Helping Professionals Select Persona Interview Questions Using Natural Language Processing

Joni Salminen[1,2(✉)], Kamal Chhirang[3], Soon-gyo Jung[1], and Bernard J. Jansen[1]

[1] Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar
jsalminen@hbku.edu.qa
[2] Turku School of Economics at the University of Turku, Turku, Finland
[3] Fulda University of Applied Sciences, Fulda, Germany

Abstract. Personas are often created based on user interviews. Yet, researchers rarely make their interview questions publicly available or justify how they were chosen. We manually extract 276 interview questions and categorize them into 10 themes, making this list publicly available for researchers and practitioners. We also demonstrate an approach of using natural language processing to assist in selecting persona interview questions for a given use case.

Keywords: Personas · Information design · Interview questions

1 Introduction

A persona is a fictitious person describing a real user or customer segment [1]. Personas are presented as profiles containing key information about a particular user segment, such as goals, needs, and wants [2]. Personas are used for multiple purposes, such as communicating information about the users within teams [3], aligning preconceptions with reality [4], and to better understand user needs for requirements engineering [5].

The approaches to persona creation include qualitative [6], quantitative (including algorithmic or data-driven personas [7–9]), and mixed methods [10]. Most often, personas are created using qualitative methods [6, 10]. In such cases, the data is collected via ethnographic means, field studies, or interviews [6]. Out of all of these, interviews are often preferred because they can be easily implemented.

After the persona creators have decided their methodological approach (e.g., qualitative) and data collection means (e.g., interviews), they typically consider what information the persona profile contains. This selection process is referred to as persona information design [11], defined as selection of information elements (attributes, characteristics) that the finalized persona profiles will communicate to their users. Ideally, the persona information design process is driven by the information needs of the eventual persona users, such as software developers, designers, and business managers.

Persona information design deals with the fundamental question of what information personas should contain. As personas are often created based on user interviews [6, 12], practitioners frequently ask for a comprehensive list of persona interview questions.

In this research, we offer a meta-synthesis of common persona interview questions, identified from top-ranking search results on Google. The results offer 276 interview questions categorized into 10 themes. We provide suggestions on how researchers and practitioners can select the appropriate questions for their particular persona use case.

Our contribution addresses the challenge in HCI, and in the field of personas, that designers often have too many interview questions to choose from. Our data collection shows that, even when using a limited number of nine sources, one ends up with a list of more than two-hundred possible questions. This excessive variety and dimensionality of potential interview questions results in a daunting paradox of choice for those creating personas, especially to those that are less inexperienced, such as students and novices, with the method and, therefore, less certain of what questions to include in their user interviews. As such, our contribution offers help and guidance for researchers and practitioners dealing with the challenge of persona information design.

2   Related Literature

Persona design is discussed in several textbooks [1, 2, 6, 10]. All of these sources mention interviews as a form of data collection for persona creation. According to Google Scholar, "persona information design" is explicitly mentioned by three studies: [9, 11, 13]. These studies briefly mention the concept, defining it as the choice of what information to include in the persona profiles (regardless of the source of data).

The generic persona creation process, when using interviews, is as follows [6]:

Use case definition → Persona information design → Interview question selection → Conducting interview → Analyzing the results → Creating the persona profiles

First, one defines the use case that gives motivation and explains why personas are needed (e.g., "we need to improve employee satisfaction and believe that using personas to understand our employees would be useful"). Second, the information design of the personas is defined – what do we want to know about the people the personas portray? (e.g., employee pain points, their life situation, their decision making). This will result in the identification of various information elements to include in the personas. Then, we operationalize these information elements by creating a list of interview questions. This is followed by the interviews, the transcription and analysis of the results, and the creation (write-up) of the persona profiles [6].

Reviewing past works reveals a consistent pattern wherein researchers have constructed personas based on user interviews but have not disclosed their interview questions. For instance, in their work to understand users for product design, DiMicco and Mann [14] created three personas from interviews but there was no mention of the questions used. On similar lines, Gao et al. [15], Chang et al. [16], Falk et al. [17], Antle [18], and Sharbatdar et al. [19] conducted studies to understand medical prenatal care users, language learners, decision makers in the oil and gas industry, children, and transportation management practitioners, respectively, and created several personas but there was no mention of the interview questions applied.

