*Innovation in Applied Nursing Informatics*
*G. Strudwick et al. (Eds.)*

*doi:10.3233/SHTI240314*

# Utilizing Digital Twin to Create Personas Representing Ovarian Cancer Patients and Their Families

Youjia Wang[a1], Khushboo Thaker[b], Vivian Hui[c], Peter Brusilovsky[b], Daqing He[b], Heidi Donovan[a], Young Ji Lee[a]

*a. School of Nursing, University of Pittsburgh, PA, USA*
*b. School of Computing and Information, University of Pittsburgh, PA, USA*
*c. Center for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong*

ORCiD ID: Youjia Wang https://orcid.org/0009-0008-1534-5753

**Abstract**. OvCa patients and caregivers perceived challenges in online health information seeking. The HELPeR recommendation system utilized digital twins to create personas reflecting real-world OvCa patients and caregivers. The aim of this study was to describe the creation of digital twins and demonstrate their use cases in the study. Digital twins of OvCa patients and caregivers were created by triangulating multiple sources, including online cancer forums, direct interviews with patients and caregivers, domain expert input, and clinical notes. 10 personas were created for both OvCa patients and caregivers who had a variety of cancer trajectories and information interests. These digital twins present a potential solution for training artificial intelligence models at the initial phase when there is a scarcity of user information.

**Keywords**. Ovarian cancer, online health information seeking, recommendation system

## 1. Introduction

Ovarian cancer (OvCa) patients and their informal caregivers (e.g., friends, family) constantly have evolving information needs throughout the cancer trajectory.[1] Online health information becomes one of the various resources from which they seek health-related information. However, OvCa patients and caregivers perceive challenges in seeking online health information, including a lack of OvCa-related knowledge, poor quality of online information, and the inherent characteristics of OvCa, such as its high complexity and aggressive disease progression.[1] To address these challenges, we created the Health E-Librarian with Personalized Recommendations (HELPeR) program, a personalized, adaptive online health information recommender system. A typical challenge for recommender engines is the 'cold start' problem, which arises when the engine lacks information about new users to make personalized recommendations.[2] To resolve the cold-start issue and reduce researchers' efforts in recruiting participants with a rare disease,[3] we adopted the digital twin (DT) approach. This approach, which

[1] Corresponding Author: Youjia Wang (yow14@pitt.edu)

involves a digital replication of living or non-living physical entities, has been increasingly utilized in healthcare.[4] It encompasses a wide range of elements, including patient profiles, and is designed to mirror its real-world counterparts in real-time.[5] As this is a new approach, we aim to describe the creation of digital twins and demonstrate their use case study.

## 2. Methods

We created digital twins of OvCa patients and caregivers by triangulating multiple sources, including online cancer forums, direct interviews with patients and caregivers, domain expert input, and clinical notes. Firstly, we crawled postings from any user who had made two or more postings in the forum. We then manually annotated the contents using our custom annotation tool (Fig 1). Annotations covered various aspects, such as disease trajectory, cancer stage, information needs, interests, user types (patient or caregiver), age, caregiving status, and socio-demographic information (e.g., marital status, insurance). This information was categorized using K-means clustering and simple matching. These clusters were then triangulated with interview transcriptions from patients and caregivers (n=20) and clinical notes from the CARE Caregiver Center at Magee Women's Hospital. Nurse experts (n=3) validated and enriched the extracted features of OvCa patients and caregivers based on their clinical experience. Since the validation of the personas was conducted in a meeting with nurse experts, we did not perform a separate measurement of agreement between these experts. All developed personas were represented with realistic scenarios and accompanying images of fictitious individuals.

