

## Hur AI-verktyget ChatGPT klarar en hemtentamen i palliativ vård

Mats Christiansen<sup>1\*</sup>, Lena Normark<sup>2</sup> & Christine Leo Swenne<sup>2</sup>

<sup>1</sup>Uppsala universitetet, Sverige, Åbo Akademi, Finland; <sup>2</sup>Uppsala universitet, Sverige

Dataprogrammet ChatGPT är en botten som utför automatiska uppgifter. Denna chattbotten skapade rubriker under hösten på grund av dess förmåga att med artificiell intelligens (AI) skapa svar på riktade frågor och texter utifrån ett beskrivet syfte. I slutet av december 2022 lades 16 frågor från en hemtentamen i palliativ vård vid sjuksköterskeprogrammet in i AI-botten för att få dem besvarade och testa botten förmåga att besvara frågorna korrekt. Vi ville även testa om man kunde se någon skillnad på svaren i rättningsmallen och på studenternas svar och på de svar som genererats från AI-botten. En bedömning av botten svar visade på att den hade klarat examinationen. Somliga svar var mycket bra och resonerande till sin natur, något gav reducerade poäng pga. bristande innehåll och ett par svar felaktiga. Ett svar stack ut då det var på engelska. Svaren hade ibland smärre grammatiska fel, men var främst betydligt längre och bättre strukturerade än studerandes. Möjligheten att upptäcka försök till vilseledande i examination hade varit liten. Det finns skillnad i svarens diskurs, där svaren från botten har en mer amerikansk stil i sin formulering.

*Nyckelord:* sjuksköterskeutbildning, examination, bedömning

### How the AI tool ChatGPT passes a home exam in palliative care

The ChatGPT computer program is a bot that performs automatic tasks. This chatbot made headlines during the fall due to its ability to use Artificial Intelligence (AI) to create answers to targeted questions and texts based on a described purpose. At the end of December 2022, 16 questions from a home exam in palliative care at the nursing program were put into the AI bot to get them answered and test its ability to answer the questions correctly but also an assessment of whether the answers could be detected. An evaluation of the bot's answers showed it had passed the examination. Some answers were excellent and reasoned, some gave reduced points due to lack of content, and some were incorrect. One answer stood out as it was in English. The answers sometimes had minor grammatical errors but were significantly more prolonged and better structured than the students' answers. The possibility of detecting attempts at deception in the examination had been limited. There is a difference in discourse in the answers, with the answers from the bot having a more American style.

*Keywords:* nursing education, examination, assessment

---

\*Författarkontakt: Mats Christiansen, e-post: mats.christiansen@pubcare.uu.se

Artiklar och reflektioner är kollegialt granskade. Övriga bidragstyper granskas av redaktionen. Se <https://hogreutbildning.se> ISSN 2000-7558

©2023 Mats Christiansen, Lena Normark & Christine Leo Swenne. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to share their work (copy, distribute, transmit) and to adapt it, under the condition that the authors are given credit, that the work is not used for commercial purposes, and that in the event of reuse or distribution, the terms of this license are made clear.

Citation: Christiansen, M., Normark, L. & Swenne, C. L. (2023). «Hur AI-verktyget ChatGPT klarar en hemtentamen i palliativ vård», *Högre utbildning*, 13(2), 56–62. <https://doi.org/10.23865/hu.v13.5331>

## INTRODUKTION

Under senhösten 2022 presenterades ChatGPT (chat.openai.com), vilket är en chattbott som med hjälp av artificiell intelligens (AI) besvarar frågor som ställs till densamma (Open AI, 2022). Svaren kan förfinas och korrigeras i botten genom att påpeka felaktigheter i svaret eller att ställa fördjupande frågor där dessa kan vara med för att ge ett mer fullödigt svar.