Therefore, a consistent pattern in prior works is that researchers rarely disclose the interview questions which were used to collect data for making personas. Not only does

this make replication of their studies difficult, but also leaves practitioners with a gap of knowledge, as academic sources do not explicitly recommend specific interview questions. Our study addresses this issue by proposing a systematic approach of identifying persona interview questions.

3   Methodology

3.1   Data Collection

We searched Google to identify lists of persona interview questions. A search using "persona interview questions" yields 13,900 results in Google (29,100,000 results without brackets). We chose six sources based on (a) high rank in Google and (b) extensiveness of the list of questions. We then manually collected all the questions from each source website and stored them in a spreadsheet. This yielded 350 questions. These were manually screened to remove duplicates, i.e., asking the same question twice. In total, 74 (21.1%) duplicates were found and removed. After the screening process, 276 questions remained. The original and the final list of questions after applying the screening criteria have been made available in the Supplementary Material[1].

3.2   Classifying the Questions

The remaining 276 questions were classified under ten themes (see Table 1). A theme taxonomy was created based on abduction, i.e., a qualitative technique which involves using both pre-existing knowledge and also analyzing the emerging properties of the data [20]. In our case, we made use of (a) the authors' pre-existing knowledge of persona information content, and (b) the emerging communalities in the question material.

In this process, one of the researchers created the themes and categorized the questions under these themes. The other two researchers independently commented on the validity and definitions given by the first researcher for each theme. In Table 1, we present the different themes which emerged during our analysis along with a short description of the theme. At this stage, the definitions became more precise, e.g., 'Life Situation' was described as day-to-day life of the users. Themes **I** (Marketing Team) and **J** (Sales Team) could be seen as subthemes of **D** (Work Life). However, since the questions of these sub-themes were very specific to their respective professions, we decided to represent them as separate themes as well in the current work.

Since personas are widely used in different fields such as sales, marketing, and e-commerce [21–25], the themes in Table 1 reflect the questions which can be used in these different fields for creating personas specific to the fields. Thus, our work can have a wide impact on not just design practice but also other professions.

---

[1] https://www.dropbox.com/s/uwshanbm674a5ui/supporting%20material_INTERACT.xlsx? dl=0.

Table 1. Themes for the interview questions.

ID	Theme	This theme contains questions about…	N (%)
A	Demographics	The most basic questions that form the basis of the persona, such as age, gender, family size, income, occupation, race, religion, education, and so on	9 (3.3%)
B	Life Situation	The general day-to-day life of the users. As well as offering insights into their personality, these questions reveal what products and services users are likely to be interested in and what would be useful to them	79 (28.6%)
C	School Life	The lives of users during their school days. What subjects did they take, what extracurricular activities they participated in, what kind of school they attended and so on	18 (6.5%)
D	Work Life	The users' working life such as their current job, position, career goals, obstacles, and solutions in their job	72 (26.1%)
E	Decision Making	The users' process of making decisions. This involves, e.g., goals to be achieved, problem awareness, evaluation of options, decision-making styles, and post-decision behavior	20 (7.2%)
F	Information Sources	The channels users obtain information from. Do they go online, prefer to learn in-person, or rely on newspapers and magazines? If they are online learners, do they visit social networks? Google? Which sources do they trust the most -- friends, family, coworkers, or industry experts?	20 (7.2%)
G	Consumer Habits	The habits and preferences of the users when buying products and making other consumption choices	36 (13.0%)
H	Pain Points	The problems faced by the users and how organizations can see and find solutions to these problems	11 (4.0%)
I	Marketing Team	Information related to the work of marketing team members	7 (2.5%)
J	Sales Team	Information related to the work of sales team members	4 (1.4%)
NOTE: N indicates the number of questions in the theme			N = 276

3.3 Clustering the Questions

The number of questions in Table 1 seem excessive – it does not appear realistic to ask 79 questions from a user relating to their life situation. Therefore, we need to trim

down the list. This was accomplished by setting a maximum number of questions to ask under each theme. This way, the total number of questions will not exceed the maximum number multiplied by the number of themes selected for a specific use case.

