

Participatory Design with Domain Experts: A Delphi Study for a Career Support Chatbot

Marianne Wilson
Edinburgh Napier University
m.wilson2@napier.ac.uk

Dimitra Gkatzia
Edinburgh Napier University
d.gkatzia@napier.ac.uk

David Brazier
Edinburgh Napier University
d.brazier@napier.ac.uk

Peter Robertson
Edinburgh Napier University
p.robertson@napier.ac.uk

ABSTRACT

We present a study of collaboration with expert participants for the purpose of the responsible design of a conversational agent. The Delphi study was used to identify and develop design and evaluation criteria for an automated career support intervention. Career support tasks present complex design problems as they are highly personalized and the definition of success for a single intervention is ambiguous. The study engaged domain experts in a structured communication process to explore the opportunities and risks of introducing a conversational agent to complement existing services provided to young people. Three rounds of questionnaires were used to build consensus across the expert panel. The questionnaire design incorporated design fictions, qualitative data from the panel, and requirement statements. The study produced a validated set of criteria that can be used for the design and evaluation of a conversational agent, that aligns with professional ethics and intended outcomes for a career support intervention. Our approach demonstrates the value of mixed method Delphi studies to facilitate participatory design of conversational user experiences by bridging knowledge gaps between technical and domain experts. The resulting evaluation criteria establish a meaningful foundation for future human-centered conversation design for career support.

CCS CONCEPTS

• **Discourse, dialogue and pragmatics;** • **Participatory Design;**
• **Retrieval tasks and goals;**

KEYWORDS

Conversational Agents, Career Support, Delphi Study, Participatory Design, Domain Experts, Responsible AI, Ethics, Evaluation

ACM Reference Format:

Marianne Wilson, David Brazier, Dimitra Gkatzia, and Peter Robertson. 2024. Participatory Design with Domain Experts: A Delphi Study for a Career Support Chatbot. In *ACM Conversational User Interfaces 2024 (CUI '24)*, July 08–10, 2024, Luxembourg, Luxembourg. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3640794.3665534>



This work is licensed under a Creative Commons Attribution International 4.0 License.

CUI '24, July 08–10, 2024, Luxembourg, Luxembourg
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0511-3/24/07
<https://doi.org/10.1145/3640794.3665534>

1 INTRODUCTION

Career education, information, advice and guidance (CEIAG) is an umbrella term for interventions to support the career choices and development of young people and adults. Career related services are often provided by educational institutions or state agencies. The social and economic value of investing in CEIAG is recognised by intergovernmental organisations [8]. Digital technologies have been deployed by career services since the 1970s, and while there are a variety of established applications [15], the use of conversational agents designed specifically for this domain is limited. Only one case study of a conversational agent is included by the Observatory on Digital technologies in Career guidance for Youth at the time of writing [27]. Career decision-making is a complex, personal and porous process. The way in which individuals navigate their careers is influenced by, and has impacts on, personal, social and economic factors and outcomes. As such, responsible design of an automated intervention requires thorough consideration of the context and objectives of the intervention, in collaboration with domain experts.

We report the design and results of a Delphi study, conducted in collaboration with a publicly funded national career support service. The purpose of the collaboration is to support the development of a conversational agent that can supplement their existing services for young people. Delphi studies are an established method for structured communication used in participatory research to build consensus between a panel of experts [23]. It entails multiple rounds of questionnaires distributed to a panel of participants, with each round presenting and building on the results of previous rounds. The method was deployed to identify a CEIAG-related task that a conversational agent could support young people with, and evaluation criteria for the success of such an intervention. It proved to be a flexible, efficient, and rigorous means to facilitate career experts' participation in the design of a conversational agent. While the results from the study are domain and task-specific, the study design provides a useful template for effective collaboration with domain experts when designing conversational agents for complex domains. As such, the paper focuses on the use of the Delphi study method in the design process.

An overview of career support services is provided as context for the Delphi design and results. Challenges in the design of conversational agents are discussed as motivation for the Delphi study, followed by details of the recruitment, data collection and analysis process. The resulting requirements for the conversational agent are provided alongside a reflective evaluation of the method.

2 BACKGROUND: CAREER SUPPORT SERVICES AND CONVERSATIONAL AGENT DESIGN

The overall aim of this project is to establish criteria for the design and evaluation of an English-language conversational agent that complements existing services for young people and meets the ethical standards required by the context. The system is intended for use by a publicly funded, national organisation that provides free-to-access career information and support to all residents. The provision of career support for individuals is considered an important mechanism for achieving national economic and social policy goals. An effective career service helps to ensure an alignment between labour market supply and demand, to support productivity and minimise unemployment. They also have an important role in supporting social equality, through reducing the impact of socio-economic background on career opportunities [11, 28]. CEIAG interventions, including those delivered by the partner organisation are informed by diverse theories from the career development literature [45]. Research in the field also emphasises the role of career support in supporting social justice [37, 40], embodied by the importance of equity in their policy and service design.

The use of automated agents for decision-making in the public sector introduces well-documented risks to equality, fairness and privacy [20, 39]. These risks have led to policy responses from a range of organisations. For example, the EU AI Act identifies high-risk areas for AI deployment, several of which overlap with public sector career services delivered in education setting [24]. UNICEF has also issued policy guidance on AI for young people [41]. In addition to general concerns about automated agents, natural language processing technologies have been identified as raising specific ethical concerns around representation, accessibility and impacts on society [3, 6]. Aligning the conversational agent with the ethical and social values of career practice is essential for meeting obligations for responsible research in this field.

A key component of the partner organisation's services are interventions with young people aged 11-18 to support the transition from school to employment, education or training. These are delivered both on-demand and routinely at key decision points, such as choosing subjects to study or preparing to leave school. There are multiple modes of intervention, including an online platform for accessing and managing career-related information as well as group and one-to-one sessions with professionally qualified career practitioners based in schools [36]. Career development theory, research and practice are consistent in recognising that career decision-making is a process rather than an event, involving ongoing experiences and interventions that interact to produce outcomes that may not be realised until years later. Standardisation of CEIAG programmes may also undermine social-justice informed approaches to career support [40]. A positive user experience will be essential for a conversational agent to be of benefit to the career service, as users already have a wide range of options for engaging with the services. However, the nature of career support means that intended users are unlikely to be able to effectively predict or evaluate the practical value of a specific intervention.

Traditional approaches to the design and evaluation of interactive systems, such user modelling, wireframing and prototyping,

are of limited value when designing conversational interactions [14, 43]. In response to these limitations, Moore et al [26] propose adopting conversation analysis as a tool for the design and evaluation of user experience. Conversational analysis identifies specific structures and strategies depending on the context and purpose of a conversation. Within CEIAG services provided by the partner organisation, at least four distinct types of conversation have been identified depending on the intervention purpose. Education, information, advice and guidance activities map to teaching, ordinary, service and counselling conversations respectively. Therefore, a clear understanding of the aims and objectives of a conversation is a pre-requisite for the design of a conversational agent that can meet users' hedonic needs and expectations. The expertise of the partner organisation staff who have experience of delivering and designing career interventions, knowledge of current services and the professional ethical standards is, therefore, critical to the design of the practical task aspects of the conversational agent. The present study addresses the following research questions, as a foundation for future work with users addressing the design of the interaction:

RQ1: In the context of existing career support services for young people, which tasks could a conversational agent support?