Redan samma dag som denna chattbott släpptes lyftes faran för skriftliga examinationer fram (Bryan, 2022). I en tråd med twittermeddelanden förklarar Bryan hur denna bott skulle kunna skriva studenters uppgifter och uppsatser. Även svenska lärosäten och deras företrädare har i olika sammanhang under vintern lyft riskerna med denna chattbott (exv. Ahlgren & Leitet, 2022; Färlin, 2023; Institutionen för informationsteknologi UU, 2022). Både svensk och internationell nyhetsmedia har visat på dess utmaningar, där fokus legat på tilltro till svar som ges, men även lyft fram risken med inkorrekt information online som denna AI-bott kan lära sig av (t.ex. Brännström, 2022; Larsson, 2022; Metz, 2022). Det har även rapporterats om felaktigheter i fakta som den rapporterar, där ett exempel är att senaste svenska riksdagsvalet ägde rum 2021 och där botten inte korrigerar sitt svar (Axén, 2022).

Denna bott är numera en realitet som lärare i utbildningar på både grund-, gymnasie- och högskolenivå behöver förhålla sig till. Den kommer med all säkerhet att användas av studenter för att snabbt få svar på frågor, där den kan ha en spännvidd från uppslagsverk till spökskrivare. I framtiden kommer botten även att kunna användas som en resurs till studenter i deras individuella studier, exempelvis som läxhjälp. Verktyg börjar presenteras för att även bidra till att bedöma om den har använts för att skriva svaren (exv. Kim, 2022).

Syftet med denna studie var att undersöka om ChatGPT skulle kunna besvara frågorna i en av våra examinationer vid sjuksköterskeprogrammet vid Uppsala universitet, samt att granska det sätt som ChatGPT besvarat frågorna på och undersöka om bedömande lärare hade kunnat upptäcka att det var en bott som besvarat dem.

## METOD

### *Beskrivning av hemtentamen*

Som en del av examinationerna i termin 6 på Sjuksköterskeprogrammet (grundnivå) förekommer en skriftlig, individuell hemtentamen i palliativ vård. Examinationen byggde vid detta tillfälle på ett patientfall som utgick från en yngre man med en hjärntumör, där patientens sjukdomstillstånd försämras successivt, vilket beskrivs i scenariot och där frågorna utgår från patientens progredierande sjukdomsbild. I patientfallet inkluderas kunskaper om palliativ vård och studentens förmåga att bedöma symptom och tecken på ohälsa samt behov hos närstående. Studenten ska även kunna bedöma behov av behandling och åtgärder, utvärdera behandling och åtgärder samt ha kunskap om vårdsystem. Frågorna som ställdes innehöll fakta, svaren innebar att studenterna behövde utveckla och reflektera, dvs. hemtentamen omfattade taxonomierna Structure of Observed Learning Outcome (SOLO) och Ideas, Connections and Extensions (ICE) (Biggs & Tang, 2011; Örebro universitet, 2020). Studenternas svar ska enligt studieuppgiften ha en grund i och referenser till kurslitteraturen, vetenskapliga artiklar och föreläsningar som ingår i kursen för att kunna bli godkända.

De poängsatta frågorna besvaras via ett digitalt examinationsverktyg Inspera Assessment (<https://www.inspera.com/>), vilket används vid lärosätet. Svaren kontrolleras för plagiat i verktyget. I höstterminens examination förekom 16 frågor vilket maximalt kunde ge 36 poäng. För godkänt resultat krävdes 25 poäng (67%). Vid denna examination hade 93 studenter lämnat in

hemtentamen och en majoritet klarade examinationen. Frågorna rättades utifrån en rättningsmall som är utarbetad av ansvariga lärare för examinationen.

#### *Svar från AI-verktyget ChatGPT*

Samtliga 16 tentamensfrågor ställdes till ChatGPT den 28 december 2022 efter att examinationen avslutats bland studenterna i syfte att undersöka om ChatGPT kunde besvara frågorna. Den första botten svar på frågorna sparades och granskades. Därefter gjordes en bedömning utifrån examinationens rättningsmall. Vid granskningen av svaren bedömdes om svaren var tillräckligt utförliga för att en student enbart genom att kopiera svaret från botten hade kunnat bli godkänd på hemtentamen. Vidare gjordes en bedömning huruvida det givna svaret hade kunnat upptäckas som skrivet av en bott. Genom dialog med botten hade svaren troligen kunnat förfinas och några svagheter hade eventuellt även korrigerats, men i detta test av ChatGPT avstod vi ifrån detta.

