

Viewpoint

# Harnessing Machine Learning to Personalize Web-Based Health Care Content

Ahmad Guni<sup>1,2</sup>, BSc, MBBS; Pasha Normahani<sup>1,2</sup>, BSc, MBBS, MSc, MRCS; Alun Davies<sup>1,2</sup>, MA, DM, DSc, FRCS, FHEA, FACPh, FLSW; Usman Jaffer<sup>1,2</sup>, BSc, MSc, PhD, FRCS

<sup>1</sup>Department of Surgery and Cancer, Imperial College London, London, United Kingdom

<sup>2</sup>Imperial Vascular Unit, Imperial College Healthcare NHS Trust, London, United Kingdom

**Corresponding Author:**

Ahmad Guni, BSc, MBBS

Department of Surgery and Cancer

Imperial College London

Exhibition Road

London, SW7 2AZ

United Kingdom

Phone: 44 7803434969

Email: [ahmad.guni@nhs.net](mailto:ahmad.guni@nhs.net)

## Abstract

Web-based health care content has emerged as a primary source for patients to access health information without direct guidance from health care providers. The benefit of this approach is dependent on the ability of patients to access engaging high-quality information, but significant variability in the quality of web-based information often forces patients to navigate large quantities of inaccurate, incomplete, irrelevant, or inaccessible content. Personalization positions the patient at the center of health care models by considering their needs, preferences, goals, and values. However, the traditional methods used thus far in health care to determine the factors of high-quality content for a particular user are insufficient. Machine learning (ML) uses algorithms to process and uncover patterns within large volumes of data to develop predictive models that automatically improve over time. The health care sector has lagged behind other industries in implementing ML to analyze user and content features, which can automate personalized content recommendations on a mass scale. With the advent of *big data* in health care, which builds comprehensive patient profiles drawn from several disparate sources, ML can be used to integrate structured and unstructured data from users and content to deliver content that is predicted to be effective and engaging for patients. This enables patients to engage in their health and support education, self-management, and positive behavior change as well as to enhance clinical outcomes.

(*J Med Internet Res* 2021;23(10):e25497) doi: [10.2196/25497](https://doi.org/10.2196/25497)

**KEYWORDS**

internet; online health information; personalized content; patient education; machine learning

## Introduction

The internet is a key medium in the consumption of health care-related content. Two-thirds of internet users in the United Kingdom and the United States access health-related information on the internet [1-3]. Furthermore, patients are increasingly motivated and able to participate in developing this growing repository of information by sharing their lived experiences [4]. Health care professionals also consume as well as create and share web-based health care information [5].

A vast array of web-based content types and delivery media and channels are available, including videos, webpages,

podcasts, images, online discussion groups and communities, and social media [6-10]. A systematic review reporting on the assessment of web-based content quality identified key domains, including accuracy, completeness, accessibility, presentation, and design, which were important overall in determining how useful and engaging content was for patients [11]. However, there is considerable variability in the quality of such content [12]. Berland et al [13] demonstrated that using search engines for common health conditions retrieved relevant content in only one out of five searches, suggesting that patients are likely to come across irrelevant content when seeking information about health. Moreover, only half of the topics that physicians thought

important to convey were accurately and appropriately covered [13].

The concept of content personalization is a powerful approach that addresses the previously described features of quality by presenting the user with relevant information that is both appropriate and engaging. A more engaged patient is more likely to understand information about their health, partake in healthy behaviors, and adhere to treatment, leading to better health outcomes [14].

Machine learning (ML) is a subset of artificial intelligence that uses algorithms to study patterns in data and develop models that improve predictions about the data over time through supervised learning, unsupervised learning, or reinforcement learning [15]. Many industries use ML techniques to analyze accrued *big data* to personalize content for users [16]. The health care sector may be well served by considering these advances in other industries to personalize experiences for people seeking health care content. ML-assisted personalization can be considered for both large groups and populations or for individuals.

In this review, we aim to first outline why the health care sector should recognize the importance of personalizing content (*Why Personalizing Web-Based Health Care Content Is Important*). We then explore the current landscape of content personalization (including ML and non-ML) both within and outside health care (*Content Personalization*). Finally, we discuss practical applications of personalization in health care, outline a model that demonstrates how ML can personalize web-based content, and consider the anticipated benefits and drawbacks (*Potential Lessons to Learn for Health Care*).

## Why Personalizing Web-Based Health Care Content Is Important

There has been an increased focus on empowering patients to engage with their own health. The delivery of information to patients has been recognized as a tenet of health care policy, resulting in almost universally positive outcomes for patients, health care staff, and communities [17]. The UK National Health Service Five Year Forward View outlined the need to facilitate patient activation by improving access to information, supporting self-management, and increasing patient control over the care they receive, with particular emphasis placed on harnessing digital technology [18]. This aligns with the *patient-centered* model [19], which improves patient satisfaction, quality of life, and quality of care provided [20]. Personalization facilitates the patient-centered model by delivering health care content that accounts for the preferences and needs of individual patients. The proliferation of easily accessible web-based content provides an opportunity to enable *patient-centered* information delivery at scale.

A randomized controlled trial of the provision of computer-based information to cancer patients reported that patients preferred to receive personalized information (based on their medical records) as opposed to generalized information [21]. They were more likely to share these resources with family members, and additionally, this approach was associated with

a reduction in anxiety levels. A similar effect was demonstrated with personalization of booklets [22] and tailored information packs [23].

It is well established that the health care content needs to vary between different patients and also change over time. Uncertainty, the inability to determine the meaning of illness-related events, has been shown to have a deleterious effect on patient experience and outcomes [24,25]; therefore, timely and accurate delivery of information is important to address information needs. However, patients' information needs vary according to stage of disease, stage of patient journey, age, previous experiences, and coping styles [26]. A blanket *one-size-fits-all* strategy for designing and delivering health care content is unlikely to be effective.

Another advantage of personalizing health care content is its potential to improve health-related choices. One of the principles of the *patient-centered* model is sharing responsibility for clinical decisions with patients (shared decision-making) [27]. Patient decision aids are evidence-based tools designed to assist in shared decision-making. They facilitate information exchange by helping patients understand the clinical conditions and the available options for treatment. They have been demonstrated to improve patient knowledge and facilitate decision-making that is more aligned with patient values and preferences [28,29].

A study on improving patient decision-making related to prostate cancer screening found that personalizing a patient decision aid based on a number of factors that patients considered important (eg, survival, unnecessary biopsy, overdiagnosis, quality of life, burden of treatment, and burden on caregivers) improved patient opinion on screening and the quality of their decision [30]. Decision quality was assessed using an instrument that allows patients to self-rate and weigh separate elements of decision quality, including the perceived clarity of options provided, relative importance and likelihood of possible outcomes, trust toward the information delivered, support received throughout the decision-making process, sense of control over the decision, and commitment toward acting on the decision [31].