RQ2: How can the ethical integrity of dialogue system for use in this domain be effectively managed?

3 METHODOLOGY

Research-through-design is an approach to academic research where design practice is applied to a specific problem as a means of generating knowledge [47]. Design best practice emphasises a clear understanding of the problem space as foundational for designing quality solutions. This is encapsulated by the Design Council's 'Double Diamond' process, which provides a person-centred template for design, focused on exploration and iteration to design solutions for well-defined problems [10]. This approach has been used to analyse the challenge of designing HCI for AI technologies [44]. Furthermore, the partner organisation policies also define best practice service design based on the Design Council's 'Double Diamond' [34]. Effective stakeholder engagement is a key mitigation against the risks associated with automation in the public sector by capturing the 'concerns, values and preferences' of stakeholders as an input to the design process [19]. Domain experts also have an established role in the evaluation of dialogue systems in the literature [9, 21]. However, effectively leveraging domain expertise for design is a complex issue. Research indicates that AI developers 'de-skill' or undervalue domain experts by reducing them to tools for the harvesting of data, even when working with professionals [32]. Participatory approaches should effectively bridge the knowledge gaps between technical and domain experts to avoid overburdening participants, while maximising the contribution made to the research [4].

The aim of this study is therefore to explore the possibilities for a conversational agent, in a way that supports the effective incorporation of domain knowledge and expertise into the design and evaluation of the practical aspects of the conversational agent. CEIAG practitioners were identified as the most appropriate domain experts for this, rather than the intended system users. Design of the practical aspects of the system requires insight into the aims and

Table 1: Design dimensions for the CEIAG conversational agent

Design Dimensions	Description
System objective / task Outputs & outcomes	Specific tasks or career support objectives that the conversational agent could facilitate. Characteristics of system outputs required to support the preferred outcomes for young people from using conversational agent
Content & information sources	The type and sources of information that the conversational agent could use to formulate responses to users.
Integration with other services Ethical considerations	The role of the conversational agent within the wider service offer for young people Ensuring that the conversational agent is aligned with the ethical standards of CEIAG professionals
Personal data & privacy	Balancing potential personalisation and accessibility benefits with the need to manage users' data in line with data protection and privacy.
Persona & Interaction	Conversation design and interaction. (Further research with young people is planned to investigate this more fully.)

objectives of CEIAG interventions and existing services that young people, who are the intended end users, are unlikely to possess.

The practical and ethical components of the design from the research questions were decomposed into design dimensions, shown in table 1. These were based on high level design dimensions identified from conversational HCI research [6, 32, 33].

4 DELPHI STUDY METHOD

A Delphi study consists of successive rounds of questionnaires issued to an expert panel, where subsequent questionnaires adapt based on the results of preceding rounds [23]. The method was developed to establish the consensus viewpoint of experts in one domain to support decision making in another domain, however, it has since been adapted for research in a wide range of disciplines [18]. Building consensus should not be confused with measuring agreement. Consensus is the result of a process that involves considering the reasons individuals have for both agreeing and disagreeing, in order to refine a statement or proposal so that it more accurately reflects the collective opinion [16]. With a Delphi study, the aim is not to measure or capture a pre-existing consensus, but to use the data collection instruments to support the panel to build consensus. Genuine anonymity is critical to prevent the views of prestigious or powerful individuals exerting undue influence. This overcomes issues with participants' unwillingness to express disagreement with majority or established views, or reluctance to renege on previously asserted opinion [46].

The use of questionnaires as the main data collection instrument means that the method has the potential to be scalable to larger panels, with over 1000 participants reported in the literature [44]. However, for mixed method studies, the resources required to analyse the qualitative responses should be considered when recruiting the panel and designing the questionnaire. Panels of between 20-25 experts who participate over 3 rounds are most common in published studies [25, 46]. Recruitment should also consider the impact of participant attrition between rounds on panel size. Although attrition is an expected occurrence in this type of research, and is not generally considered detrimental to the validity of the results unless it is the result of a failure to adequately address dissent within the panel [22, 25].

Delphi studies require commitment from participants over a longer timeframe than would be expected for a design workshop, focus group or interviews. However, delivering the Delphi in an online, asynchronous format allows participants to respond at their own convenience [46]. This is an important consideration when working with geographically dispersed panels who have significant demands on their time, as in the current example of career practitioners working in the public sector. The method is particularly useful for working with qualitative data in a way that transcends domain and disciplinary boundaries [29]. The multiple rounds provide an opportunity for the expert panel to refine, reject or validate the researchers' analysis, and thereby the trustworthiness of the findings [13]. The qualitative data in the present study was provided by career experts' responses to open ended questions. However, the iterative cycles of analysis and data collection meant that this could be gradually structured by the researcher into explicit requirement statements for use in system design and evaluation, that incorporated the participants' views in their own words.

4.1 Delphi Study Participants & Structure

Participants for the Delphi study were recruited using purposive sampling from a target population of CEIAG employees with experience of CEIAG practice, service design and policy. Purposive and convenience sampling is appropriate for Delphi studies because domain knowledge is the key criterion for participation, rather than representative sample of a population required for survey methods [2]. This included both an open call to school-based practitioners and targeted recruitment through a gatekeeper. This resulted in expressions of interest from 29 practitioners with experience of practice with young people, service design and policy. As regional characteristics can have an impact on CEIAG service requirements, the geographic distribution of participants was also reviewed to ensure adequate representation of urban and rural communities. The resulting panel was highly experienced, with an average of 15 years working in the field. The majority also had experience across multiple roles within CEIAG. 23 participants went on to complete the first questionnaire, the dropout rate thereafter was low, with 20 participants completing the third and final round. This is in line with median response rates in information science Delphi studies,

where 3 rounds are also most common [25]. While early proponents of the method advocated continuing rounds until responses are stable [22], a pre-determined number of rounds allows participants to make an informed decision about the level of commitment involved, with 3 being regarded as sufficient for most studies [2].