We set the maximum to five questions. For example, if five themes were selected, the maximum number of questions would be $5 \times 5 = 25$ questions, which is reasonable number of questions for a user interview [26, 27].

Next, we will need to decide how to choose the representative five questions from each theme. For this study, we opted for using Natural Language Processing (NLP). Specifically, we used *Roberta-Large*[2], a transformer-based model. In brief, the transformer enables us to transform the questions into numerical format in a 1024-dimensional space where each dimension describes a position in the vector space relative to other questions. The NLP model computes the semantic similarity between each question under each theme – the meaning of two questions is similar when the numerical distance is low, and dissimilar when the distance is high [28].

To avoid a situation in which the very similar questions would appear among the Top-5 questions of a theme, we apply clustering on the word embeddings given by the transformer model. By creating five clusters from each theme (apart from 'Sales Information,' which only contains four questions, all of which will be selected), we can ensure that the chosen questions are separate from the other questions.

4 Results

4.1 Use Case Definition

In this part, we address the question, "*How to use the themes to select the final interview questions?*". For this, we created three fictitious scenarios to demonstrate how the themes can be used for persona information design:

- **Scenario 1:** A startup company wants to develop a software product for enterprise users. (Use case: *product development*)
- **Scenario 2:** A corporation wants to understand how their salespeople work and how to help them in their daily work. (Use case: *employee support*)
- **Scenario 3:** Marketers want to understand their potential target market in order to better market products for them. (Use case: *marketing*)

We selected the themes for each scenario (see Table 2) by considering the typical information needs of stakeholders facing the decision-making scenario. For the first scenario, as the target users are enterprise users, 'Work Life' is a relevant theme. Other relevant themes include 'Decision Making' and 'Pain Points,' as these themes help to understand the target users' needs and thinking. For the second scenario, important themes are 'Life Situation' (provides background information about the persons), 'Work Life' (informs about how they see their work role), 'Pain Points' (what concerns they have), and 'Sales Team' (specific questions about the sales profession). For the third scenario, the chosen themes include 'Demographics' (important for a marketing use

---

[2] https://huggingface.co/roberta-large.

cases), 'Decision Making' (to understand how the target market thinks), 'Information Sources' (to understand where they get their information from), and 'Consumer Habits' (to understand consumer behavior). Since the first scenario has three themes, the number of questions will be $3 \times 5 = 15$. The second scenario has four themes, so the number of questions will be $4 \times 5 = 20$. The third scenario also has four themes, but 'Sales Team' only contains four questions, so there will be a total of 19 questions.

Table 2. Identification of interview question themes based on the example use case scenarios.

	Scenario 1	Scenario 2	Scenario 3
Demographics			x
Life situation		x	
School life			
Work life	x	x	
Decision making	x		x
Information Sources			x
Consumer habits			x
Pain points	x	x	
Marketing team			
Sales team		x	

4.2 The Final Questions

Table 3 shows the final questions for the three persona scenarios. As described in Sect. 3.3, a representative question was selected from each selected them based on the similarity scores obtained using NLP. The similarity scores for each question are provided in the Supplementary Material[3].

Some fine-tuning was conducted for the final presentation. The questions that referred to 'they' instead of 'you' were changed to 'you' for consistency (e.g., "What job are they currently doing?" → "What job are you currently doing?"). If the top question in a theme was a follow-up from another question, we skipped it and selected the question with the next highest similarity score. We also skipped personal and potentially sensitive questions regarding income ("How much are they worth?", "How much do they earn?") and religion ("Were you raised in a religious household? If yes, which religion were you raised under?") as not everyone might be comfortable in answering these. Since Cluster 4 of 'Life Situation' only contained questions about religion, we chose instead two questions from the largest cluster (Cluster 3).

