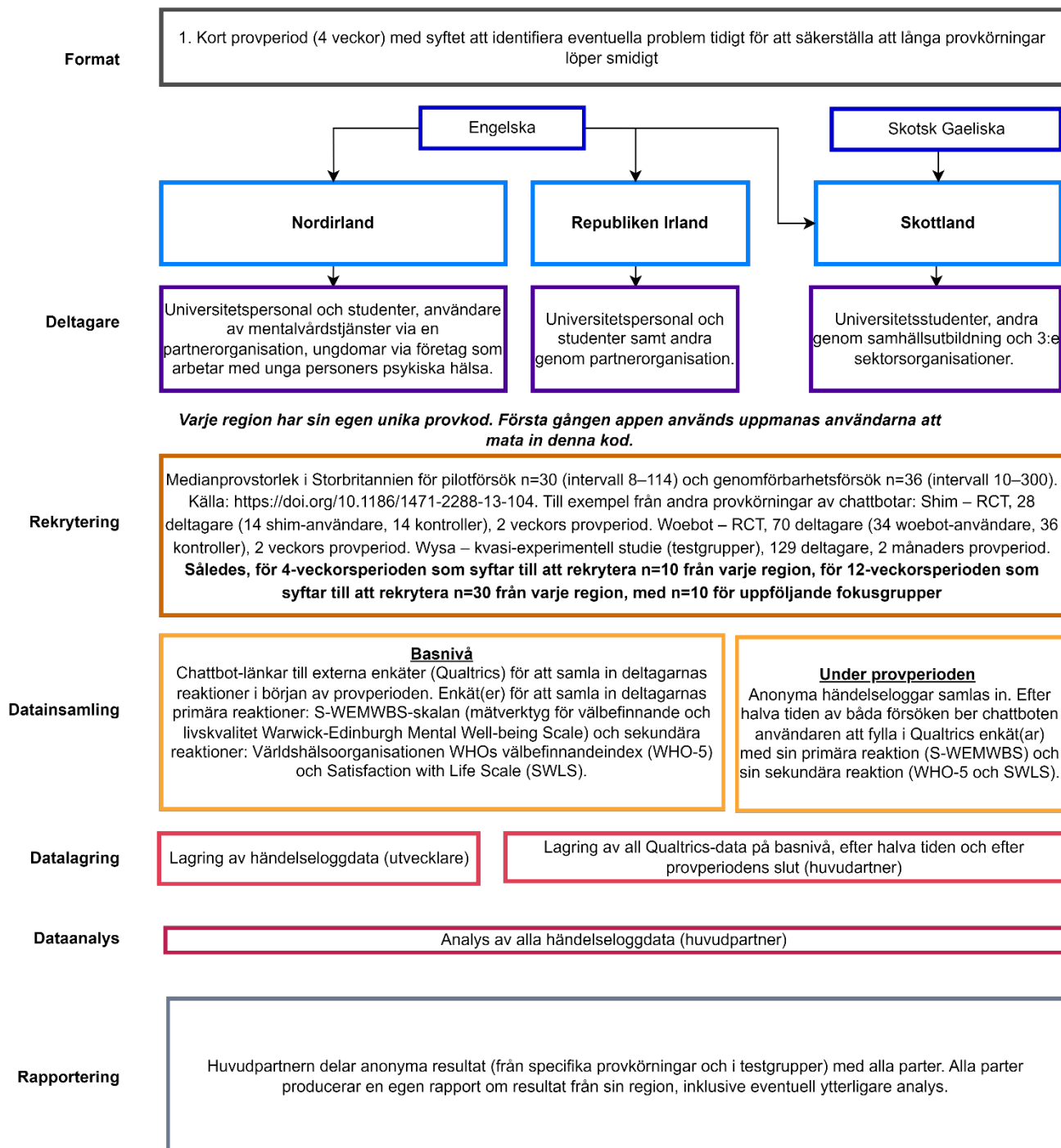
Vi valde först en kort 4-veckors testperiod och finjusterade sedan protokollet inför en längre 12-veckorsperiod. Under dessa provkörningar kan frågeformulär och skalor användas för att korsundersöka eller mäta olika resultat med jämna mellanrum under provkörningsperioden. När man väljer sätt att mäta resultaten bör man tänka på orsaken och effekten av den digitala lösningen. Man kanske också vill mäta på flera olika sätt.