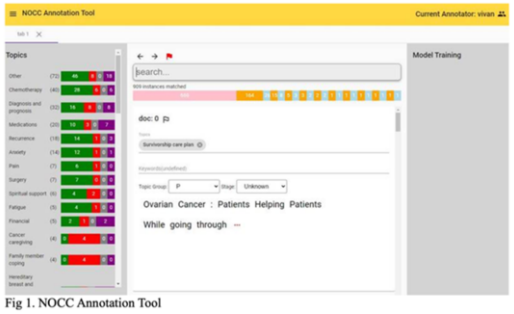

Fig 1. NOCC Annotation Tool

## 3. Results

A total of ten personas were developed for the HELPeR program, comprising seven OvCa patients and three informal caregivers (Fig 2). The age range of the patient personas varies from 42 to 65, encompassing a diversity of professional statuses, from those currently working or on medical leave to retirees. The majority of these patients (n=6) are in advanced stages (stage III or IV) of cancer, with their trajectories spanning initial diagnosis, survivorship, remission, to progression or recurrence. The needs of these patients are multifaceted, including symptom management (e.g., peripheral neuropathy, early menopause, chemo brain, chronic pain), treatment information (e.g., general treatment, hormonal maintenance therapy), practical needs (e.g., financial assistance, managing medical bills and insurance, support services),

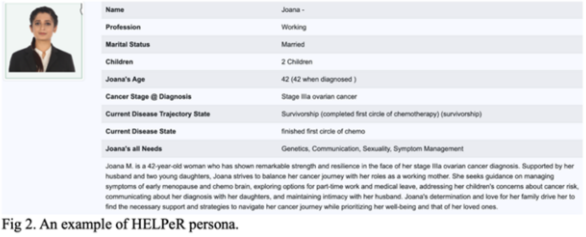

Fig 2. An example of HELPeR persona.

communication needs (e.g., discussing cancer diagnosis with children), and personal needs (e.g., genetics, sexuality/intimacy).
The caregivers, ranging in age from 45 to 79, are primarily husbands or daughters of the patients. They are caring for individuals in advanced stages of cancer, often dealing with recurrence, progression, or end-of-life situations. Their needs encompass areas such as communication, hospice care/end-of-life care, advance care planning, self-care, and spiritual support. All these personas are currently being utilized to train the recommender algorithm.

## 4. Conclusions

The HELPeR personas successfully mirrored the real-life situations of most OvCa patients and caregivers, capturing a diverse range of cancer trajectories and unique information interests. These digital twins present a potential solution for training artificial intelligence models at the initial phase when there is a scarcity of user information. The accuracy and efficiency with which HELPeR recommends articles based on these personas will be a significant indicator of its ability to recommend pertinent information to actual OvCa patients and caregivers. Currently, we have not diversified the cultural aspects of the personas; the focus has been on clinical aspects. Future work will aim to create more diverse personas, including variations in ages, races, and those in the early stages of the disease.

**Acknowledgement**
The project was supported by the National Institutes of Health R01 LM01308 (MPI: Lee and He).

## References

[1] Chi Y, Hui V, Kunsak H, et al. Challenges of ovarian cancer patient and caregiver online health information seeking. *Proceedings of the Association for Information Science and Technology*. 2021;58(1):688-690. doi:10.1002/pra2.530
[2] Lika B, Kolomvatsos K, Hadjiefthymiades S. Facing the cold start problem in Recommender Systems. *Expert Systems with Applications*. 2014;41(4):2065-2073. doi:10.1016/j.eswa.2013.09.005
[3] Ovarian cancer - about the disease. Genetic and Rare Diseases Information Center. Accessed December 11, 2023. https://rarediseases.info.nih.gov/diseases/7295/ovarian-cancer.
[4] El Saddik A. Digital twins: the convergence of multimedia technologies. *IEEE Multimed* 2018; 25: 87–92.
[5] Albanians L. A comprehensive guide to implementing digital twins in Healthcare. Medium. September 25, 2023.https://medium.com/@Larisa10/a-comprehensive-guide-to-implementing-digital-twins-in-healthcare-dfd832a6f72#:~:text=These%20Digital%20Twins%20represent%20individual,treatment%20plans%20and%20remote%20monitoring.