#### RESULTATET

Resultatet visade att tre svar hade fått noll poäng då ett svar besvarades på engelska samt två på grund av att svaret speglade ett annat, felaktigt sammanhang. I övrigt var botten svar ungefär 2–3 gånger så omfattande i jämförelse med studenternas svar på hemtentamen i examinationsverktyget Inpera Assessment. Texten från ChatGPT på svenska innehöll i genomsnitt 175–180 ord. Ett exempel på ett bra svar avsåg ett scenario med sambons undran om samliv och sexualitet i samband med palliativ vård: *Ge förslag på information och åtgärder du kan göra som sjuksköterska*. Botten svar hade renderat 1,5 av 2 p:

Som sjuksköterska kan du ge Lisa och Kalle information om att intimitet och sexualitet är viktiga aspekter i alla stadier av livet, inklusive i palliativt skede. Du kan också ge information om hur de kan fortsätta att upprätthålla intimitet och sexualitet under dessa omständigheter, till exempel genom att prata om deras behov och önskemål med varandra, att använda kommunikation och beröring för att skapa närhet, att använda hjälpmedel som kan underlätta intimiteten (till exempel stödprodukter för sängen), och att ta hjälp av professionell hjälp (till exempel en sexterapeut).

Det är också viktigt att du ger Lisa och Kalle tid och rum att prata om och hantera de känslor och bekymmer som kan uppstå när man saknar intimitet och sexualitet. Du kan också erbjuda dem stöd och rådgivning om hur man kan hantera dessa känslor och bekymmer, och hjälpa dem att hitta strategier för att hantera eventuella hinder som kan förhindra intimitet och sexualitet.

Det är viktigt att du också respekterar Lisa och Kalles integritet och privatafär, och att du ger dem den tid och den plats de behöver för att utforska sin intimitet och sexualitet på ett självbestämmande sätt. Du bör också se till att du har den kompetens och den kunskap som krävs för att ge rådgivning och stöd om dessa ämnen på ett professionellt och empatiskt sätt.

Ser vi till svarets diskursiva stil finns en annan ton än vad som förekommer i studenternas svar. Det är en diskurs som är mer ödmjuk och känslig.

Ett exempel på ett felaktigt svar var svaret på frågan: *Kalle har blivit så törstig [på föregående sida har det beskrivits att patienten står på kortisonbehandling]. Under ditt besök dricker han flera glas vatten på en kort stund. Lisa berättar att han kissar ofta. Vad misstänker du nu? Motivera.*

Based on the information provided, it is possible that Kalle may be experiencing dehydration. Dehydration occurs when the body does not have enough fluids, and can be caused by a variety of factors such as excessive sweating, vomiting, diarrhea, or not drinking enough fluids.

One sign of dehydration is increased thirst, which Kalle is experiencing. Another sign is increased urination, which Lisa mentions that Kalle is also experiencing. It is important to properly hydrate the body to maintain good health, and if Kalle is experiencing dehydration, it may be necessary for him to drink more fluids and possibly seek medical attention if his symptoms are severe or do not improve.

Bedömning: 0 av 2p. Korrekt svar: misstanke om diabetes då patienten står på en behandling med kortison. Diabetes är en komplikation till kortisonbehandling som en snart färdig sjuksköterska förväntas känna till och reagera på. Dessutom var detta svar skrivet på engelska av botten, trots att frågan ställts på svenska.

Vidare hade vi exempel på svar där upprepade grammatiska fel ("en dödsfall") hade varit en möjlig grund till att en bedömande lärare kunde ha reagerat, förutom förekommande faktafel.

I Sverige finns det olika typer av efterlevandestöd som kan erbjudas till den som har förlorat en närstående. Detta syftar till att hjälpa den som sörjer att hantera sin sorg och sin nya situation, samt att ge rådgivning och stöd i olika frågor som kan uppstå efter en dödsfall.

Efterlevandestöd kan till exempel bestå av olika typer av samtalsstöd, gruppstöd, rådgivning och stöd i kontakt med olika myndigheter och försäkringsbolag. Det kan också innefatta stöd i form av praktisk hjälp, till exempel med att ordna begravningen eller att hantera den avlidnes ekonomi.