The study was designed so that each round would incrementally progress towards explicit design requirements based on the consensus of the expert panel. An overview of the research design is shown in figure 1. The questionnaire for each round provided with an opportunity to provide qualitative responses throughout each round, as well as Likert scale responses that were used to gauge the level and topics of consensus within the group. The length of the Likert scale was progressively shortened: round 1 used a 5-point Likert, round 2 used a 4-point Likert, and round 3 offered only binary agree or disagree responses. This approach aimed to strike a balance between allowing the panel to explore the topic and producing clear guidance for the design. The median of responses was used to identify where the panel’s opinion converged [16]. The definition of consensus used was where no more than 1 participant deviated from the median response. However, qualitative data supplemented this, by clarifying the source and extent of dissensus. Rounds 2 and 3 included a summary of the quantitative results of the previous round, as well as excerpts from the qualitative data, identified during analysis. Incorporating the results of the preceding round into the data collection process is a key feature of the Delphi method. Informal feedback from participants indicated that interpreting box plots (as shown in Figure 2) was challenging. Therefore, for round 3, the quantitative results were presented using stacked bar charts (as depicted in Figure 4).. This exemplifies the Delphi study as a participatory research method, where data collection adapts based on the preferences of participants. Similarly, the term ‘chatbot’ was adopted for data collection to reflect the language used by CEIAG practitioners when discussing conversational agents.

Although the format of the responses remained relatively stable, the content of the questions was distinctive to each round. As the aim of the first round was to increase participants’ awareness of the possibilities for conversational agents, they were asked to respond to four design fictions. In round 2, participants were presented with direct quotes taken from analysis of qualitative data collected during the first round. The final round presented explicit requirement statements synthesised from the responses to the second round.

4.2 Round 1: Design Fictions

Design fictions have been successfully used in participatory design of NLP systems with non-technical domain experts [30, 35]. Although there is debate about how to incorporate these in research [1], design fictions are valued as a flexible, low-resource approach for engaging critically with speculative designs [38] and drawing out thematic patterns [1]. For this study, design fictions were written by the researcher as a means to neutrally present a range of design options to the panel in a way that did not require significant pre-existing knowledge of conversational agents. It should be noted that this portion of the research was conducted before OpenAI launched ChatGPT into the public domain, therefore views on general purpose generative agents were not solicited. As the panel were not required to have pre-existing knowledge of chatbots, the

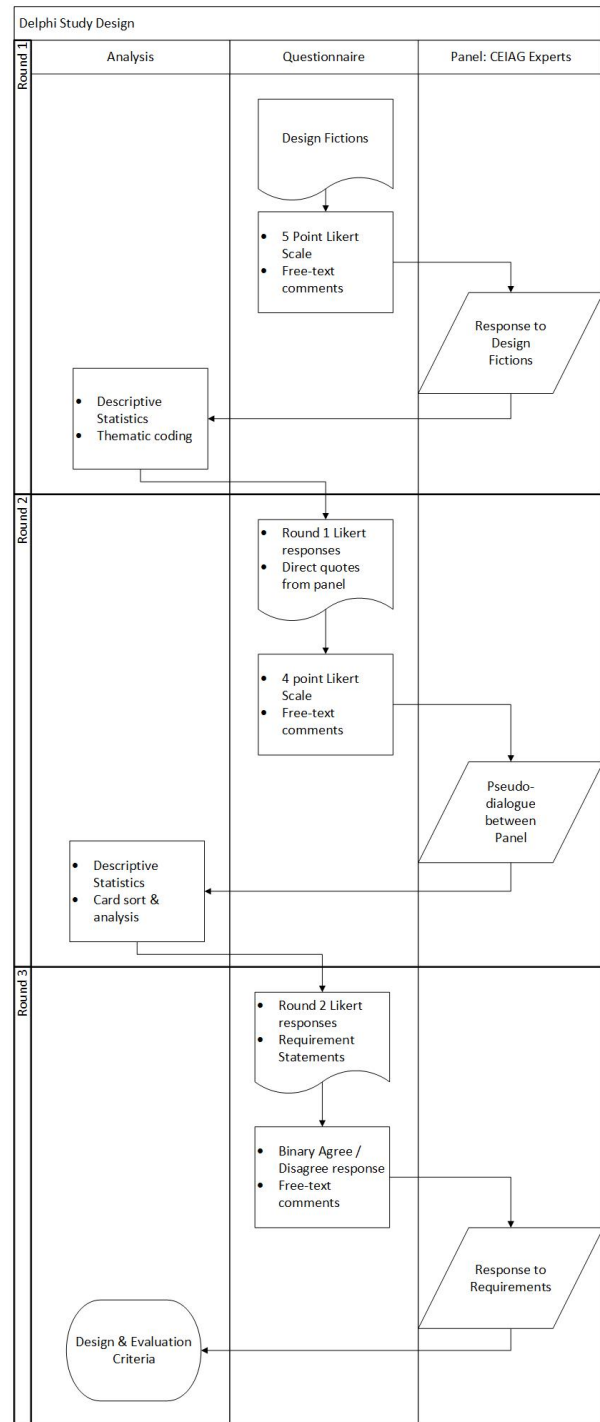


Figure 1: Overview of the Delphi Study with Career Support Experts



Figure 2: Likert scale responses to design fictions

questionnaire opened with a brief introduction to the technology, included in Appendix A. The questionnaire presented four fictions, based on CEIAG tasks that required distinct styles of conversation. These were: Information – signposting users to a range of career support services; Advice - supporting exploration of career information; Education – an agent that guided students through career education activities in a classroom setting; and Guidance - supporting personal statements writing. An example design fiction, for the Information use case, is included in Appendix B. The scenarios used in the design fictions were identified through secondary analysis of the partner organisation’s data on contacts with customers and existing research on digital CEIAG and research on the application of conversational agents in analogous domains. Each fiction included a description of the circumstances, interaction, and outcomes for an imagined young person using a conversational agent to complete a career related task. Examples of positive and negative design were included, presented neutrally in order to elicit responses that reflected the panel’s priorities and concerns. After each fiction, participants were required to use a 5-point Likert scale to respond to 7 statements regarding the potential usefulness and impact on the quality of services, shown in figure 2. Given the wide range of options presented, a neutral option was included to allow the panel to express neutrality, prevarication or ambivalence where appropriate. Free text response options were provided for comments after each scenario. Participants were also asked open ended questions about their own views on the potential uses, benefits, risks and design options for the chatbot.

All scenarios resulted in 6 median responses of ‘Agree’ and 1 ‘Neutral’ across the 7 Likert scale statements. This indicates that the panel agree that a conversational agent has the potential to be useful for CEIAG. The Education and Guidance scenarios elicited less agreement when compared to Information and Advice scenarios. The qualitative data reflected this preference, with one participant expressing that the conversational agent should ‘help

users slice through the sheer volume of information’ involved in career decision making. The quality and volume of qualitative data provided by participants indicated a significant level of engagement and reflection. Although free text comments were optional, 21 of the participants submitted responses to at least half of the free text response options available. These were analysed in two passes of qualitative coding, assisted by use of NVIVO, qualitative data analysis software. An initial pass of deductive thematic coding [31] was followed by inductive coding based on the design dimensions shown in table 1 in section 3. The deductive coding was used to identify both the representative and dissenting views, with a particular focus on identifying elements in the design fiction that elicited strong responses from the panel. These were then mapped to the design dimensions during inductive coding to ensure adequate coverage in the resulting dataset [12]. The terminology and definitions of the design dimensions were reviewed and refined to better reflect the panel’s language.