It is increasingly recognized that delivering health information without consideration for personalization and the relevance of content experienced limits the potential to change health behavior [32,33]. A meta-analysis on behavior choices from 40 web-based interventions, which used personalized strategies including interactive multimedia content, tailored feedback, discussion groups, and personalized management plans, showed a positive impact on behavior outcomes related to smoking cessation, alcoholism, physical activity, diet, and chronic disease management [34]. These findings are corroborated by other meta-analyses evaluating tailored content for similar health-related behavior outcomes [35-38]. However, given the significant heterogeneity in the intervention modality, design, and features, it is challenging to identify the specific factors that are most associated with behavior change.

With a greater understanding of these factors, there is significant scope to integrate personalized content into both large-scale public health initiatives as well as individual treatment plans to encourage self-management, adherence to treatment, and positive lifestyle changes.

## Content Personalization

### Content Personalization in Health Care—Current State

The paths patients take to encounter web-based content can be described by a number of discrete patient journeys. First, patients can independently find web-based information using internet search engines. Although this offers patients a plethora of information, quality (as previously discussed in the *Introduction* section) is variable [12]. Without strict content moderation and regulation, patients may struggle to parse out factual and relevant content, instead relying on content that is superficially engaging (*clickbait*) or appears credible. Furthermore, subtle differences in search terms can significantly alter the quality of the retrieved information [39].

Health care organizations and services hold repositories of quality-controlled content and can serve as gateway sites for other similar websites [40]. These provide credible and accurate information but hold limited quantities of content and may not be directly relevant to every patient. Health care professionals can assess individual information-seeking needs during consultations and refer patients to high-quality and engaging content [5]. However, this solution lacks scalability because most web-based health care searching encounters are *unsupervised* by health care providers. Limitations on how patients access health information can be addressed with content personalization, which mandates an understanding of what factors may be important in personalizing content.

Patients' information needs are affected by several factors that may influence how patients respond to web-based content, as discussed in *Why Personalizing Web-Based Health Care Content Is Important* section. For example, in the context of age, older patients often report difficulty in accessing useful web-based content because of complex website layouts, lack of navigational aids or instructional tools, and too much information being presented [41]. Younger patients may be more prone to uncertainty and worry about their health, resulting in information-seeking behavior [42]. A study that allowed cancer patients to self-tailor web-based educational content based on text, visual, and audio-visual modes demonstrated increased satisfaction among younger patients in comparison with nontailored content [43].

With regard to factors that affect the decision to select or reject web-based content, a study found several content and design features that influenced whether patients trust web-based information related to hormone replacement therapy [44]. An initial poor impression of design factors—including inappropriate website name, complex layout, poor navigation aids, dull design, small print, and excessive text—constituted 94% of cited reasons for rejection. Content features were then comparatively more important in selecting trustworthy websites. This consisted of informative content, accessible explanations, illustrations, breadth of topics covered, unbiased information, age-related information, clear language, discussion groups, and a frequently answered questions section. Source factors were also key, such as explicit author or organization credibility and authors with similar social identities.

Other studies have evaluated the design and content factors that influence patients' engagement in web-based videos, particularly on the video streaming website YouTube, which is one of the most popular websites with over 2 billion daily views [45]. These include educational resources on a range of medical topics for both patients and health care professionals [46-51]. These studies also assessed the quality of content uploaded on YouTube, which is not strictly regulated and is liable to misinformation [7]. However, the correlation between engagement and quality of content is conflicting [7], suggesting that other factors are important for gaining user attention in educational resources.

An analysis of 390 scientific communication videos on YouTube found that user-generated content, videos with regular presenters, and rapidly paced videos were more engaging than their counterparts [52].

Similarly, another study concluded that patient experience videos were more popular than videos created by health care professionals, as assessed by the video power index [53]. The video power index is an innovative tool that measures video performance by assessing its effectiveness on all platforms, comparing it with industry leaders, and aiding strategies to engage target audiences [54]. In terms of webpage content, Finnegan et al [55] found that engaging content categories were first-person narrative articles, articles that answer questions posed by readers, and articles with videos embedded in the webpage. These are all potential factors that can be considered when personalizing video content toward patients.

Sorice et al [56] examined patients' preferred social media content related to plastic surgery on six social media platforms (Facebook, Instagram, Pinterest, Snapchat, Twitter, and YouTube) [56]. Patients used Facebook and YouTube as the most favored posts relating to before and after photographs and the surgery practice information. Second, the content that engaged plastic surgeons and patients differed. The authors concluded that this information should guide the web-based activity of plastic surgeons to effectively target the desired patients.

A systematic review evaluating factors associated with engaging web-based content revealed the following key categories: textual information, discussion boards and web-based groups, video content, visual or pictographs, device accessibility, stage of patient journey, credibility, and completeness of information [57]. A framework was developed for each category describing the factors that should be considered when designing an effective content. Evidently, the manner in which users engage with health care content is influenced by both design and content factors, many of which are likely not yet identified.

### Content Personalization Outside Health Care—Current State

With increasing volumes of web-based data available for extraction, storage, and processing, ML is useful in improving the efficiency and accuracy of data processing models without human input. Its application spans a wide range of disciplines, including marketing, engineering, computer science, finance, bioinformatics, and health care. In the context of personalizing

health care content, ML applications may fall into the following categories: facilitating market segmentation, content analysis, and recommender systems.

In marketing, maximizing user—or customer—engagement is obviously a key driver. Customer segmentation and personalization of content in these segments in a competitive environment is easy to appreciate. Furthermore, 59% of customers believe that personalization influences their purchasing habits [58]. A study reporting over 30,000 campaigns by one company revealed that targeted campaigns resulted in greater customer retention, engagement, and conversion into active users compared with generic campaigns [59]. Audience segmentation for web-based marketing aims to split the customer population based on characteristic features (eg, demographic, psychographic, geographic, behavior, and product preference) [60]. Individual customer segments can be targeted with specific content and products predicted to elicit the most attention, resulting in sales and profits [61]. However, customer segmentation performed by human marketers is limited by the amount of data that can be amassed, analytical methods that can be used, and the number of conclusions drawn. ML using clustering techniques can process larger volumes of data and uncover complex patterns to draw more practical conclusions and create better-defined segments for targeting. Infamously, this approach can also be used to target groups with messages that may affect behavior, such as political elections [62], but is less likely to be a useful method to personalize health care content for individuals, as there will still be differences in the needs and preferences of individuals within segments.

Recommender systems are used by the entertainment, e-commerce, and marketing industry to personalize content discovery and information retrieval in the context of massive item repositories [63-66]. Established methods include collaborative filtering, which applies the behavior of similar users to suggest new items of interest; content-based methods, which analyze content similarities with previous user preferences to produce recommendations; and hybrid methods, which combine both. Although the research landscape has predominantly focused on collaborative filtering [67], increased interest has gathered around content-based filtering with techniques emerging to identify content features [68], including user-generated tags and reviews [69], and advances in video [70] and image [71] analysis capabilities.