Det finns olika organisationer och myndigheter som erbjuder efterlevandestöd, till exempel Socialstyrelsen, Försäkringskassan, Försäkringskassans kris- och efterlevandekontor, Röda Korsets efterlevandestöd och olika församlingar och föreningar. Det är också vanligt att kommuner och landsting har egna verksamheter som erbjuder efterlevandestöd. Det är viktigt att tänka på att det finns ingen "rätt" eller "fel" sätt att sörja på, och att alla reagerar olika på en dödsfall. Efterlevandestöd är till för att ge stöd och hjälp till den som sörjer, och det kan vara bra att ta kontakt med någon av ovan nämnda organisationer eller myndigheter om man känner att man behöver hjälp att hantera sin sorg eller om man har frågor eller funderingar som man vill få svar på.

Totalt genererade svaren från ChatGPT 28 poäng. Eftersom gränsen för godkänt är 25 poäng skulle de svar som genererats av ChatGPT ha godkänts under förutsättning att studenten lagt in eller bett om någon form av litteraturförankring i svaret. Dessa essäsvare var genomgående präglade av en inre logik, hade väl avvägda stycken och i vissa fall listade svar i punkter eller siffror.

## DISKUSSION

Det är tveksamt om de bedömande lärarna hade kunnat identifiera om svaren kom från AI-verktyget ChatGPT. Vid en examination med ett större antal individuella examinationer, och att bedömande lärare hade bedömt lika, hade svar med inre logik och bra innehåll kunnat ha passerat.

En granskning av de 16 svar på frågorna som vi samlade in från AI-verktyget ChatGPT visar att det förekommer ett svar på engelska trots att frågan var ställd på svenska. I detta fall var även svaret felaktigt innehållsligt. De övriga 15 svaren var på svenska. Det hade för bedömande lärare varit ett frågetecken med ett engelskspråkigt svar då inga övriga svar angavs på annat än

svenska. Botten hade vidare ett svar vilket avser en skild kontext (personalfrågor i stället för vård). Även detta svar hade identifierats i rättningen och poängsatts till noll poäng. I bägge dessa fall hade svaren stuckit ut, men frågan är om det hade varit nog för att bli misstänksam i en större omfattning.

Svaren har en del grammatiska fel såsom en/ett med följande inkongruenser samt avsaknad av bestämd form, men som är svar som inte nödvändigtvis hade varit något som en bedömande lärare hade reagerat på. Dels förekommer studenter som kan ha svenska som andraspråk, dels är det främst vanligt att det även i en datoriserad inlämning sker felskrivningar där tanken kanske gått snabbare än fingrarna över tangenterna.

Texten från ChatGPT på svenska innehöll i genomsnitt 175–180 ord, vilket i genomsnitt var ungefär två till tre gånger så lång text som studenter skrivit som besvarade frågorna i det digitala examinationsverktyget Inspira Assessment. Dessa essäsvare var genomgående präglade av en inre logik, de hade väl avvägda stycken och i vissa fall listade svar i punkter eller siffror. En sådan framställning är det enklare att duperas av. Som bedömande lärare är det ”drömsvaren” att få bedöma. Det går snabbt att få en överblick över svaret och att de delar eller argument som krävs kan återfinnas.

Diskursen i svaren som genererats av AI-verktyget ChatGPT är emellertid mer amerikansk i sin kommunikativa stil, snarare än svensk (Staples, 2015). Det finns en artighet och strävan efter att inte komma i konflikt med det svar som den anger avseende sexualitet och intimitet. Begrepp som respekt för integritet och privatsfär nämns, i studenternas svar var detta en mindre förekommande aspekt utan ett större fokus på att ge förslag till hjälp och assistans. Denna kulturella skillnad finns exempelvis i tonen på det bra svaret

*... Det är viktigt att du också respekterar Lisa och Kalles integritet och privatsfär, och att du ger dem den tid och den plats de behöver för att utforska sin intimitet och sexualitet på ett självbestämmande sätt.*

För botten kan man kanske fundera över om vad som är dess modersmål. Det finns litteratur som visat på att personer har en annan kommunikativ stil när de använder sitt andra språk jämfört med sitt modersmål (Han et al., 2020). Frågan är förstås om vi kan tänka oss att det även lyser igenom i en botts svar.