4.3 Round 2: Pseudo-dialogue

This process resulted in 50 short quotes from the qualitative data, that were used as statements for Likert scales in the second questionnaire. The use of direct quotes was intended to create a pseudo-dialogue between the career experts, in their own words. This also reduced the extent to which researchers intruded into this dialogue. Some examples are provided in figure 3 to illustrate the style and content of the statements. All of the statements are included in Appendix C. The use of abbreviations and terminology that are familiar to the panel highlights the way in which this approach allows the experts to speak to each other ‘in their own words’. An overview of the quantitative results of round 1, as shown in figure 2 were also provided to the panel as context, as is standard for the Delphi method.

As the statements in this round were based on opinions that one or more members of the panel had felt strongly enough about to

Likert Statements

Useful Universal - I think that this would be a useful addition to the universal school-age service offer.

Useful Independent - This would be a useful resource for young people access independently.

Useful Prep for Adviser - This would be useful to prepare young people for an Adviser led intervention.

Useful Group Work - This would be a useful resource for use during group work.

Improve Quality of Support - This would improve the quality of support that young people can access from careers advisers.

Improve Quality of Self-Service - This would improve the quality of self-service support available to young people.

Improve Accessibility - This would improve the accessibility of CEIAG support

- The chatbot should NOT be used for career decisions or CMS*.
 - It's important to give the correct introduction to the service and point out that it is a chatbot and has its restrictions and that a helpline adviser or their CA* can offer a personal service.
 - Some groups may struggle to use the chatbot, therefore there is a need to be careful that one group does not end up with the 'Rolls Royce service'
 - The chatbot should not include any matching activities, these can be counterproductive, and theories have developed past this way of looking at career.
 - There is a risk in the chatbot doing things a customer should be learning to do by themselves - e.g. navigating information.
- * These abbreviations are commonly used within CEIAG practice for 'Career Management Skills' and 'Careers Advisers'.

Figure 3: Examples of direct quotes used for round 2 Likert statements

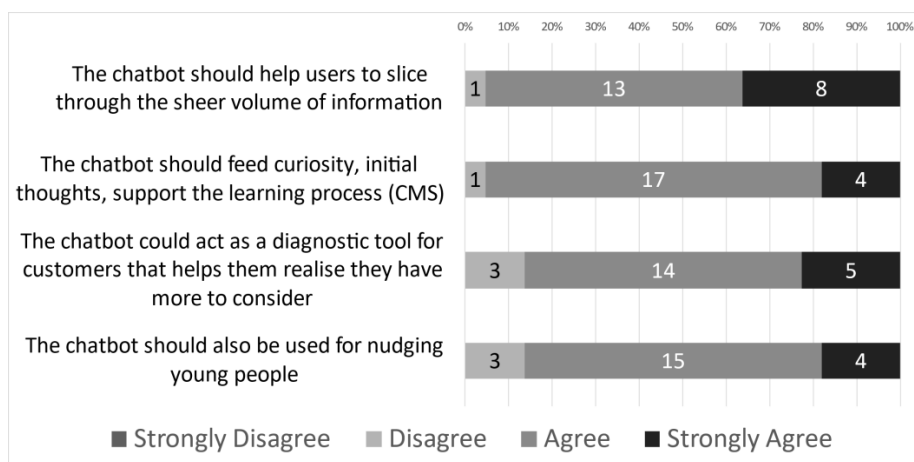


Figure 4: Summary of round 2 results for topic on Main Purpose of Chatbot

voluntarily provide, the neutral option was removed. This meant that the panel were required to provide an explicit opinion for each statement in this round. Due to the number and heterogeneity of the Likert statements, a physical card sort was used to create small groups of 3-5 statements on related topics. Details of the topic groups and associated statements are included in Appendix B. Participants responses to statements for each topic group were analysed independently. For each topic group, the statements were ranked based on the panel’s overall level of agreement, which was measured by assigning numerical values for each individual Likert response (where Strongly Agree = 4, and Strongly Disagree = 1). Descriptive statistics were used to identify the level of consensus for each statement, an example of this for the top is shown in CMS. Deductive analysis of the qualitative data identified participants’ reasons for agreeing or disagreeing [5]. These were then used to produce a requirement statement that aimed to reflect the panel’s

views for each topic. These were further refined to reduce duplication across the list.

4.4 Round 3: Requirement Statements

This resulted in 16 requirement statements shown in table 2. To illustrate the requirement synthesis, the statement that resulted from the results shown in figure 4 was: ‘The chatbot should support young people to navigate information in a way that encourages curiosity and exploration’. This combines the key elements of the statements with highest rates of panel agreement using language that is meaningful to the domain experts (‘curiosity’) and system designers (‘exploration’). For some complex topics, two contradictory statements were produced to demonstrate mutually exclusive approaches to the design. For example, two requirement statements presented alternative options for handling a situation where a user declines to be referred to other services when their query is too

complex for the conversational agent. The requirement statements were presented to the panel for validation in round 3. Alongside each statement, the round 2 quantitative results relevant to that requirement were presented as shown in figure 4. Quotes from qualitative data were also included where additional context was required. As this was the final round of questionnaire, it was important to be able to clearly gauge the panel's views. Therefore, participants were asked to respond to each requirement statement with a binary agree or disagree. This binary choice meant that participants could signal doubts clearly, as there would be no further scope for the panel to feedback. The limitations of this were mitigated by including the option to provide free text comments for each statement.

5 RESULTS

Consensus, and thereby full validation of requirements, was defined as no more than one participant disagreeing with the majority response, and no significant objections or caveats in the qualitative data. The panel's responses to the statements are shown in table 2. The panel reached consensus on 10 out of the 16 of the requirements, including at least one for each design dimension. A further two statements generated near consensus, where only two participants disagreed. The qualitative data for 3 of the statements that did not generate consensus (Req. IDs 4, 9 & 12), indicated that their rejection was rooted in ambiguity in the statement wording. These can therefore be refined for clarity, rather than rejected. Comparison of responses to statements related to data privacy, password protection and personalisation trade-offs (Req. IDs 14 & 15) indicate a moderate preference, but not consensus for password-free access with lower customisation. Similarly, the combined responses to the two statements that referred to options for handling referral to a different type of service (Req. IDs 11 & 13) provide guidance that can be further refined during the design process but cannot yet be considered as fully validated. Overall, the response to the final round produced 10 fully validated requirements and 3 that can be considered partially validated, pending re-wording. The analysis also provided guidance for the topics where requirements were not fully validated. These will be explored more fully in subsequent design activities with users and CELAG experts. The majority of the requirements produced through the Delphi process are functional requirements that relate to how the conversational agent should help users to meet the task requirements. However, requirements related to the development process (Req IDs 2, 12 & 16), and the context of deployment (Req IDs 6 & 9) were also captured as part of this process.