Överväg att använda olika metoder för att samla in feedback. Man kan uppmana användaren att ge feedback i själva appen eller att fylla i externa undersökningar, följa upp deltagare via telefon, e-post eller personligen och fokusgrupper för att samla in viktig feedback om användarupplevelsen. På följande sidor kan du se den testdesign som användes för att testa ChatPal-appen.

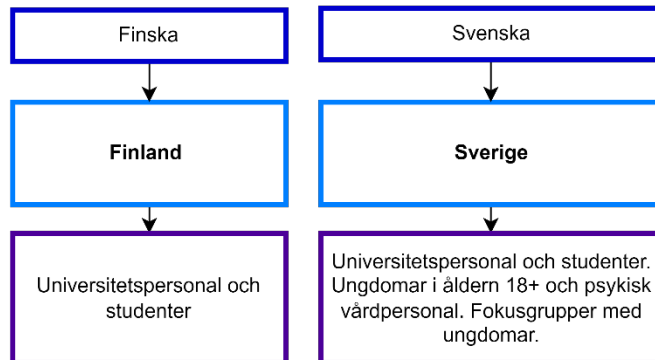
Forskarkoordinerade regionala provkörningar

App tillgänglig på mobila enheter (Android & Apple) och på nätet på fyra olika språk.

Varje appanvändare får ett unikt deltagar-ID som automatiskt läggs till i undersökningar så att det kan länkas till händelseloggdata.



2. Lång provperiod (12 veckor) för att mäta engagemang, viljan att använda appen och chattbotens effektivitet för att främja ett positivt mentalt välbefinnande



Inriktning på olika kohorter av elever och grupper (t.ex. HBTQ, funktionsnedsättning). Marknadsföring genom sociala medier, reklam genom lokala organisationer, flygblad och affischer, uppmuntra nya universitetsstudenter att ladda ner som en del av introduktionen osv.

Efter provperioden

Chattbot-länkar till externa enkäter (Qualtrics) för att samla in resultaten av provkörningarna. Enkät(er) som samlar in deltagarnas primära reaktioner (S-WEMWBS) och sekundära (WHO-5, SWLS) och ett frågeformulär om chattbotens användbarhet. Alla partner kan sätta ihop fokusgrupper efter provperioden för att samla in ytterligare feedback efter 12 veckor.

Andra typer av data t.ex. intervju/fokusgruppsdata som samlas in av specifika regioner lagras lokalt i varje land.

Korta och långa provkörningsrapporter ska skrivas i samma stil som en forskningsrapport (introduktion, metoder, resultat, diskussion, slutsats). Inklusivt kvantitativ (t.ex. användningsstatistik) och kvalitativ dataanalys (t.ex. tematisk analys från fokusgrupper/intervjuer, analys av fritext skriven av användarna i chattboten). Identifiera gemensamma teman för varje region och jämför resultat inom de olika områdena.

Provkörning med testgrupper

Appen tillgänglig i alla regioner på engelska, skotsk gaeliska, finska och svenska. Marknadsföring i sociala medier och flygblad som ska visas i butiker och företag för att locka till nedladdningar.

Alla som fritt laddar ner appen och ger sitt samtycke till inspelning och analys av deras anonyma uppgifter.

Användarna i testgrupperna ska välja "inte tillämpligt" när de uppmanas att ange en kod i början av provkörningen.

Inga uppgifter samlas specifikt in före eller efter provkörningen.

Under provperioden: Anonyma händelseloggar

PROVKÖRNING (FORTSÄTTNING)

Informerat samtycke bör erhållas från deltagare som anmäler sig till provkörningen. Det enklaste sättet att göra detta är att införliva digitalt samtycke i appen, med korta koncisa påståenden som anger vad provkörningen innebär och hur deras uppgifter kommer att användas. Deltagarna ska aktivt välja att samtycka till lagring och analys av deras uppgifter, eftersom detta är god etisk sed.