En begränsning är att vi inte gjorde något försök att granska botten svar i något av de verktyg som tagits fram för ändamålet. Det är i skrivande stund heller inget som förekommer i lärosätets digitala examinationsmiljö, vilket för närvarande enbart påpekar potentiellt plagiat.

### *Implikationer*

En bott likt denna klarar av att formulera längre, essäliknande svar. Tidigare har krav på resonerande svar varit en väg för att fånga upp försök till fusk då det krävt mer utvecklade och resonerande svar, men med en bott likt ChatGPT blir det svårt att använda det som kriterium.

Somliga svar är mycket välformulerade, välskrivna och de hade kunnat passera utan anmärkning, medan andra hade upptäckts på grund av faktafel. Botten skriver dock svar som genomgående är längre än vad som kunde förväntas i flera av svaren.

Diskussioner har börjat föras – även *med* ChatGPT-botten – om hur den skulle kunna användas i undervisningen i stället för att se den som ett hot eller problem (O'Connor & ChatGpt, 2023). Den kan användas som en resurs att ställa frågor till, om man håller i minnet att svaren kan vara felaktiga som vi sett ovan. Men felaktiga kan fungera som en utgångspunkt

för diskussioner där studenter i stället får korrigera svaren (O'Connor, 2021). Den föreslår själv att den kan bidra med faktakunskaper till fallstudier och skrivuppgifter (ChatGpt, 2023).

Vår slutsats blir att botten klarar att ge oss svar enligt flera olika nivåer i olika taxonomier, varför andra typer av examinationsformer kan komma att behövas. Om en student hade valt att kopiera svaren till en examination och i tillägg angivit referenser hade resultatet blivit acceptabelt. Även om denna fallstudie mest blir ett inlägg i den alarmistiska delen av att använda ChatGPT, öppnar ChatGPT-botten för nya pedagogiska möjligheter i undervisningen och som stöd i inläringen för studenterna i framtiden.

#### FÖRFATTARBIOGRAFI

**Mats Christiansen** arbetar som universitetsadjunkt vid sjuksköterskeutbildningarna, Institutionen för folkhälso- och vårdvetenskap, Uppsala universitet. Han är med.lic., specialistsjuksköterska inom vård av äldre och aukt. sexolog. Vidare är Mats doktorand i vårdvetenskap vid Fakulteten för pedagogik och välfärdsstudier, Åbo Akademi, Vasa i Finland.

**Lena Normark** är universitetsadjunkt med klinisk inriktning vid sjuksköterskeutbildningarna, Institutionen för folkhälso- och vårdvetenskap, Uppsala universitet. Hon är leg. sjuksköterska och med.mag. med fokus på pedagogik och handledning.

**Christine Leo Swenne** är programansvarig och studierektor för Sjuksköterskeprogrammet, Institutionen för folkhälso- och vårdvetenskap, Uppsala universitet. Leo Swenne är docent och handleder och examinerar uppsatser på grundutbildningsnivå och avancerad nivå, föreläser och håller seminarier i vetenskaplig metodik, operationssjukvård och har reflekterande yrkesmässig handledning med studenter i programmet.