6 DISCUSSION AND FUTURE WORK

The Delphi study has therefore addressed RQ1 by identifying information navigation as a career support task that a conversational agent could usefully support young people with. Furthermore, this includes a clear definition of task success that can be used to support user-evaluation of the practical aspects of the conversational agent. Req. ID 1 articulates a clear target outcome for young people interacting with the conversational agent: increased curiosity and exploration. Existing metrics for measuring curiosity [33, 42] and exploration [17] from career development research can therefore be

incorporated into the evaluation of the conversational agent. Req IDs 2-7 provide guidance on how this aim can be achieved. RQ2 has been partially addressed by the remaining requirement statements. Req IDs 8, 9, 10, 12, 14 and 16 have been validated by the panel as approaches that will reduce the risk of harm to young people as a result of the introduction of a conversational agent to the partner organisation's services. Open questions remain regarding ethical handling of conversational breakdown. Similarly, consensus was not reached on the role of password protection, although a preference for password-free access was identified. User experience implications and users' preferences are important factors in this design decision, therefore this will be included in planned further research with the system's intended users.

Although the contents and results of this study are specific to the current design problem, the design dimensions used here to define the scope of the collaboration may be useful for the design of research with experts in other domains. Similarly, the resulting requirements may serve as useful inputs to the ideation phase of designing conversational interfaces for analogous domains. However, the aim of this study was to produce design criteria that reflect the specific intended use context, rather than generalisable requirements.

The Delphi study method described here may be useful for researchers seeking to engage domain experts in exploratory research. The use of design fiction in the first round was effective in eliciting a rich qualitative dataset of expert's insights. These then formed the second round that allowed the panel to respond directly to their peers' opinions. The third round offered the researchers with an opportunity to structure the participants' views to meet the needs of the research. The iterative nature of the study, where the results of each round were analysed before the next questionnaire was designed meant that the study was responsive to participants' data and preferences. It also embeds a form of member check, as participants' responses also served to validate or correct the analysis of the preceding rounds. Member checking is a commonly used approach for demonstrating the validity of qualitative analysis [13].

However, the nature of the method does restrict the format and extent of the influence that participants have over the design, when compared to co-located, synchronous design activities. The Delphi study is not proposed as an alternative to co-design methods that provide participants with an opportunity to directly shape the prototypes and interaction design [7]. However, this study has shown it to be an efficient and effective method for collaborating with stakeholders in the early discovery and problem definition stages of development [10].

The maximum total time commitment for individual experts was 3 hours, with participation at their convenience. The final questionnaire also asked for feedback about participating in the research. Although only four responses to this question were received, they indicated that the Delphi study had been successful in mitigating known issues regarding the devaluing and overburdening of domain experts in software development. Participants stated that they had enjoyed the experience, which is further evidenced by the low dropout rate and high levels of engagement in the form of qualitative responses. One participant stated that the opportunity to reflect on their own practice during the study was highly valued.

Table 2: Summary of response to requirement statements

Req. ID	Design Dimension	Topic	Requirement Statement	Agree	Dis-agree
1	System objectives or task focus	Main purpose of chatbot	The chatbot should support young people to navigate information in a way that encourages curiosity and exploration.	19	1
2	System objectives or task focus	Accessibility advantages	The chatbot should be thoroughly tested to ensure it meets the needs of young people using it as an independent self-service route to access support.	20	0
3	Outputs & outcomes	Risk of pigeonholing users	The chatbot should ensure that the range of information presented is broad enough to encourage users to explore their options further.	20	0
4**	Outputs & outcomes	Chatbot boundaries	It should be clear to users that the chatbot is not intended as a tool for career decision-making.	18	2
5	Content & information sources	Information Sources	The chatbot should focus on guiding users through existing partner organisation-managed information, but it may be appropriate to direct users to carefully selected external information sources where required.	20	0
6	Integration with other services	Mitigating risks	The chatbot should ensure that all users are aware of how to access other sources of support from partner organisation (e.g. Helpline, appointment with careers adviser)	20	0
7	Integration with other services	Relationship with Existing Services	The chatbot should function well as a tool for independent use. It should not require significant changes to existing services in order for potential benefits to be realised for young people.	19	1
8	Integration with other services	Risk of overwhelming users	Users should be made aware of alternative sources of support, and how to access them before any potentially overwhelming responses are provided.	19	1
9**	Integration with other services	Accessibility advantages	To ensure that the chatbot increases young people's access routes to career support, it should only be introduced as an additional complement to, not replacement for, any aspect of existing services.	18	2
10	Ethical considerations	Risk of pigeonholing users	Ensuring that users understand the scope and limitations of the chatbot is important for aligning with the partner organisation's approach to career support.	20	0
11*	Ethical considerations	Mitigating risks	If it is unclear what level of support a user required, it would be preferable for the chatbot to encourage the user to contact the helpline or a careers adviser, before continuing the interaction, even though this may result in some unnecessary calls / appointments	17	3
12**	Ethical considerations	Barriers to Access	The Equality Impact Assessment for the chatbot should include individuals who face digital literacy and/or technology infrastructure barriers. Any negative impacts identified should be fully mitigated before the chatbot is introduced.	16	4
13*	Ethical considerations	Mitigating risks	There may be occasions where it is clear that a user requires a level of support beyond the scope of the chatbot but continues the interaction even after being advised to contact the helpline or an adviser. In these circumstances it would be preferable for the chatbot to reiterate other sources of support available and end the conversation in order to avoid the risk of confusing the young person, even although this may mean their experience with the chatbot is perceived negatively.	11	9
14	Personal data & privacy	Personal data & privacy	Customisation of responses should be based on high level, non-sensitive information provided by users during the conversation only. (For example: whether user is in school/unemployed/college etc; non-specific location).	19	1
15*	Personal data & privacy	Personal data & privacy	The chatbot should not be password protected, and therefore should not store or process personal or sensitive data.	17	3
16	Persona & Interaction	Risk of overwhelming users	The chatbot should be tested with young people to determine the appropriate volume and complexity of information to be included in chatbot responses.	20	0

* indicates non-validated requirements. ** indicates partially validated requirements.

This indicates that Delphi studies can not only support researchers to access domain expertise, but it can also benefit participants.

While there were some limitations to the validation of the requirements, the qualitative data meant that these issues were simple to identify and remedy. The varied approach to questionnaire design and analysis across the three rounds of the Delphi was challenging in that it required a high degree of adaptability and responsiveness from the researchers. However, it yielded significant benefits in the ability to efficiently engage a cohort of domain experts, without requiring them to acquire significant technical knowledge. Direct feedback and attrition rates indicate that the domain experts did not find participation burdensome. They remained engaged from initial exploration through to validation of the criteria for the design and evaluation of the chatbot. The mixed methods Delphi study described here is a highly effective tool for participatory design with domain experts. It effectively supports the design of a conversational agent that will complement the existing services and reflect the professional and ethical standards of CEIAG practitioners. Further work with users, including a pilot study, is planned to ensure that the conversational agent meets users' interactional preferences, as well as the practical outcomes identified by the domain experts.