As one of the largest platforms for creating and sharing content, the YouTube recommender system uses deep learning to generate and rank candidate videos by incorporating a rich set of user and video features, such as the user's history, context, and interaction with similar videos [72]. This facilitates access to a small set of engaging personalized content from an ever-increasing repository of videos. Other studies have demonstrated several content factors that can also influence personalization. For example, a study incorporated textual content features including video metadata and nontextual features consisting of audio, scenes, and motion to enhance personalized recommendations for videos; this was more accurate in effective personalized video recommendation from large video data sets (Netflix and MovieLens) over existing models that use single specific content features [73].

Social media recommender systems provide insights into how companies personalize other media content discovery for users. Instagram analyzes content that users have previously interacted with and uses natural language processing to identify similar accounts to recommend content that the user is likely to interact with on their *Explore* page [74]. In addition, content analysis of social media pages reveals several factors that also influence user engagement and may further refine content personalization. In a study on over 13,000 Instagram posts, using an image application programming interface (API) to extract visual features from posts, several creator-related, context, and content factors predicted user engagement [75]. In particular, images containing people, scenery, and emoticons associated with positive emotions engaged users more strongly. Other content features on Instagram that correlate with user attention are photos with faces [76] and filters enhancing warmth, exposure, and contrast [77]. An analysis using a natural language processing API on over 100,000 messages on Facebook found that emotional and philanthropic content enhances engagement, whereas informative content reduces engagement in isolation, but further invokes attention when combined with persuasive features [78].

Advances in recommender systems have further improved the personalized recommendations. For instance, movie recommender systems traditionally use higher semantic features (eg, tags, plot, genre, and actors) suggested by users or experts to personalize recommendations [79]. A recent work using a deep learning neural network found that extracting low-level stylistic features (eg, colors, texture, and lighting) outperformed traditional semantic-based methods in recommending content [70]. With developments in algorithmic approaches and deep learning [68], high- and low-level content features can be integrated to generate more personalized content recommendations.

Recently, open-source services that leverage ML have become available on commercial platforms with the Google Cloud Artificial Intelligence as a foremost example [80,81]. These services require minimal ML expertise and consist of custom models using AutoML and pretrained models, which include video intelligence API (analyze video metadata), natural language API (analyze text), vision API (image segmentation and classification), and speech API (transcribing audio). Similar platforms exist with Amazon Rekognition image and video analysis [82], Microsoft Azure video indexer, text analytics and personalizer [83], and IBM Watson video content analysis and natural language understanding [84]. Amazon's predictive user engagement service offers to improve user engagement by analyzing real-time activity to personalize recommendations and notifications for users [85,86]. The prospect of designing custom ML may have been prohibitive for many industries previously, but these open-source platforms provide an opportunity to adopt it into the mainstream of a variety of disciplines for large-scale data processing.

## Potential Lessons to Learn for Health Care

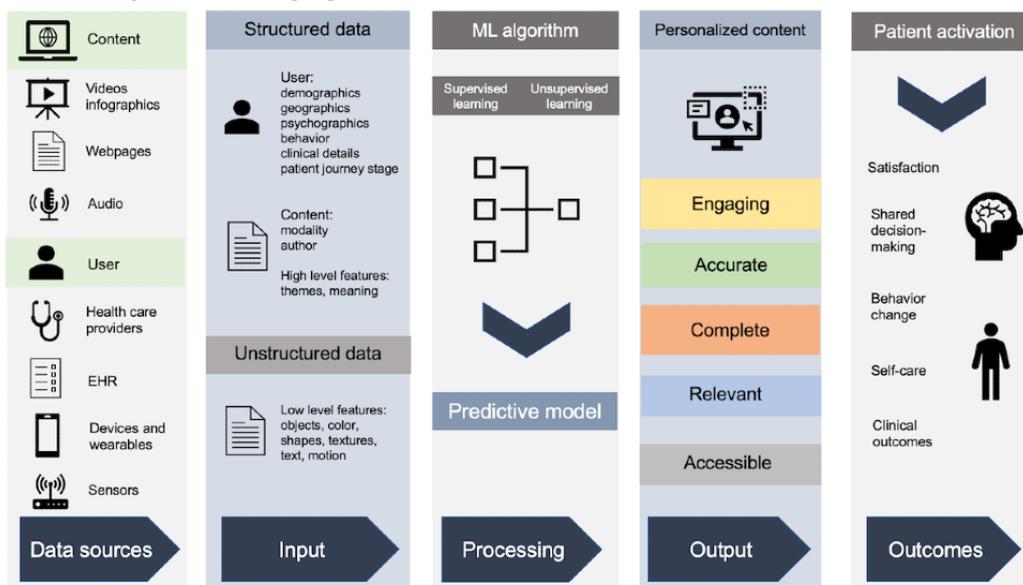
The previous sections described user segmentation, targeted advertisements, and personalization based on recommender systems using ML techniques. With the vast amount of web-based health care content readily accessible to patients, cross-disciplinary collaboration and the use of open-source platforms indicate that these techniques may be feasible. If this is achieved, the aim of personalizing web-based content and enhancing outcomes is possible. However, clinical studies and clinical applications related to this are sparse.

Big data in health care can transform the field of health marketing (an established concept in public health medicine), drawing principles from traditional marketing to create, communicate, and deliver information in a patient-centered manner [87]. This aims to identify population segments and market health care messages to them in terms of the segments that are likely to respond [88]. A systematic review of health marketing research identified a number of studies that used hierarchical and nonhierarchical clustering techniques to segment health consumers in unique ways [88]. However, the studies did not explore whether these segments were meaningful

(predictive segmentation) or whether personalized interventions affected outcomes. Furthermore, there was a reliance on rudimentary data such as survey, service, and basic clinical data, which limits the clustering process as opposed to truly *big* data. Although these strategies may have beneficial effects for groups of people, it is difficult to imagine their utility to individuals.

We propose a model that leverages ML algorithms to personalize content for an individual person (Figure 1). Health care big data consists of diverse data types, including clinical data, electronic patient records, biometrics, sensor-generated data, population data, social media posts, and webpages [16]. Electronic health records are accumulating data at an exponential rate. With the increasing use of medical devices, sensors, wearable technology, and social media, more personal data can be recorded [89]. These consist of potential sources of structured and unstructured data that may be fed into ML algorithms. Structured data include labeled user features such as demographics, geographics, psychographics, behavior, and clinical details, as well as content features consisting of modality, themes, and author information. Unstructured data, comprising 80% of all health care data [90], can be processed by video, image, and natural language processing APIs into structured formats [91]. ML algorithms using supervised and unsupervised learning can process these data to produce a predictive model for content personalization.

**Figure 1.** Suggested model using machine learning to personalize web-based health care content. EHR: electronic health record; ML: machine learning.



User features can be matched to content features (whether video, text, infographics, or audio) to create a model predicting which content is likely to be engaging to which people. Content features need not be limited to the content or design features identified in the *Content Personalization* section. Meta-level information encompassing object identification (colors, shapes, and texture), person or face identification, motion features, patterns, textual analysis, medical tags, higher semantic meaning, and significantly more may be extracted and analyzed. The content for patients can be created with these specific features in mind. Recommender systems could automatically predict other content that is useful and engaging to patients, conveying education that is likely to affect them.