#### REFERENSER

- Ahlgren, F. & Leitet, J. (2022, Dec 14). *Nytt AI-verktyg ändrar förutsättningarna för studier och examination*. Linnéuniversitetet. Hämtad January 3 från <https://lnu.se/mot-linneuniversitetet/aktuellt/nyheter/2022/nytt-ai-verktyg-paverkar-studier-och-examinationer/>
- Axén, G. [gunnar.axen]. (2022, December 27). *Som gammal early adopter (anammare av något nytt) så följer jag utvecklingen av AI. Därför testar jag AI-motorn GPT 3 och funderar på när landets politiska kommentatorer kan komma att rationaliseras bort...* Facebook. <https://www.facebook.com/gunnar.axen/posts/pfbid02szGNoxmYtjQ5TeqNZyn3jbdzwZYGUMPkBmLnACsrNvvpNBdw2tB5NWv4YB7p8vUPKI>
- Biggs, J. & Tang, C. (2011). *Teaching for quality learning at university : what the student does* (4. uppl.). Open University Press.
- Bryan, K. [Afinetheorem]. (2022, November 30). *I have helped run an AI-based entrepreneurship program for years, written papers on the econ of AI, and follow the field quite closely. Nonetheless, I am \*shocked\* by how good OpenAI's new chat (<https://chat.openai.com/chat>) is. E.g., you can no longer give take-home exams/homework*. Twitter. <https://twitter.com/Afinetheorem/status/1598081835736891393?s=20&t=BdJK3MpM7XWmxcfTejxSXg>
- Brännström, S. L. (2022, December 14). AI-revolutionen är här: "Lätt att bli förbluffad". *Svenska Dagbladet*. <https://www.svd.se/a/APb83q/open-ai-s-chat-gpt-valte-internet-wallenberg-satsar-pa-svenskt-program>
- ChatGpt. (2023). *Chat: ge mig tre exempel på hur ChatGPT kan vara en resurs inom sjuksköterskeutbildningen* <https://chat.openai.com/chat/94cc602f-b9d1-44eb-b7a8-69cc5e55c117>
- Färlin, J. (2023, Jan 7). Uppsatsskrivande bottar – det nya fusket? *Uppsala Nya Tidning*. <https://unt.se/nyheter/uppsala/artikel/universitetet-rustar-mot-ai-hotet-uppsatsskrivande-bottar-det-nya-fusket-prorektorn-studenterna-ligger-fore-oss/rmz5k911>



- Han, Y., Segalowitz, N., Khalil, L., Kehayia, E., Turner, C. & Gatbonton, E. (2020). Do nurses use discourse markers differently when using their second language as opposed to their first while interviewing patients? *The Canadian Modern Language Review*, 76(2), 91–113. <https://doi.org/10.3138/cmlr-2018-0268>
- Institutionen för informationsteknologi UU [ITinstitutionen]. (2022, December 22). *Thomas Schön is talking in SVT about AlphaCode and ChatGPT, two powerful AI tools. This is something everyone and especially us as researchers and educators in this topic has to relate and adopt to!* Facebook. [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiMjKO8j6z8AhUwx4sKHBroAdEQFn0ECAwQAQ&url=https%3A%2F%2Fwww.facebook.com%2FITinstitutionen%2F&usg=AOvVaw1lwPCPV-SdffPoL\\_rv-Qn-](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiMjKO8j6z8AhUwx4sKHBroAdEQFn0ECAwQAQ&url=https%3A%2F%2Fwww.facebook.com%2FITinstitutionen%2F&usg=AOvVaw1lwPCPV-SdffPoL_rv-Qn-)
- Kim, S. (2022). How to detect OpenAI's ChatGPT output. *Geek Culture*. <https://medium.com/geekculture/how-to-detect-if-an-essay-was-generated-by-openais-chatgpt-58bb8adc8461>
- Larsson, L. (2022, December 6). Chatroboten visar att vi inte kan lita på vem som är människa. *Dagens Nyheter*. <https://www.dn.se/kultur/linus-larsson-chattroboten-visar-att-vi-inte-kan-lita-pa-vem-som-ar-manniska/>
- Metz, C. (2022, December 12). The new chatbots could change the world. Can you trust them? *New York Times*, 1. <https://www.nytimes.com/2022/12/10/technology/ai-chat-bot-chatgpt.html>
- O'Connor, S. (2021). Artificial intelligence and predictive analytics in nursing education [Editorial]. *Nurse Educ Pract*, 56, 103224. <https://doi.org/10.1016/j.nepr.2021.103224>
- O'Connor, S. & ChatGpt. (2023). Open artificial intelligence platforms in nursing education: tools for academic progress or abuse? [Editorial]. *Nurse Educ Pract*, 66, 103537. <https://doi.org/10.1016/j.nepr.2022.103537>
- Open AI. (2022). ChatGPT: optimizing language models for dialogue. <https://openai.com/blog/chatgpt/>
- Staples, S. (2015). *The discourse of nurse-patient interactions: contrasting the communicative styles of U.S. and international nurses*. John Benjamins Publishing Company.
- Örebro universitet. (2020, 2021-06-07). *Lär dig mer om kunskapsnivåer*. <https://www.oru.se/om-universitetet/hogskolepedagogiskt-centrum/resurser/kursmal/lar-dig-mer-om-kunskapsnivaer/>