När det gäller rekrytering av deltagare kan man söka via till exempel universitet, företag som arbetar med psykisk hälsa eller andra hälsovårdstjänster och organisationer från tredje sektorn. Snöbollsmetoder är användbara för att uppmuntra andra att gå med i provkörningen, eftersom deltagarna kan dela detaljer med vänner, familj och kollegor. Publicering på nätet i nyhetsartiklar och annonsering i sociala medier kan uppmuntra deltagande och traditionella marknadsföringsmetoder, som att sätta upp affischer i lokala butiker och samhällslokaler, är ett användbart sätt att rikta in sig på icke-digitala personer.

Ett tillvägagångssätt som fungerar bra är att engagera deltagarna personligen för att bjuda in dem att delta i provkörningar och att följa upp personligen där det är möjligt. Detta kommer sannolikt att uppmuntra folk att fortsätta och kan vara användbart för att samla in feedback om produkten. Se till att deltagarna har tillräckligt stöd och tillgång till sjukvårdspersonal under hela provperioden om några problem skulle uppstå. Deltagare som rekryterades till ChatPal-provkörningarna fick en regionspecifik provkod så att antalet deltagare i varje område kunde spåras och registreras. Chattboten skickade ut en länk till externa undersökningar så att de kunde fylla i enkäter och förmedla sina primära och sekundära reaktioner och ge ytterligare feedback om appen. Alla appanvändare får ett personligt ID, som är anonymt, och som används för att länka chattbotens händelseloggdata till undersökningsdata för analys. Detta innebar att resultatdata (t.ex. olika mått på psykiskt välbefinnande) som samlades in i undersökningarna kunde kopplas till användningsdata om hur dessa

användare faktiskt använde chattboten under provperioden (registrering via händelseloggdata). Detta gjorde det möjligt för forskarna att associera förändringar i resultat med appanvändningen, eftersom resultaten möjligen kan förbättras även om chattboten inte användes (av andra skäl/variabler).

Avancerad analys utöver enkel statistik kan appliceras på händelseloggdata för att få nya insikter, till exempel att tillämpa oövervakad maskininlärning (klassificering) för att identifiera olika grupper av appanvändare baserat på användningsstatistik.

Skanna QR-koden nedan om du vill veta mer om provkörningarna av ChatPal:



Video där professor Catrine Kostenius och Frida Lindström (Luleå tekniska universitet) får sällskap av Courtney Potts (Ulster University), för att diskutera de provkörningar som har genomförts för ChatPal-appen.

Våra rekommendationer

- Planera att testa chattboten i en pilotstudie, en genomförbarhetsstudie eller en randomiserad studie med slutanvändare.
- Använd onlineannonsering och traditionella marknadsföringsmetoder för att rekrytera deltagare, inriktade på både digitala och icke-digitala personer.

Tips för en lyckad kampanj i sociala medier

- Ägna minst 3–4 månader åt att bygga upp målgrupper innan kampanjen lanseras.
- I allmänhet är videokampanjer mer engagerande än statiska bilder och lockar en större publik.
- Människor kan engagera sig mer i kampanjer om välrenommerade organisationer/institutioner som har varit involverade i design- och utvecklingsprocessen nämns i annonserna (till exempel hälso- och sjukvårdsorganisationer och universitet).
- Se till att introduktionsprocessen är så enkel som möjligt. Utforma kampanjerna på ett sätt som innebär att deltagarna bara behöver klicka en gång för att ladda ner appen och börja provköra den.
- Tillåt deltagarna att aktivt välja alternativ för lagring och analys av deras uppgifter där det är möjligt.
- Genomför ett kort test för att identifiera och lösa eventuella problem för att säkerställa att den längre provperioden löper smidigt.
- Överväg att använda flera skalor för att korsundersöka resultat, eller för att mäta olika utfall.
- Använd skalor som är validerade, robusta, lätta att använda, lämpliga för flera användningsområden och som kan översättas.
- Analysera händelseloggdata för att mäta hur många som har börjat använda appen.