Metrics related to view count, likes, shares, and positive comments have traditionally been used as an indicator of popularity, but they may only provide a superficial measure of engagement and fail to capture key outcomes for patients. Similarly, no single outcome metric is likely to be sufficient. Possible surrogate measures to consider include shared decision-making [27], patient satisfaction [92], objective clinical outcomes and symptoms [93], changes in attitude and behavior [94,95], and physiological signals [96]. These factors can aid in content personalization.

Harnessing data from personal digital devices such as wearables, phones, and computers has led to research into digital phenotyping and personal sensing, which refer to the analysis

of data streams from personal devices to build a human phenotype by identifying behaviors, traits, thoughts, and feelings [97,98]. This field has been adopted predominantly in psychiatry, where the objective identification of behavior patterns can aid in the diagnosis and stratification of mental health conditions, as well as their treatment (digital health interventions) [97]. In a recent study of internet-based cognitive behavioral therapy, ML was used to identify different behavior patterns among segments of patients, consisting of low engagers, late engagers, high engagers with rapid disengagement, and the highest engagers [99]. Each patient subtype was more likely to engage with different intervention tools (eg, core modules, goal-based activities, mood trackers, and mindfulness tools), leading to varying improvements in depression and anxiety symptoms. The authors concluded that this information could be used to tailor specific intervention types to different patient subtypes to improve engagement and adherence to treatment.

There are clear similarities between these digital health concepts and the proposed model to personalize web-based health care content. In particular, ML can be used to analyze data streams that include sensor measurements, user activity on personal devices, and user-generated content to identify individual behavior patterns. This can then be used to personalize interventions, of which personalized content could form a part of the intervention, or, at the very least, to inform patients about their health and engage them in making healthy behavior choices.

The successful implementation of big data and ML in personalizing web-based content requires the input and collaboration of several multidisciplinary stakeholders [100]. Health care professionals must produce accurate and engaging user-centered content, which is consumed by patients who can use recommender systems to discover related content and are also able to create content on their own. ML algorithms based on the model described in Figure 1 were designed by computer scientists and ML engineers and further optimized by several data streams provided by patients and health care organizations. There should be ongoing collaborative research between clinicians and computer scientists to take advantage of developments in ML, such as the use of deep learning.

However, current inadequacies in the digital infrastructure of health care systems can pose a significant challenge to this process. For example, as outlined in the UK government policy paper on their future digital strategy plan [101], patient data are often stored in disparate systems between different hospitals and health care settings that are unable to communicate with each other. One of the priorities should, therefore, be to create data standards that facilitate the interoperability of patient health

records, which would enable seamless access, storage, and processing at scale. It is promising that government agencies have already taken steps to outline frameworks to achieve secure access, interoperability, and sharing of health-related patient data [101,102].

Other drawbacks of big data and personalized health care must also be considered in addition to the benefits. Maintaining the privacy and security of sensitive patient data is paramount and poses significant challenges with the volume of data recorded from an increasing number of sources. No single legal or ethical framework covers all aspects of health information privacy [103]. Furthermore, many laws are outdated and insufficient for the current era of big data, which includes user-generated data (eg, wearables and sensors) and nonhealth information that can lead to health inferences (eg, social media habits) [104]. Therefore, governments and health care bodies must also act as key stakeholders to ensure that laws are updated to allow ML to be harnessed for the benefit of patients while maintaining privacy and security. This may necessitate the development of oversight agencies to strictly regulate the use of ML, as well as collaboration with cybersecurity experts [100]. The principles of consent in digital data research and use need to be established and will require input from governments, national data regulators, medical ethicists, legal experts, and, most importantly, patients [105].

There are several principles for maintaining private and secure data, including collecting data from trusted sources, encrypting and anonymizing stored data, maintaining strict authorization and access control, and securing processing environments [106]. However, a cybersecurity report in 2016 revealed a 320% year-on-year increase in breaches of protected health information in US hospitals, with 81% of breached records resulting from hacking attacks [107]. This compromised over 16 million individual patient health records, indicating a pressing need to continue monitoring and developing security systems in the face of both malicious and unintentional data breaches.

## Conclusions

The proliferation of web-based content and increased participation of patients in interacting with said content provides an opportunity to understand what features of content are engaging to people. Harnessing ML technologies to process *big data* in health care will allow health care providers and other users to create and contribute to personalized content. These insights may be leveraged to facilitate patient activation and enable patients to make healthy choices, ultimately improving outcomes.

## Authors' Contributions

UJ constructed themes for the manuscript. AG and PN wrote the paper. PN, AHD, and UJ critically revised the manuscript. All authors reviewed and approved the manuscript.

## Conflicts of Interest

None declared.