---

[3] https://www.dropbox.com/s/uwshanbm674a5ui/supporting%20material_INTERACT.xlsx?dl=0.

Table 3. Questions inferred by mapping persona information themes with scenarios and then using NLP to infer representative questions from each theme.

Scenario 1	Scenario 2	Scenario 3
Work Life	Life Situation	Demographics
What job are you currently doing? What is a typical workday like for you? What's important to you and what's driving the change? What kind of customer or user information do you need? How often do you buy high ticket items? (for work)	Tell us about your family life What hobbies (if any) do you have? Do you regularly go on vacations? What type of indulgent or luxurious purchases do you make? Do you tend to break or follow rules?	What is your name? Are you male or female? Which city were you born in? What is your current occupation? What is your income?
Decision Making	Work Life	Decision Making
What are the goals you're trying to achieve? Who do you consult with and trust for advice and information? What is your decision-making process when planning on buying [your product/service]? (If a customer) Why did you choose [your company] over another company? (If not a customer) Why did you choose X Company over us? What is most important to you when selecting a vendor?	What job are you currently doing? What is a typical workday like for you? What's important to you and what's driving the change? What kind of customer or user information do you need? How often do you buy high ticket items? (for work)	What are the goals you're trying to achieve? Who do you consult with and trust for advice and information? What is your decision-making process when planning on buying [your product/service]? (If a customer) Why did you choose [your company] over another company? (If not a customer) Why did you choose X Company over us? What is most important to you when selecting a vendor?
Pain Points	Pain Points	Information Sources
What is the most frustrating part of your day? What is the worst customer service experience you've ever had? What do you enjoy most? What makes you nervous? What is your least favorite part of your job?	What is the most frustrating part of your day? What is the worst customer service experience you've ever had? What do you enjoy most? What makes you nervous? What is your least favorite part of your job?	How adept are you at using technology? Do you use any social media websites? Which search engine do you use the most? What's your process for finding something online? Who do you ask for product/service recommendations? What websites or publications do you read regularly?

Table 3. (*continued*)

Scenario 1	Scenario 2	Scenario 3
	Sales Team	Consumer Habits
	What types of customers do you typically meet? Why do different types of customers typically make a purchase? What reasons do customers cite for selecting your business over a competitor? What are the most common objections you hear?	Think back on a recent purchase. How did you research the purchase? What factors were most important to you in evaluating your options? What doubts did you have? Did anyone help you make the final decision? What's your preferred method of communication? Emailing, texting, using an app (such as WhatsApp), or do you prefer to pick up the phone? Do you shop online? Are you willing to make the purchase by alternative means, or is there only one means by which you are happy or able to buy? What could we do to reach more people just like you?

4.3 Evaluation and Guidance for Application

Even though we have tried to be exhaustive in the questions for each scenario, some questions may or may not make sense to all researchers. For instance, Question 5 (the high-ticket item) does not seem directly relevant in Scenario 1, but other questions do seem to make sense. For Scenario 3, the two first questions can be considered optional, as this information may either not be pertinent (name) or can be observed in the interview situation (gender). Overall, a best practice is to conduct a manual 'sanity check' for the questions before conducting the interviews. This is because an "AI-generated" approach that uses algorithms might miss specific contextual and cultural nuances important for a given use case [29, 30].

5 Discussion

5.1 Research Contribution

The current work can help both practitioners and researchers to find inspiration and justification when creating or selecting their persona interview questions. Make all the questions – original, deduplicated, and final lists – publicly available. We hope this inspires others to make their interview questions more transparent, and to build upon this effort to support the research community.

Based on the use case, persona creators can select different combinations of persona information themes. For example, picking A+B+F (see IDs in Table 1) that combines demographics, life situation, and information sources might yield very different personas

than E+H (decision making and pain points). By combining the different themes, one can create alternative "lenses" of user understanding. Future development could include an online tool that recommends persona interview questions based on the user's selection of themes. Currently, the same can be accomplished by sorting the spreadsheet provided in the Supplementary Material[4] for theme and cluster, and then selecting the questions with the highest similarity score.