- Rapportera om appens förmåga att hålla kvar användarna och hur folk gör när de använder appen, och analysera sambandet mellan chattbotanvändning och resultat.

Förväntade resultat: Provkörningsplan, marknadsföringskampanj för att rekrytera deltagare, användarhändelseloggdata för analys, dokument eller rapport som beskriver resultat av provkörningen inklusive eventuella förändringar i resultat.

SPRIDNING AV CHATTROBOTTJÄNSTER

Det sista steget är att sprida chattbottjänsten. Förbättringar kan göras av den slutliga chattboten baserat på feedback från testdeltagare. Det kan vara värt att överväga att sätta ihop en kommitté som träffas regelbundet för att diskutera underhållet i chattboten. Resultaten från provkörningarna bör offentliggöras för att visa eventuella förbättringar/förändringar i resultatet till följd av användning av en chattbot. Ett viktigt långsiktigt övervägande är slutanvändarnas förtroende för digital teknik, eftersom detta kommer att påverka användningen och adoptionen av chattboten. Liksom inför provkörningarna rekommenderar vi ett kort och koncist uttalande om hur slutanvändardata används, lagras och analyseras med en länk till integritetspolicyn. Standardalternativet bör vara "tacka nej", så att användarna aktivt får välja om de vill delta eller inte. Riktlinjer för att informera domänexperter om hur man rekommenderar eller förskriver en chattbot bidrar till att skapa förtroende hos vårdpersonalen och öka deras kompetens, vilket påverkar deras inställning till att adoptera och förskriva nya innovativa digitala hälsotekniker. Se här till exempel

<https://chatpal.interreg-npa.eu/general-information-on-chatpal-faqs/>. I Storbritannien har organisationen Organization for the Review of Health and Care Apps (ORCHA) som bedömer och ackrediterar digital hälsoteknik publicerat en rapport om Storbritanniens attityder och beteende med titeln "The People's View of Digital Mental Health"

<https://info.orchhealth.com/digital-for-mental-health-attitudes-and-behaviour-report>

I den här rapporten beskrev ORCHA resultaten av en onlineundersökning med 2 000 invånare i Storbritannien, där man fann att över en tredjedel (36 %) av de vuxna gärna skulle vilja få en rekommendation på en app för psykisk hälsa från sin läkare. Den siffran steg till 45 % för personer i åldern 35–44 och 55 % för 18–24-åringar.

Därför är det viktigt att ha stöd och backning från vårdgivare. Genom att utbilda och hjälpa personalen att acceptera digital omvandling kan de förstå tjänsten så att de kan rekommendera appen, vilket kan öka sannolikheten för att folk använder appen på längre sikt.

Varumärkesanpassning kan påverka förtroendet för digital teknik, eftersom folk är mer benägna att lita på en app som är godkänd av ett hälsovårdsorgan, som National Health Service (NHS) i Storbritannien, snarare än en kommersiell organisation. Till exempel tillhandahåller ORCHA en rigorös bedömningsprocess som gör det möjligt för den brittiska sjukvården att välja kvalitetssäkrade appar. ORCHA utvärderar appar, poängsätter dem baserat på data och integritet, säkerhet och användbarhet/tillgänglighet. Om appen når ett riktmärke (med 60 % eller mer) finns den med i ORCHAs digitala appbibliotek som vårdgivare kan använda för att rekommendera appar till kunder.