## References

1. O'Neill B, Ziebland S, Valderas J, Lupiáñez-Villanueva F. User-generated online health content: a survey of internet users in the United Kingdom. *J Med Internet Res* 2014;16(4):e118 [FREE Full text] [doi: [10.2196/jmir.3187](https://doi.org/10.2196/jmir.3187)] [Medline: [24784798](https://pubmed.ncbi.nlm.nih.gov/24784798/)]
2. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. *Arch Intern Med* 2005;165(22):2618-2624. [doi: [10.1001/archinte.165.22.2618](https://doi.org/10.1001/archinte.165.22.2618)] [Medline: [16344419](https://pubmed.ncbi.nlm.nih.gov/16344419/)]
3. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. *J Med Internet Res* 2009;11(1):e4 [FREE Full text] [doi: [10.2196/jmir.1035](https://doi.org/10.2196/jmir.1035)] [Medline: [19275980](https://pubmed.ncbi.nlm.nih.gov/19275980/)]
4. Kamel BMN, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. *Health Info Libr J* 2007 Mar;24(1):2-23. [doi: [10.1111/j.1471-1842.2007.00701.x](https://doi.org/10.1111/j.1471-1842.2007.00701.x)] [Medline: [17331140](https://pubmed.ncbi.nlm.nih.gov/17331140/)]
5. Podichetty VK, Booher J, Whitfield M, Biscup RS. Assessment of internet use and effects among healthcare professionals: a cross sectional survey. *Postgrad Med J* 2006 Apr;82(966):274-279 [FREE Full text] [doi: [10.1136/pgmj.2005.040675](https://doi.org/10.1136/pgmj.2005.040675)] [Medline: [16597816](https://pubmed.ncbi.nlm.nih.gov/16597816/)]
6. Pershad Y, Hange PT, Albadawi H, Oklu R. Social Medicine: Twitter in Healthcare. *J Clin Med* 2018 May 28;7(6) [FREE Full text] [doi: [10.3390/jcm7060121](https://doi.org/10.3390/jcm7060121)] [Medline: [29843360](https://pubmed.ncbi.nlm.nih.gov/29843360/)]
7. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J* 2015 Sep;21(3):173-194. [doi: [10.1177/1460458213512220](https://doi.org/10.1177/1460458213512220)] [Medline: [24670899](https://pubmed.ncbi.nlm.nih.gov/24670899/)]
8. Boulos MN, Maramba I, Wheeler S. Wikis, blogs and podcasts: a new generation of web-based tools for virtual collaborative clinical practice and education. *BMC Med Educ* 2006;6:41 [FREE Full text] [doi: [10.1186/1472-6920-6-41](https://doi.org/10.1186/1472-6920-6-41)] [Medline: [16911779](https://pubmed.ncbi.nlm.nih.gov/16911779/)]
9. Jin J, Yan X, Li Y, Li Y. *Int J Med Inform* 2016 Feb;86:91-103. [doi: [10.1016/j.ijmedinf.2015.11.002](https://doi.org/10.1016/j.ijmedinf.2015.11.002)] [Medline: [26616406](https://pubmed.ncbi.nlm.nih.gov/26616406/)]
10. Rupert DJ, Moultrie RR, Read JG, Amoozegar JB, Bornkessel AS, O'Donoghue AC, et al. Perceived healthcare provider reactions to patient and caregiver use of online health communities. *Patient Educ Couns* 2014 Sep;96(3):320-326. [doi: [10.1016/j.pec.2014.05.015](https://doi.org/10.1016/j.pec.2014.05.015)] [Medline: [24923652](https://pubmed.ncbi.nlm.nih.gov/24923652/)]
11. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *J Am Med Assoc* 2002;287(20):2691-2700. [doi: [10.1001/jama.287.20.2691](https://doi.org/10.1001/jama.287.20.2691)] [Medline: [12020305](https://pubmed.ncbi.nlm.nih.gov/12020305/)]
12. Benigeri M, Pluye P. Shortcomings of health information on the internet. *Health Promot Int* 2003 Dec;18(4):381-386. [doi: [10.1093/heapro/dag409](https://doi.org/10.1093/heapro/dag409)] [Medline: [14695369](https://pubmed.ncbi.nlm.nih.gov/14695369/)]
13. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *J Am Med Assoc* 2001;285(20):2612-2621 [FREE Full text] [doi: [10.1001/jama.285.20.2612](https://doi.org/10.1001/jama.285.20.2612)] [Medline: [11368735](https://pubmed.ncbi.nlm.nih.gov/11368735/)]
14. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. *Health Aff (Millwood)* 2013 Feb;32(2):207-214. [doi: [10.1377/hlthaff.2012.1061](https://doi.org/10.1377/hlthaff.2012.1061)] [Medline: [23381511](https://pubmed.ncbi.nlm.nih.gov/23381511/)]
15. Murphy K. *Machine Learning: A Probabilistic Perspective*. Cambridge, Massachusetts: MIT press; 2012.
16. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst* 2014;2:3 [FREE Full text] [doi: [10.1186/2047-2501-2-3](https://doi.org/10.1186/2047-2501-2-3)] [Medline: [25825667](https://pubmed.ncbi.nlm.nih.gov/25825667/)]
17. Farrell C. *Patient and public involvement in health : the evidence for policy implementation*. Department of Health. 2004. URL: <https://www.semanticscholar.org/paper/Patient-and-Public-Involvement-in-Health-%3A-The-for-Farrell/19d9d7a92bf348b04179b2b83a70cbc77ef6773d> [accessed 2021-07-24]
18. Five year forward view. National Health Service. 2014. URL: <https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf> [accessed 2020-10-22]
19. Levenstein JH, McCracken EC, McWhinney IR, Stewart MA, Brown JB. The patient-centred clinical method. 1. A model for the doctor-patient interaction in family medicine. *Fam Pract* 1986 Mar;3(1):24-30. [doi: [10.1093/fampra/3.1.24](https://doi.org/10.1093/fampra/3.1.24)] [Medline: [3956899](https://pubmed.ncbi.nlm.nih.gov/3956899/)]
20. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. *Med Care Res Rev* 2013 Aug;70(4):351-379. [doi: [10.1177/1077558712465774](https://doi.org/10.1177/1077558712465774)] [Medline: [23169897](https://pubmed.ncbi.nlm.nih.gov/23169897/)]
21. Jones R, Pearson J, McGregor S, Cawsey AJ, Barrett A, Craig N, et al. Randomised trial of personalised computer based information for cancer patients. *Br Med J* 1999 Nov 06;319(7219):1241-1247 [FREE Full text] [doi: [10.1136/bmj.319.7219.1241](https://doi.org/10.1136/bmj.319.7219.1241)] [Medline: [10550090](https://pubmed.ncbi.nlm.nih.gov/10550090/)]
22. Jones RB, Pearson J, Cawsey AJ, Bental D, Barrett A, White J, et al. Effect of different forms of information produced for cancer patients on their use of the information, social support, and anxiety: randomised trial. *Br Med J* 2006 Apr 22;332(7547):942-948 [FREE Full text] [doi: [10.1136/bmj.38807.571042.68](https://doi.org/10.1136/bmj.38807.571042.68)] [Medline: [16597660](https://pubmed.ncbi.nlm.nih.gov/16597660/)]
23. O'Connor G, Coates V, O'Neill S. Randomised controlled trial of a tailored information pack for patients undergoing surgery and treatment for rectal cancer. *Eur J Oncol Nurs* 2014 Apr;18(2):183-191. [doi: [10.1016/j.ejon.2013.10.011](https://doi.org/10.1016/j.ejon.2013.10.011)] [Medline: [24275208](https://pubmed.ncbi.nlm.nih.gov/24275208/)]
24. Mishel MH. Reconceptualization of the uncertainty in illness theory. *Image J Nurs Sch* 1990;22(4):256-262. [doi: [10.1111/j.1547-5069.1990.tb00225.x](https://doi.org/10.1111/j.1547-5069.1990.tb00225.x)] [Medline: [2292449](https://pubmed.ncbi.nlm.nih.gov/2292449/)]