ACKNOWLEDGMENTS

We are grateful for the funding provided by the Scottish Graduate School of Social Science and Skills Development Scotland, as well as the ongoing support from Sandra Cheyne at Skills Development Scotland. Our appreciation extends to the reviewers for their insightful feedback, colleagues who offered guidance throughout the research, and most importantly, the research participants who generously contributed their time and perspectives, making this study possible.

REFERENCES

- [1] Eric P. S. Baumer, Mark Blythe, and Theresa Jean Tanenbaum. 2020. Evaluating Design Fiction: The Right Tool for the Job. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, July 03, 2020. ACM, Eindhoven Netherlands, 1901–1913. <https://doi.org/10.1145/3357236.3395464>
- [2] Ian Belton, Alice MacDonald, George Wright, and Iain Hamlin. 2019. Improving the practical application of the Delphi method in group-based judgment: A six-step prescription for a well-founded and defensible process. *Technological Forecasting and Social Change* 147, (October 2019), 72–82. <https://doi.org/10.1016/j.techfore.2019.07.002>
- [3] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmar-Garet Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, 2021. Association for Computing Machinery, New York, NY, USA, 610–623. <https://doi.org/10.1145/3442188.3445922>
- [4] Abeba Birhane, William Isaac, Vinodkumar Prabhakaran, Mark Diaz, Madeleine Clare Elish, Jason Gabriel, and Shakir Mohamed. 2022. Power to the People? Opportunities and Challenges for Participatory AI. In *EAAMO '22: Equity and Access in Algorithms, Mechanisms, and Optimization*, October 06, 2022. Association for Computing Machinery. <https://doi.org/10.1145/3551624.3555290>
- [5] Norman Blaikie and Jan Priest. 2017. *Social Research: Paradigms in Action*. Polity Press, Cambridge. Retrieved from [https://ebookcentral.proquest.com/lib/ed/detail.action?docID=\\$4783896](https://ebookcentral.proquest.com/lib/ed/detail.action?docID=$4783896)
- [6] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (Technology) is Power: A Critical Survey of “Bias” in NLP. In *58th Annual Meeting of the Association for Computational Linguistics*, 2020. <https://doi.org/10.18653/v1/2020.acl-main.485>
- [7] Francisco Maria Calisto, João Fernandes, Margarida Morais, Carlos Santiago, João Maria Abrantes, Nuno Nunes, and Jacinto C. Nascimento. 2023. Assertiveness-based Agent Communication for a Personalized Medicine on Medical Imaging Diagnosis. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 19, 2023. Association for Computing Machinery, New York, NY, USA, 1–20. <https://doi.org/10.1145/3544548.3580682>
- [8] CEDEFOP; ETF; European Commission. *Investing in career guidance: Revised edition 2021*. Inter-Agency Working Group on Career Guidance WGCG. Retrieved October 17, 2023 from <https://www.cedefop.europa.eu/en/publications/2230>
- [9] Jan Deriu, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, and Mark Cieliebak. 2021. Survey on evaluation methods for dialogue systems. *Artificial Intelligence Review* 54, 1 (2021), 755–810. <https://doi.org/10.1007/s10462-020-09866-x>
- [10] Design Council. The Double Diamond. Retrieved February 11, 2024 from <https://www.designcouncil.org.uk/our-resources/the-double-diamond/>
- [11] Valerie Egddell and Peter Robertson. 2020. A critique of the Capability Approach's potential for application to career guidance. *International Journal for Educational and Vocational Guidance* (October 2020), 1–17. <https://doi.org/10.1007/s10775-020-09445-0>
- [12] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods* 5, 1 (March 2006), 80–92. <https://doi.org/10.1177/160940690600500107>
- [13] Lisa Given. 2008. Member check. In *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States. <https://doi.org/10.4135/9781412963909>
- [14] Hendrik Heuer and Daniel Buschek. 2021. Methods for the Design and Evaluation of HCI+NLP Systems. In *Proceedings of the First Workshop on Bridging Human-Computer Interaction and Natural Language Processing*, February 26, 2021. 28–34. Retrieved October 12, 2021 from <https://arxiv.org/abs/2102.13461v1>
- [15] Tristram Hooley and Tom Staunton. 2021. The Role of Digital Technology in Career Development. In *The Oxford Handbook of Career Development*, Peter J. Robertson, Tristram Hooley and Phil McCash (eds.). Oxford University Press, Oxford, 296–312. <https://doi.org/10.1093/oxfordhb/9780190069704.013.22>
- [16] Chia-Chien Hsu and Brian A. Sandford. 2007. The Delphi Technique: Making Sense of Consensus. *Practical Assessment, Research, and Evaluation* 12, (August 2007). <https://doi.org/10.7275/PDZ9-TH90>
- [17] Zhou Jiang, Alexander Newman, Huang Le, Alfred Presbitero, and Connie Zheng. 2019. Career exploration: A review and future research agenda. *Journal of Vocational Behavior* 110, (February 2019), 338–356. <https://doi.org/10.1016/j.jvb.2018.08.008>
- [18] Adrianna Kezar and Daniel Maxey. 2016. The Delphi technique: an untapped approach of participatory research. *International Journal of Social Research Methodology* 19, 2 (March 2016), 143–160. <https://doi.org/10.1080/13645579.2014.936737>
- [19] Pascal D. König and Georg Wenzelburger. 2021. The legitimacy gap of algorithmic decision-making in the public sector: Why it arises and how to address it. *Technology in Society* 67, (November 2021), 101688. <https://doi.org/10.1016/j.techsoc.2021.101688>
- [20] Maciej Kuziemski and Gianluca Misuraca. 2020. AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy* 44, 6 (2020), 101976. <https://doi.org/10.1016/j.telpol.2020.101976>
- [21] Xiaomin Li, Chuanqi Tao, Jerry Gao, and Hongjing Guo. 2022. A Review of Quality Assurance Research of Dialogue Systems. In *2022 IEEE International Conference On Artificial Intelligence Testing (AITest)*, August 2022. 87–94. <https://doi.org/10.1109/AITest55621.2022.00021>
- [22] Harold A Linstone and Murray Turoff. 2002. *The Delphi Method: Techniques and Applications*.
- [23] Harold A. Linstone and Murray Turoff. 2011. Delphi: A brief look backward and forward. *Technological Forecasting and Social Change* 78, 9 (2011), 1712–1719. <https://doi.org/10.1016/j.techfore.2010.09.011>
- [24] Luciano Floridi, Matthias Holweg, Mariarosaria Taddeo, Javier Amaya Silva, Jakob Mökander, and Yuni Wen. 2023. *capAI A procedure for conducting conformity assessment of AI systems in line with the EU Artificial Intelligence Act*. <http://10.2139/ssrn.4064091>
- [25] Brady D. Lund. 2020. Review of the Delphi method in library and information science research. *JD* 76, 4 (February 2020), 929–960. <https://doi.org/10.1108/JD-09-2019-0178>
- [26] Robert J. Moore, Margaret H. Szymanski, Raphael Arar, and Guang-Jie Ren. 2018. *Studies in Conversational UX Design*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-95579-7>
- [27] Observatory on Digital technologies in Career guidance for Youth (ODiCY). 2023. *CiCi, the 24/7 careers chatbot*. Organisation for Economic Co-operation and Development. Retrieved from <https://www.oecd.org/stories/odicy/practices/ci-ci-the-24-7-careers-chatbot-9ccfd77d/>
- [28] Wendy Patton. 2008. Recent Developments in Career Theories: The Influences of Constructivism and Convergence. In *International Handbook of Career Guidance*, J.A. Athanasou and R. Van Esbroeck (eds.). Springer Netherlands, 133–156. https://doi.org/10.1007/978-1-4020-6230-8_7
- [29] Alison Jane Pickard. 2013. Delphi study. In *Research Methods in Information*. Facet, 149–156. <https://doi.org/10.29085/9781783300235.017>