Våra rekommendationer

- Gör förbättringar baserat på användarfeedback innan ni släpper den slutliga chattbottjänsten.
- Fundera på hur chattboten kan integreras i nuvarande tjänster och hur dataflödet från den kan informera och förbättra dessa tjänster (t.ex. kanske humördata och användningsdata från chattboten kan delas i förväg med en terapeut för att bättre informera och förbättra kvaliteten på terapisessionerna).
- Sätt ihop en kommitté som ska underhålla chattboten på lång sikt.
- Vidta åtgärder för att säkerställa att vårdpersonal och användare litar på chattbottjänsten.
 - Offentliggör resultaten från provkörningarna och visa på eventuella förändringar i resultatet till följd av användning av en chattbot.
 - Arbeta med vårdpersonal för att tydligt förklara tjänsteutbudet så att de litar på produkten och kan rekommendera tjänsten.

- Kör appen genom en bedömnings-/ackrediteringsprocess så att hälso- och sjukvårdspersonalen kan vara säkra på den teknik de rekommenderar.
- Framhåll stöd från hälso- och sjukvårdsorgan om tillämpligt.

Förväntade resultat: En slutlig version av en hälsovårdsschattbot, evidensbas för att underlätta för vårdpersonal och klienter/slutanvändare att lita på chattboten.

Referenser

Baumel, A., Faber, F., Kane, J. and Muench, F. (2016) Enlight Quality Assessment Section. *Enlight-Quality-Assessment-Section.pdf* (mindtools.io)

Bond, R., Moorhead, A., Mulvenna, M., et al. (2019) Exploring temporal behaviour of app users completing ecological momentary assessments using mental health scales and mood logs. *Behav Inf Technol* 38:1016–1027. <https://doi.org/10.1080/0144929x.2019.1648553>

Bickmore, T., Trinh, H., Olafsson, S., et al. (2018) Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. *J Med Internet Res* 20(9):e11510. <https://doi.org/10.2196/11510>

Cameron, G., Cameron, D., Megaw, G., et al. (2018a) Best practices for designing chatbots in mental healthcare: a case study on iHelp. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI '18)*. BCS Learning & Development Ltd., Swindon, GBR, Article 129, 1–5. <https://doi.org/10.14236/ewic/HCI2018.129>

Cameron, G., Cameron, D., Megaw, G., et al. (2018b) Back to the future: lessons from knowledge engineering methodologies for chatbot design and development. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI '18)*. BCS Learning & Development Ltd., Swindon, GBR, Article 153, 1–5. <https://doi.org/10.14236/ewic/HCI2018.153>

Deibel, D. and Evanhoe, R., 2021. *Conversations with Things: UX design for Chat and Voice*. Rosenfeld Media.
Goldstein, S.P., Evans, B.C., Flack, D. et al. (2017) Return of the JITAI: Applying a Just-in-Time Adaptive Intervention Framework to the Development of m-Health Solutions for Addictive Behaviors. *Int J Behav Med* 24:673–682. <https://doi.org/10.1007/s12529-016-9627-y>

Mori, M., MacDorman, K.F., and Kageki, N. (2012) The Uncanny Valley [From the Field]. *IEEE Robot Autom Mag.* 19(2):98-100. <https://doi.org/10.1109/MRA.2012.2192811>.

Potts, C., Ennis, E., Bond, R.B. et al. (2021) Chatbots to Support Mental Wellbeing of People Living in Rural Areas: Can User Groups Contribute to Co-design?. *J Technol Behav Sci* 6:652–665. <https://doi.org/10.1007/s41347-021-00222-6>

Stoyanov, S., Hides, L., Kavanagh, D. and Wilson, H. (2016) Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth and uHealth* 4(2):e72. <https://doi.org/10.2196/mhealth.5849>

Sweeney, C., Potts, C., Ennis, E., et al. (2021) Can Chatbots Help Support a Person's Mental Health? Perceptions and Views from Mental Healthcare Professionals and Experts. *ACM Trans Comput Healthc* 2(3):1-16. <https://doi.org/10.1145/3453175>

Checklista inför utformningen av en chattbot för hälsovård

Intressentcentrerad design

Förståelse av hälso- och sjukvårdsdomänen

- Gör en litteraturgenomgång.
- Planera för kunskapsdelningsdiskussioner eller diskussionsgrupper med vårdpersonal och teknikexperter.
- Identifiera användningsfall som godkänts av sjukvårdspersonal.
- Bestäm chattbotens omfattning och syfte.