25. Arora NK. Interacting with cancer patients: the significance of physicians' communication behavior. *Soc Sci Med* 2003 Sep;57(5):791-806. [doi: [10.1016/s0277-9536\(02\)00449-5](https://doi.org/10.1016/s0277-9536(02)00449-5)] [Medline: [12850107](https://pubmed.ncbi.nlm.nih.gov/12850107/)]
26. Nanton V, Docherty A, Meystre C, Dale J. Finding a pathway: information and uncertainty along the prostate cancer patient journey. *Br J Health Psychol* 2009 Sep;14(Pt 3):437-458. [doi: [10.1348/135910708X342890](https://doi.org/10.1348/135910708X342890)] [Medline: [18718111](https://pubmed.ncbi.nlm.nih.gov/18718111/)]
27. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. *N Engl J Med* 2012 Mar 01;366(9):780-781. [doi: [10.1056/NEJMp1109283](https://doi.org/10.1056/NEJMp1109283)] [Medline: [22375967](https://pubmed.ncbi.nlm.nih.gov/22375967/)]
28. Normahani P, Sounderajah V, Harrop-Griffiths W, Chukwuemeka A, Peters NS, Standfield NJ, et al. Achieving good-quality consent: review of literature, case law and guidance. *BJS Open* 2020 May 31:757-763 [FREE Full text] [doi: [10.1002/bjs5.50306](https://doi.org/10.1002/bjs5.50306)] [Medline: [32475083](https://pubmed.ncbi.nlm.nih.gov/32475083/)]
29. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2014;1:CD001431. [doi: [10.1002/14651858.CD001431.pub4](https://doi.org/10.1002/14651858.CD001431.pub4)] [Medline: [24470076](https://pubmed.ncbi.nlm.nih.gov/24470076/)]
30. Salkeld G, Cunich M, Dowie J, Howard K, Patel MI, Mann G, et al. The role of personalised choice in decision support: a randomized controlled trial of an online decision aid for prostate cancer screening. *PLoS One* 2016;11(4):e0152999 [FREE Full text] [doi: [10.1371/journal.pone.0152999](https://doi.org/10.1371/journal.pone.0152999)] [Medline: [27050101](https://pubmed.ncbi.nlm.nih.gov/27050101/)]
31. Kalsoft M, Cunich M, Salkeld G, Dowie J. Assessing decision quality in patient-centred care requires a preference-sensitive measure. *J Health Serv Res Policy* 2014 Apr;19(2):110-117 [FREE Full text] [doi: [10.1177/1355819613511076](https://doi.org/10.1177/1355819613511076)] [Medline: [24335587](https://pubmed.ncbi.nlm.nih.gov/24335587/)]
32. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient Educ Couns* 2009 Feb;74(2):156-173. [doi: [10.1016/j.pec.2008.08.023](https://doi.org/10.1016/j.pec.2008.08.023)] [Medline: [18947966](https://pubmed.ncbi.nlm.nih.gov/18947966/)]
33. Zufferey MC, Schulz PJ. Self-management of chronic low back pain: an exploration of the impact of a patient-centered website. *Patient Educ Couns* 2009 Oct;77(1):27-32. [doi: [10.1016/j.pec.2009.01.016](https://doi.org/10.1016/j.pec.2009.01.016)] [Medline: [19321286](https://pubmed.ncbi.nlm.nih.gov/19321286/)]
34. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. *J Health Commun* 2013;18(9):1039-1069. [doi: [10.1080/10810730.2013.768727](https://doi.org/10.1080/10810730.2013.768727)] [Medline: [23750972](https://pubmed.ncbi.nlm.nih.gov/23750972/)]
35. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. *Ann Behav Med* 2006 Jun;31(3):205-223. [doi: [10.1207/s15324796abm3103\\_2](https://doi.org/10.1207/s15324796abm3103_2)] [Medline: [16700634](https://pubmed.ncbi.nlm.nih.gov/16700634/)]
36. Neville LM, O'Hara B, Milat AJ. Computer-tailored dietary behaviour change interventions: a systematic review. *Health Educ Res* 2009 Aug;24(4):699-720 [FREE Full text] [doi: [10.1093/her/cyp006](https://doi.org/10.1093/her/cyp006)] [Medline: [19286893](https://pubmed.ncbi.nlm.nih.gov/19286893/)]
37. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. *Int J Behav Nutr Phys Act* 2009;6:30 [FREE Full text] [doi: [10.1186/1479-5868-6-30](https://doi.org/10.1186/1479-5868-6-30)] [Medline: [19490649](https://pubmed.ncbi.nlm.nih.gov/19490649/)]
38. Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. *Patient Educ Couns* 1999 Feb;36(2):107-117. [doi: [10.1016/s0738-3991\(98\)00128-1](https://doi.org/10.1016/s0738-3991(98)00128-1)] [Medline: [10223016](https://pubmed.ncbi.nlm.nih.gov/10223016/)]
39. Fabricant P, Dy C, Patel R, Blanco J, Doyle S. Internet search term affects the quality and accuracy of online information about developmental hip dysplasia. *J Pediatr Orthop* 2013 Jun;33(4):361-365. [doi: [10.1097/BPO.0b013e31827d0dd2](https://doi.org/10.1097/BPO.0b013e31827d0dd2)] [Medline: [23653022](https://pubmed.ncbi.nlm.nih.gov/23653022/)]
40. Shepperd S, Charnock D, Gann B. Helping patients access high quality health information. *Br Med J* 1999 Sep 18;319(7212):764-766 [FREE Full text] [doi: [10.1136/bmj.319.7212.764](https://doi.org/10.1136/bmj.319.7212.764)] [Medline: [10488009](https://pubmed.ncbi.nlm.nih.gov/10488009/)]
41. Bolle S, Romijn G, Smets EM, Loos EF, Kunnevan M, van Weert JC. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. *J Med Internet Res* 2016 Jul 25;18(7):e208 [FREE Full text] [doi: [10.2196/jmir.5618](https://doi.org/10.2196/jmir.5618)] [Medline: [27457709](https://pubmed.ncbi.nlm.nih.gov/27457709/)]
42. Basevitz P, Pushkar D, Chaikelson J, Conway M, Dalton C. Age-related differences in worry and related processes. *Int J Aging Hum Dev* 2008;66(4):283-305. [doi: [10.2190/AG.66.4.b](https://doi.org/10.2190/AG.66.4.b)] [Medline: [18507331](https://pubmed.ncbi.nlm.nih.gov/18507331/)]
43. Nguyen MH, Smets EM, Bol N, Loos EF, van Laarhoven HW, Geijsen D, et al. Tailored web-based information for younger and older patients with cancer: randomized controlled trial of a preparatory educational intervention on patient outcomes. *J Med Internet Res* 2019 Oct 01;21(10):e14407 [FREE Full text] [doi: [10.2196/14407](https://doi.org/10.2196/14407)] [Medline: [31573911](https://pubmed.ncbi.nlm.nih.gov/31573911/)]
44. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? *Soc Sci Med* 2007 May;64(9):1853-1862. [doi: [10.1016/j.socscimed.2007.01.012](https://doi.org/10.1016/j.socscimed.2007.01.012)] [Medline: [17328998](https://pubmed.ncbi.nlm.nih.gov/17328998/)]
45. The top 500 sites on the web. Alexa. 2020. URL: <https://www.alexa.com/topsites> [accessed 2020-10-07]
46. Szmuda T, Rosvall P, Hetzger TV, Ali S, Słoniewski P. Youtube as a source of patient information for hydrocephalus: a content-quality and optimization analysis. *World Neurosurg* 2020 Jun;138:469-477. [doi: [10.1016/j.wneu.2020.02.149](https://doi.org/10.1016/j.wneu.2020.02.149)] [Medline: [32147561](https://pubmed.ncbi.nlm.nih.gov/32147561/)]
47. Szmuda T, Syed MT, Singh A, Ali S, Özdemir C, Słoniewski P. YouTube as a source of patient information for Coronavirus Disease (COVID-19): a content-quality and audience engagement analysis. *Rev Med Virol* 2020 Sep;30(5):e2132 [FREE Full text] [doi: [10.1002/rmv.2132](https://doi.org/10.1002/rmv.2132)] [Medline: [32537771](https://pubmed.ncbi.nlm.nih.gov/32537771/)]