- [30] Ronda Ringfort-Felner, Matthias Laschke, Shadan Sadeghian, and Marc Hassenzähl. 2022. Kiro: A Design Fiction to Explore Social Conversation with Voice Assistants. In *Proceedings of the ACM on Human-Computer Interaction*, January 14, 2022. Association for Computing Machinery. <https://doi.org/10.1145/3492852>
- [31] Johnny Saldaña. 2015. *The coding manual for qualitative researchers*. Sage Publications, Inc., Thousand Oaks. Retrieved February 1, 2021 from <https://app.talis.com/ed/player#/modules/5f48c0462b1a275adc78b1de/resources/5f48d6382b1a275adc78b341>
- [32] Nithya Sambasivan and Rajesh Veeraraghavan. 2022. The Deskillling of Domain Expertise in AI Development. *Conference on Human Factors in Computing Systems - Proceedings (2022)*. <https://doi.org/10.1145/3491102.3517578>
- [33] Mark L. Savickas and Erik J. Porfeli. 2012. Career Adapt-Abilities Scale: Construction, reliability, and measurement equivalence across 13 countries. *Journal of Vocational Behavior* 80, 3 (June 2012), 661–673. <https://doi.org/10.1016/j.jvb.2012.01.011>
- [34] Scottish Government. 2019. The Scottish Approach to Service Design How to design services for and with users.
- [35] Dilruba Showkat and Eric P.S. Baumer. 2022. "It's Like the Value System in the Loop": Domain Experts' Values Expectations for NLP Automation. In *DIS 2022 - Proceedings of the 2022 ACM Designing Interactive Systems Conference: Digital Wellbeing*, June 13, 2022. Association for Computing Machinery, Inc, 100–122. <https://doi.org/10.1145/3532106.3533483>
- [36] Skills Development Scotland. Our Careers Service in Schools. Retrieved April 4, 2021 from <https://www.skillsdevelopmentscotland.co.uk/what-we-do/scotland-careers-services/our-careers-service-in-schools/>
- [37] Tom Staunton and Karla Rogosic. 2021. Labour market information and social justice: a critical examination. *International Journal for Educational and Vocational Guidance* 21, 3 (October 2021), 697–715. <https://doi.org/10.1007/S10775-021-09466-3/TABLES/2>
- [38] Bruce Sterling. 2009. Design fiction. *interactions* 16, 3 (May 2009), 20–24. <https://doi.org/10.1145/1516016.1516021>
- [39] Luca Tangi, Colin van Noordt, and A. Paula Rodriguez Müller. 2023. The challenges of AI implementation in the public sector. An in-depth case studies analysis. In *Proceedings of the 24th Annual International Conference on Digital Government Research (DGO '23)*, July 11, 2023. Association for Computing Machinery, New York, NY, USA, 414–422. <https://doi.org/10.1145/3598469.3598516>
- [40] Rie Thomsen. 2016. The practice portrait – a method for promoting social justice in practice. *Journal of the National Institute for Career Education and Counselling* 36, 1 (2016), 30–37.
- [41] UNICEF. 2021. *Policy guidance on AI for children*. Retrieved November 10, 2022 from <https://www.unicef.org/globalinsight/reports/policy-guidance-ai-children>
- [42] M. Fernanda Wagstaff, Gabriela L. Flores, Rawia Ahmed, and Sarah Villanueva. 2021. Measures of curiosity: A literature review. *Human Resource Development Quarterly* 32, 3 (2021), 363–389. <https://doi.org/10.1002/hrdq.21417>
- [43] Qian Yang, Justin Cranshaw, Saleema Amershi, Shamsi T Iqbal, and Jaime Teevan. 2019. Sketching NLP: A case study of exploring the right things to design with language intelligence. In *Conference on Human Factors in Computing Systems - Proceedings*, 2019. ACM, New York, NY, USA. <https://doi.org/10.1145/3290605.3300415>
- [44] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Conference on Human Factors in Computing Systems - Proceedings (April 2020)*. <https://doi.org/10.1145/3313831.3376301>
- [45] Julia Yates. 2020. Career Development Theory: An Integrated Analysis. In *The Oxford Handbook of Career Development*, Peter J. Robertson, Tristram Hooley and Phil McCash (eds.). Oxford University Press, Oxford. <https://doi.org/10.1093/oxfordhb/9780190069704.013.10>
- [46] Jhon Wilder Zartha Sossa, William Halal, and Raul Hernandez Zarta. 2019. Delphi method: analysis of rounds, stakeholder and statistical indicators. *Foresight* 21, 5 (September 2019), 525–544. <https://doi.org/10.1108/FS-11-2018-0095>
- [47] John Zimmerman and Jodi Forlizzi. 2014. Research through design in HCI. In *Ways of Knowing in HCI*. Springer, New York, NY, 167–189. https://doi.org/10.1007/978-1-4939-0378-8_8

APPENDIX A: INTRODUCTION TO CHATBOTS

Dialogue systems (also known as chatbots or conversational agents) allow users to interact using natural, conversational language in order to access digital information and resources. This survey is about screen-based / text-only systems.

The most common versions of these are fairly limited, often used to increase efficiency through reducing the volume of simple customer service enquiries. However, there are examples of more specialised systems, for example used for mental health support, or educational tutoring.

When developing a dialogue system, you have to ‘teach’ them three things

- deciphering what has been typed – ‘natural language understanding’,
- how to figure out a response – ‘information representation’
- how to put that response into words – ‘natural language generation’.

This means system developers have to decide what information is relevant for the system’s task, including the kind of language it should be able to understand and translate that into data for the computer. Then they system has to have some method of identifying what data is relevant to the system’s response. This is usually achieved by applying statistical methods to the available data. Then decisions have to be made about natural language generation, including the tone and ‘persona’ of the system.