Tekniska begränsningar

- Bedöm styrkor och svagheter hos den teknik som är tänkt att användas.
- Titta på ramar för utveckling av öppen källkod.
- Överväg att välja en utvecklingsmiljö där åtkomsten till alla uppgifter är privat och fristående.
- Tänk på att implementera nyckelordsutlösare.
- Planera chattbotens svar på oväntade kommentarer/gränsfall.

Behöver analyseras

- Identifiera berörda intressenter.
- Samla in användarbehov genom att arbeta med slutanvändare och intressenter (t.ex. undersökningar eller diskussionsgrupper).

Definiera krav

- Samproducera lösningar för att möta användarnas behov och användningsfall som stöds av yrkesverksamma, inom vad som är tekniskt genomförbart.
- Prioritera och förfina kravlistan.
- Bestäm slutliga chattbot-funktioner baserat på kraven.

Utformning av dialoger

- Fatta de slutliga besluten när det gäller chattbotens modalitet och användning av multimedia eller röst, typ av interaktion, personifiering och extern kunskap.
- Skapa samtalsmanus baserade på de slutliga funktionerna i chattboten tillsammans med experter på samtalsdesign och utifrån bästa praxis.
- Se till att användarna blir medvetna om chattbotens begränsningar i början av samtalet.
- Tänk på hur användarinmatningar lagras och framställs av chattboten.
- Översätt innehållet till andra språk (om det är aktuellt).
- Bedöm och signera manus baserat på samtalsflöde, informationskvalitet, kulturella aspekter och persona (se fullständig checklista i verktygssatsen ovan).

Utveckling av chattboten

- Ställ in anonym inspelning av varje interaktion i appen (användarhändelseloggdata) för framtida analys.
- Överväg att använda s.k. ecological momentary assessment (EMA), dvs. frågor som presenteras och besvaras av användare "i nuet".
- Tänk på att lägga till multimediamaterial och globala aviseringar för att påminna användarna om att engagera sig i appen.

Testning av chattboten

- Genomför omfattande provkörningar av chattboten, använd till exempel ett tvärvetenskapligt team som rapporterar eventuella problem som utvecklingsteamet behöver åtgärda.
- Genomför användbarhetstester med en liten grupp slutanvändare (10–20 deltagare).
- Använd feedback från användare för att göra förbättringar av efterföljande versioner av appen.

Provkörning

- Planera att provköra chattboten med hjälp av slutanvändare (i en pilotstudie, en genomförbarhetsstudie eller en randomiserad studie).
- Rekrytera testdeltagare med hjälp av traditionell marknadsföring (affischer, flygblad osv.) och via onlineannonsering (se listan med "tips för en lyckad kampanj i sociala medier" i verktygssatsen).
- Som standard, se till att användarna aktivt ska välja att samtycka till lagring och analys av deras uppgifter.
- Överväg ett kort provperiod med ett litet antal deltagare för att identifiera och lösa eventuella problem inför en längre provperiod.
- Bestäm hur resultaten ska mätas och lämpliga skalor att använda med upprepade intervaller under försöken.
- Analysera användarhändelseloggdata från provkörningarna för att mäta hur många som har börjat använda appen och hur många som stannar kvar, och utforska sambandet mellan chattbotanvändning och resultat.

Spridning av chattbottjänster

- Gör förbättringar baserat på användarfeedback innan ni släpper den slutliga chattbottjänsten.
- Offentliggör resultaten från provkörningarna och visa på eventuella förändringar i resultatet till följd av användning av chattboten.
- Fastställ hur chattboten kan integreras och förbättra dagens hälsotjänster genom att samarbeta med vårdpersonal.
- Sätt ihop en kommitté som ska diskutera det långsiktiga underhållet av chattboten.
- Vidta åtgärder för att säkerställa att vårdpersonal och användare litar på chattbottjänsten (låt ORCHA eller liknande organisationer granska och ackreditera appen).

www.chatpal.interreg-npa.eu