48. Fode M, Nolsøe AB, Jacobsen FM, Russo GI, Østergren PB, Jensen CF, EAU YAU Men's Health Working Group. Quality of information in YouTube videos on erectile dysfunction. *Sex Med* 2020 Sep;8(3):408-413 [FREE Full text] [doi: [10.1016/j.esxm.2020.05.007](https://doi.org/10.1016/j.esxm.2020.05.007)] [Medline: [32593674](https://pubmed.ncbi.nlm.nih.gov/32593674/)]
49. Ajumobi AB, Malakouti M, Bullen A, Ahaneku H, Lunsford TN. YouTube™ as a source of instructional videos on bowel preparation: a content analysis. *J Cancer Educ* 2016 Dec;31(4):755-759. [doi: [10.1007/s13187-015-0888-y](https://doi.org/10.1007/s13187-015-0888-y)] [Medline: [26224241](https://pubmed.ncbi.nlm.nih.gov/26224241/)]
50. Abdulghani HM, Haque S, Ahmad T, Irshad M, Sattar K, Al-Harbi MM, et al. A critical review of obstetric and gynecological physical examination videos available on YouTube: content analysis and user engagement evaluation. *Medicine (Baltimore)* 2019 Jul;98(30):e16459 [FREE Full text] [doi: [10.1097/MD.00000000000016459](https://doi.org/10.1097/MD.00000000000016459)] [Medline: [31348247](https://pubmed.ncbi.nlm.nih.gov/31348247/)]
51. Keelan J, Pavri-Garcia V, Tomlinson G, Wilson K. YouTube as a source of information on immunization: a content analysis. *J Am Med Assoc* 2007 Dec 5;298(21):2482-2484. [doi: [10.1001/jama.298.21.2482](https://doi.org/10.1001/jama.298.21.2482)] [Medline: [18056901](https://pubmed.ncbi.nlm.nih.gov/18056901/)]
52. Welbourne DJ, Grant WJ. Science communication on YouTube: factors that affect channel and video popularity. *Public Underst Sci* 2016 Aug;25(6):706-718. [doi: [10.1177/0963662515572068](https://doi.org/10.1177/0963662515572068)] [Medline: [25698225](https://pubmed.ncbi.nlm.nih.gov/25698225/)]
53. Ferhatoglu MF, Kartal A, Ekici U, Gurkan A. Evaluation of the reliability, utility, and quality of the information in sleeve gastrectomy videos shared on open access video sharing platform YouTube. *Obes Surg* 2019 May;29(5):1477-1484. [doi: [10.1007/s11695-019-03738-2](https://doi.org/10.1007/s11695-019-03738-2)] [Medline: [30706318](https://pubmed.ncbi.nlm.nih.gov/30706318/)]
54. Improve video performance with shareable's new Video Power Index (VPI). Shareable. 2020. URL: <https://www.shareable.com/blog/2017/06/27/improve-video-performance-with-shareable-s-new-video-power-index-vpi> [accessed 2020-10-07]
55. Finnegan G, Holt D, English PM, Glismann S, Thomson A, Salisbury DM, et al. Lessons from an online vaccine communication project. *Vaccine* 2018 Oct 22;36(44):6509-6511. [doi: [10.1016/j.vaccine.2018.05.007](https://doi.org/10.1016/j.vaccine.2018.05.007)] [Medline: [29921491](https://pubmed.ncbi.nlm.nih.gov/29921491/)]
56. Sorice SC, Li AY, Gilstrap J, Canales FL, Furnas HJ. Social media and the plastic surgery patient. *Plast Reconstr Surg* 2017;140(5):1047-1056. [doi: [10.1097/prs.0000000000003769](https://doi.org/10.1097/prs.0000000000003769)]
57. Oktay LA, Abdelwahed A, Houbby N, Lampridou S, Normahani P, Peters N, et al. Factors affecting engagement in online healthcare patient information: a systematic review of the literature. *J Med Internet Res* 2020:In Press. [doi: [10.2196/preprints.19896](https://doi.org/10.2196/preprints.19896)]
58. STUDY: Rethinking retail, insights from consumers and retailers into an omni-channel shopping experience. Infosys. 2013. URL: <https://www.infosys.com/newsroom/press-releases/Documents/genome-research-report.pdf> [accessed 2020-10-15]
59. Appboy data analysis of over 30,000 campaigns. Braze. 2016. URL: <https://www.braze.com/blog/in-app-conversion-5-tips-infographic/> [accessed 2020-10-15]
60. Coolil B, Aksoy L, Keiningham TL. Approaches to customer segmentation. *J Relation Mark* 2008 Jan 14;6(3-4):9-39. [doi: [10.1300/j366v06n03\\_02](https://doi.org/10.1300/j366v06n03_02)]
61. An J, Kwak H, Jung S, Salminen J, Jansen BJ. Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data. *Soc Netw Anal Min* 2018 Aug 23;8(1):54. [doi: [10.1007/s13278-018-0531-0](https://doi.org/10.1007/s13278-018-0531-0)]
62. Murray G, Scime A. Microtargeting and electorate segmentation: data mining the American National Election Studies. *J Polit Mark* 2010 Jul;9(3):143-166. [doi: [10.1080/15377857.2010.497732](https://doi.org/10.1080/15377857.2010.497732)]
63. Gomez-Uribe CA, Hunt N. The Netflix recommender system: algorithms, business value, and innovation. *ACM Trans Manage Inf Syst* 2016 Jan 14;6(4):1-19. [doi: [10.1145/2843948](https://doi.org/10.1145/2843948)]
64. Davidson J, Liebold B, Liu J, Nandy P, Van Vleet T, Gargi U, et al. The YouTube video recommendation system. In: *Proceedings of the fourth ACM conference on Recommender Systems*. 2010 Presented at: RecSys '10: Fourth ACM Conference on Recommender Systems; September 26 - 30, 2010; Barcelona Spain p. 293-296. [doi: [10.1145/1864708.1864770](https://doi.org/10.1145/1864708.1864770)]
65. Wei K, Huang J, Fu S. A survey of e-commerce recommender systems. In: *Proceedings of the International Conference on Service Systems and Service Management*. 2007 Presented at: International Conference on Service Systems and Service Management; June 9-11, 2007; Chengdu, China. [doi: [10.1109/icsssm.2007.4280214](https://doi.org/10.1109/icsssm.2007.4280214)]
66. Linden G, Smith B, York J. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Comput* 2003 Jan;7(1):76-80. [doi: [10.1109/mic.2003.1167344](https://doi.org/10.1109/mic.2003.1167344)]
67. Jannach D, Zanker M, Ge M, Gröning M. Recommender systems in computer science and information systems – a landscape of research. In: *E-Commerce and Web Technologies*. Berlin: Springer; 2012:76-87.
68. Lops P, Jannach D, Musto C, Bogers T, Koolen M. Trends in content-based recommendation. *User Model User Adap Interact* 2019 Mar 7;29(2):239-249. [doi: [10.1007/s11257-019-09231-w](https://doi.org/10.1007/s11257-019-09231-w)]
69. Chen L, Chen G, Wang F. Recommender systems based on user reviews: the state of the art. *User Model User Adap Interact* 2015 Jan 22;25(2):99-154. [doi: [10.1007/s11257-015-9155-5](https://doi.org/10.1007/s11257-015-9155-5)]
70. Deldjoo Y, Elahi M, Quadrana M, Cremonesi P. Using visual features based on MPEG-7 and deep learning for movie recommendation. *Int J Multimed Info Retr* 2018 Jun 14;7(4):207-219. [doi: [10.1007/s13735-018-0155-1](https://doi.org/10.1007/s13735-018-0155-1)]
71. McAuley J, Targett C, Shi Q, Van Den HA. Image-based recommendations on styles and substitutes. In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2015 Presented