5.2 Limitations, Future Research, and Theorization

Whilst we focus on three persona scenarios, there are many more use cases for personas where the approach we suggested could be tested. Moreover, it would be interesting to see other researchers' take on how to organize the themes into original persona types. In this regard, one idea we offer is the concept of *instrumental* and *intrinsic* personas. It can be said that when personas are instrumental, they contain information for a specific goal. While this is compatible with Cooper's idea of goal-oriented design [1], there are other, alternative ways of seeing personas, most prominently the idea that personas are immersive portrayals of users that have value in unexpected ways.

Therefore, the question of "what is a complete persona profile?" can be addressed in different ways – the goal-oriented approach postulate that a persona profile is complete when it contains all the necessary information that the persona user needs [13]. In turn, the intrinsic view postulates that defining the user's information needs a prior may not be feasible or even possible, resulting in either the presumption that "completeness" is impossible, or at least it is understood much more broadly than one may think.

Finally, the persona technique may benefit from different data sources [31] and interviews are only one type of data. Therefore, creating personas using only interviews may not result in representative personas.

6 Conclusion

Interviews are a popular data collection technique for persona creation. Literature offers little guidance on what questions to include when interviewing users. In turn, the blogosphere offers hundreds, if not thousands of possible questions. Those that are new to personas, are therefore pinned between too little and too much information. We demonstrated an approach of collecting a large base of questions, manually categorizing them to themes, and then applying NLP techniques to create a manageable list of questions that corresponds to specific use cases. This approach can be applied by researchers and practitioners alike when choosing interview questions for persona creation.

References

1. Cooper, A.: The Inmates are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity. Sams - Pearson Education, Indianapolis, IN (1999)

---

[4] ibid.

2. Adlin, T., Pruitt, J.: The Essential Persona Lifecycle: Your Guide to Building and Using Personas. Morgan Kaufmann Publishers Inc., San Francisco (2010)

3. Friess, E.: Personas and decision making in the design process: an ethnographic case study. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1209–1218 (2012)

4. Nielsen, L., Jung, S., An, J., Salminen, J., Kwak, H., Jansen, B.J.: Who are your users?: comparing media professionals' preconception of users to data-driven personas. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction, Brisbane, Queensland, Australia, pp. 602–606. ACM (2017). https://doi.org/10.1145/3152771.3156178

5. Aoyama, M.: Persona-scenario-goal methodology for user-centered requirements engineering. In: Proceedings of the 15th IEEE International Requirements Engineering Conference (RE 2007), Delhi, India, pp. 185–194 (2007). https://doi.org/10.1109/RE.2007.50

6. Nielsen, L.: Personas - User Focused Design. Springer, New York (2019)

7. An, J., Kwak, H., Salminen, J., Jung, S., Jansen, B.J.: Imaginary people representing real numbers: generating personas from online social media data. ACM Trans. Web (TWEB) **12**, 27 (2018). https://doi.org/10.1145/3265986

8. McGinn, J.J., Kotamraju, N.: Data-driven persona development. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, pp. 1521–1524. ACM (2008). https://doi.org/10.1145/1357054.1357292

9. Salminen, J., Guan, K., Jung, S., Chowdhury, S.A., Jansen, B.J.: A literature review of quantitative persona creation. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI 2020, Honolulu, Hawaii, USA, pp. 1–14. ACM (2020). https://doi.org/10.1145/3313831.3376502

10. Mulder, S., Yaar, Z.: The User is Always Right: A Practical Guide to Creating and Using Personas for the Web. New Riders (2006)

11. Salminen, J., Guan, K., Nielsen, L., Jung, S.-G., Jansen, B.J.: A template for data-driven personas: analyzing 31 quantitatively oriented persona profiles. In: Yamamoto, S., Mori, H. (eds.) HCII 2020. LNCS, vol. 12184, pp. 125–144. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50020-7_8

12. Nielsen, L., Hansen, K.S., Stage, J., Billestrup, J.: A template for design personas: analysis of 47 persona descriptions from danish industries and organizations. Int. J. Sociotechnol. Knowl. Dev. **7**, 45–61 (2015). https://doi.org/10.4018/ijskd.2015010104