The complexity of language understanding and generation means that ensuring the system has a clearly defined scope is important for ensuring that users have a positive experience. This is the focus of my research- designing a dialogue system that can be used as part of the existing CEIAG services for Scottish young people. The first stage is deciding on the what the purpose of the chatbot should be in order to ensure it will be genuinely useful for young people. Also, as the system will, in a way, be acting on behalf of SDS, at a minimum, it should be aligned with their ethical standards and policies. To be useful, it needs to support, improve or extend existing services in some way.

APPENDIX B: EXAMPLE OF DESIGN FICTION FOR SIGNPOSTING USE CASE

Situation

A young person is approaching school leaving age, and is confused about what their options are if they decide not to stay on. They are interested in a few different career areas, and are aware of various routes such as work, apprenticeships, college & university, but unsure of the details and differences between them. They would like some practical information about what each route entails (e.g. financial support, requirements for entry, application dates) to help them evaluate the short- and medium-term differences between the routes.

Accessing Chatbot

During group work, their careers adviser had mentioned a chatbot that could help them to find out information about the routes and schemes available post school. Their careers adviser had explained that this was a signposting service that would direct young people to specific information & organisations, not an advice or guidance service. They use their mobile to access the service through a web browser while on the bus on a Sunday morning.

The Interaction

The chatbot provides a brief overview of what it can do, and directs users to Skills Development Scotland’s helpline and school careers adviser if their query cannot be answered by the chatbot. They answer some questions about their current qualifications, location, and the kind of opportunities they are interested in. The chatbot responds to each user input with a standard acknowledgement and asks some clarifying questions. If the young person takes longer than average to respond to a question, the system sends

a message directing them to the helpline if they need some help. When the user indicates that they have no more information they wish to provide, the chatbot summarises the information they have been given, to check for accuracy and provide opportunity for correction.

Outputs

The chatbot sends a single message containing links to specific pages that contain information about application processes and financial support on the user's local college website, and other websites specific to partner organisation. Links to information pages of local schemes that aim to support school leavers accessing careers in the third-sector and green economy are also provided. These are links to specific, relevant pages within the site, not links to the home pages of each organisation. The user is given the option of having this information (and/or a full transcript of the chat) emailed to them. They are directed to their school careers adviser and the helpline for further support during the conversation close.

APPENDIX B: ROUND 2 TOPICS AND STATEMENTS

Topic: Main Purpose of Chatbot

- The chatbot should help users to slice through the sheer volume of information.
- The chatbot should feed curiosity, initial thoughts, support the learning process (CMS).
- The chatbot could act as a diagnostic tool for customers that helps them realise they have more to consider.
- A chatbot could be a good way of generating ideas and conversation around job roles.
- The chatbot should also be used for nudging young people.

Topic: Barriers to Access

- Technical issues or lack of digital skills are very real issues that could lead to frustrations or abandonment.
- Chatbot roll out should factor in digital exclusion and equity as there is the danger of those with more support/resources gleaning the most benefit.
- Connectivity issues and access to ICT in schools may be an issue.
- Some groups may struggle to use the chatbot, therefore there is a need to be careful that one group does not end up with the 'Rolls Royce service'.

Topic: Accessibility Advantages

- The chatbot could be useful in a situation where a young person wants to find out what support is available but doesn't want to disclose to a person.
- We know from insight that young people in particular can be reluctant to pick up a phone so they may prefer a chatbot as a first interface with SDS.
- The chatbot could be, more accessible for some people in some ways.
- The chatbot could extend the scope and variety of access routes to CIAG support.
- The chatbot is likely to get some young people engaging who wouldn't have engaged in any other way.

Topic: Information Sources

- A chatbot should allow a young person to be guided through the information on MyWoW in a meaningful way.
- The chatbot should draw from a wider range of resources than MyWoW.
- The chatbot should be diagnostic to narrow down which areas of MyWoW the young person should be using.

Topic: Risk of Overwhelming Users

- If the information presented to customers is too complex and there is too much of it, then some customers may be put off.
- The chatbot should ONLY be for simpler questions to free up time for more detailed guidance support.
- Experience would tend to suggest if people are overwhelmed by information they tend to walk away rather than navigate a new system or help function.
- The user may be further confused by the information presented.

Topic: Risk of Pigeonholing Users

- Often careers conversation go into many life experiences which affect the individual's view of self, etc, that the chatbot may not be able to respond to appropriately.
- There is a risk that young people might feel pigeon-holed by the information provided by a chatbot.
- There is a risk that young people would access the chatbot and then think the job is done, leaving no room for CMS.
- There is a risk in the chatbot doing things a customer should be learning to do by themselves - e.g. navigating information.

Topic: Mitigating Risks

- Young people using the chatbot should be able to request an appointment with an adviser if they still have issues/question when using the chatbot.
- It's important to give the correct introduction to the service and point out that it is a chatbot and has its restrictions and that a helpline adviser or their CA can offer a personal service.
- The chatbot should recognise when an adviser engagement is needed and help them to arrange this.
- A clear narrative should be given on the purpose of the chatbot.
- The young person should always have a face-face appointment to discuss the information they have gathered.

Topic: Personal Data & Privacy

- Having to register or remember passwords can create a barrier.
- An option to filter by location to receive information specific to where they live (or intend to live).
- The chatbot should clarify who the user is (school, unemployed, redundant etc) to begin with.
- The chatbot should have access to existing MyWoW profiles and personal information and then users should have the option to add to this or remove anything they felt wasn't relevant.

Topic: Chatbot Boundaries

- The chatbot should not claim to help you make a decision.

- The chatbot should NOT be used for career decisions or CMS.
- The chatbot should present all information equally so as to not steer a young person down a particular path.
- The chatbot should not include any matching activities, these can be counterproductive and theories have developed past this way of looking at career.

Topic: Chatbot Content

- The chatbot should provide information about sources of funding and benefits.
- The chatbot should suggest games and activities to help with the learning process.
- A chatbot should help young people to find out the skills different courses demand, in simple terms.
- The chatbot should provide information about how to find a job, or apply to college.
- The chatbot should provide information about Foundation Apprenticeships - this may remove the need for each person registering interest needing to be contacted by an Adviser.

Topic: Relationship with Existing Services

- The chatbot can supplement the service we offer, allowing coaches to coach and give less information.
- Robust career education messages should be provided alongside the chatbot.
- A chatbot could be a good way of generating ideas and conversation around job roles.
- The chatbot could provide some engagement during the times when a careers adviser is not available.

Topic: Relationship to other interventions

- Before appointments with a Careers Guidance Practitioner, the chatbot could be used to explore some ideas and gain a better understanding of what happens during appointment with a practitioner.
- The chatbot has some potential to be effectively used in a whole class / group setting.
- The chatbot should be introduced by careers adviser, not as a stand alone tool.
- The chatbot should only be used when the advice and guidance has already taken place and the young person is confident with their decision but needs some specific direction or advice in relation to their next steps.