13. Salminen, J., Santos, J.M., Kwak, H., An, J., Jung, S., Jansen, B.J.: Persona perception scale: development and exploratory validation of an instrument for evaluating individuals' perceptions of personas. Int. J. Hum Comput Stud. **141**, 102437 (2020). https://doi.org/10.1016/j.ijhcs.2020.102437

14. DiMicco, J.M., Mann, N.: User research to inform product design: turning failure into small successes. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 872–879 (2016)

15. Gao, Y., Li, X., Lin, Y.-H., Liu, X., Pang, L.: Nuwa: enhancing the pregnancy experience for expectant parents. In: Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI EA 2014, Toronto, Ontario, Canada, pp. 257–262. ACM Press (2014). https://doi.org/10.1145/2559206.2580928

16. Chang, Y.-J., Li, L., Chou, S.-H., Liu, M.-C., Ruan, S.: Xpress: crowdsourcing native speakers to learn colloquial expressions in a second language. In: Extended Abstracts on Human Factors in Computing Systems, CHI 2013, pp. 2555–2560 (2013)

17. Falk, K., Kamara, A.K., Braathen, E.P., Helle, K., Moe, P.T., Kokkula, S.: Digitizing the maintenance documentation; a system of systems in oil and gas industry. In: 2020 IEEE 15th International Conference of System of Systems Engineering (SoSE), pp. 493–500. IEEE (2020)

18. Antle, A.: Supporting children's emotional expression and exploration in online environments. In: Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community, pp. 97–104 (2004)
19. Sharbatdar, N., Lamine, Y., Milord, B., Morency, C., Cheng, J.: Capturing the practices, challenges, and needs of transportation decision-makers. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–7 (2020)
20. Richardson, R., Kramer, E.H.: Abduction as the type of inference that characterizes the development of a grounded theory. Qual. Res. **6**, 497–513 (2006)
21. Al-Qirim, N.: Personas of e-commerce adoption in small businesses in New Zealand. J. Electron. Commer. Organ. (JECO) **4**, 18–45 (2006)
22. Duda, S.: Personas—who owns them. In: von Gizycki, V., Elias, C.A. (eds.) Omnichannel Branding: Digitalisierung als Basis erlebnis- und beziehungsorientierter Markenführung, pp. 173–191. Springer, Wiesbaden (2018). https://doi.org/10.1007/978-3-658-21450-0_8
23. Li, K., Deolalikar, V., Pradhan, N.: Mining lifestyle personas at scale in e-commerce. In: 2015 IEEE International Conference on Big Data (Big Data), pp. 1254–1261. IEEE (2015)
24. Salminen, J., Jansen, B.J., An, J., Kwak, H., Jung, S.: Are personas done? Evaluating their usefulness in the age of digital analytics. Persona Stud. **4**, 47–65 (2018). https://doi.org/10.21153/psj2018vol4no2art737
25. Thoma, V., Williams, B.: Developing and validating personas in e-commerce: a heuristic approach. In: Gross, T., et al. (eds.) INTERACT 2009. LNCS, vol. 5727, pp. 524–527. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03658-3_56
26. Baker, S.E., Edwards, R.: How Many Qualitative Interviews is Enough. National Centre for Research Methods, UK (2012)
27. Bolderston, A.: Conducting a research interview. J. Med. Imaging Radiat. Sci. **43**, 66–76 (2012). https://doi.org/10.1016/j.jmir.2011.12.002
28. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. In: Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 26, pp. 3111–3119. Curran Associates, Inc. (2013)
29. Amershi, S., et al.: Guidelines for human-AI interaction. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2019)
30. Kocielnik, R., Amershi, S., Bennett, P.N.: Will you accept an imperfect AI? Exploring designs for adjusting end-user expectations of AI systems. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2019)
31. Jansen, B., Salminen, J., Jung, S., Guan, K.: Data-Driven Personas. Morgan & Claypool Publishers (2021)