- at: SIGIR '15: The 38th International ACM SIGIR conference on research and development in Information Retrieval; August 9 - 13, 2015; Santiago Chile p. 43-52. [doi: [10.1145/2766462.2767755](https://doi.org/10.1145/2766462.2767755)]
72. Deep neural networks for YouTube recommendations. Google. 2016. URL: <https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/45530.pdf> [accessed 2020-10-30]
73. Du X, Yin H, Chen L, Wang Y, Yang Y, Zhou X. Personalized video recommendation using rich contents from videos. *IEEE Trans Knowl Data Eng* 2020 Mar 1;32(3):492-505. [doi: [10.1109/tkde.2018.2885520](https://doi.org/10.1109/tkde.2018.2885520)]
74. Powered by AI: Instagram's explore recommender system. Facebook AI. 2019. URL: <https://ai.facebook.com/blog/powered-by-ai-instagram-s-explore-recommender-system/> [accessed 2020-10-29]
75. Jaakonmäki R, Müller O, vom Brocke BJ. The impact of content, context, and creator on user engagement in social media marketing. *ScholarSpace* 2017:1-9. [doi: [10.24251/HICSS.2017.136](https://doi.org/10.24251/HICSS.2017.136)]
76. Bakhshi S, Shamma D, Gilbert E. Faces engage us: photos with faces attract more likes and comments on Instagram. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26 - May 1, 2014; Toronto Ontario Canada p. 965-974. [doi: [10.1145/2556288.2557403](https://doi.org/10.1145/2556288.2557403)]
77. Bakhshi S, Shamma D, Kennedy L, Gilbert E. Why we filter our photos and how it impacts engagement. *ICWSM*. 2015. URL: <http://comp.social.gatech.edu/papers/icwsml5.why.bakhshi.pdf> [accessed 2021-07-24]
78. Lee D, Hosanagar K, Nair H. The effect of social media marketing content on consumer engagement: evidence from Facebook. *Stanford University, Graduate School of Business, CA*. 2014. URL: <https://econpapers.repec.org/paper/eclstabus/3087.htm> [accessed 2021-07-24]
79. Elahi M, Deldjoo Y, Bakhshandegan MF, Cella L, Cereda S, Cremonesi P. Exploring the semantic gap for movie recommendations. In: *Proceedings of the Eleventh ACM Conference on Recommender Systems*. 2017 Presented at: RecSys '17: Eleventh ACM Conference on Recommender Systems; August 27 - 31, 2017; Como Italy p. 326-330. [doi: [10.1145/3109859.3109908](https://doi.org/10.1145/3109859.3109908)]
80. Bisong E. An overview of Google Cloud platform services. In: *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Berkeley, CA: Apress; 2019:7-10.
81. AI and machine learning products. Google Cloud. URL: <https://cloud.google.com/products/ai> [accessed 2020-10-07]
82. Amazon Rekognition. Amazon Web Services. URL: <https://aws.amazon.com/rekognition> [accessed 2020-10-07]
83. Azure products. Microsoft. URL: <https://azure.microsoft.com/en-us/services/> [accessed 2020-10-07]
84. IBM Watson products and solutions. IBM. URL: <https://www.ibm.com/watson/products-services> [accessed 2020-10-07]
85. Predictive user engagement. Amazon Web Services. 2019. URL: <https://docs.aws.amazon.com/solutions/latest/predictive-user-engagement> [accessed 2020-10-07]
86. Predictive user engagement solution. GitHub. URL: <https://github.com/aws-labs/predictive-user-engagement> [accessed 2020-10-07]
87. What is health marketing? Centers for Disease Control and Prevention. 2011. URL: <https://www.cdc.gov/healthcommunication/toolstemplates/WhatIsHM.html> [accessed 2020-10-15]
88. Swenson ER, Bastian ND, Nembhard HB. Healthcare market segmentation and data mining: a systematic review. *Health Mark Q* 2018;35(3):186-208. [doi: [10.1080/07359683.2018.1514734](https://doi.org/10.1080/07359683.2018.1514734)] [Medline: [30470165](https://pubmed.ncbi.nlm.nih.gov/30470165/)]
89. Maddox TM, Matheny MA. Natural language processing and the promise of big data: small step forward, but many miles to go. *Circ Cardiovasc Qual Outcomes* 2015 Sep;8(5):463-465. [doi: [10.1161/CIRCOUTCOMES.115.002125](https://doi.org/10.1161/CIRCOUTCOMES.115.002125)] [Medline: [26286870](https://pubmed.ncbi.nlm.nih.gov/26286870/)]
90. O'Dowd E. Unstructured healthcare data needs advanced machine learning tools. *HIT Infrastructure*. 2018. URL: <https://hitinfrastructure.com/news/unstructured-healthcare-data-needs-advanced-machine-learning-tools> [accessed 2020-10-26]
91. Ohno-Machado L. Realizing the full potential of electronic health records: the role of natural language processing. *J Am Med Inform Assoc* 2011;18(5):539 [FREE Full text] [doi: [10.1136/amiajnl-2011-000501](https://doi.org/10.1136/amiajnl-2011-000501)] [Medline: [21846784](https://pubmed.ncbi.nlm.nih.gov/21846784/)]
92. Gupta D, Rodeghier M, Lis CG. Patient satisfaction with service quality as a predictor of survival outcomes in breast cancer. *Support Care Cancer* 2014 Jan;22(1):129-134. [doi: [10.1007/s00520-013-1956-7](https://doi.org/10.1007/s00520-013-1956-7)] [Medline: [24013568](https://pubmed.ncbi.nlm.nih.gov/24013568/)]
93. Armstrong AW, Kim RH, Idriss NZ, Larsen LN, Lio PA. Online video improves clinical outcomes in adults with atopic dermatitis: a randomized controlled trial. *J Am Acad Dermatol* 2011 Mar;64(3):502-507. [doi: [10.1016/j.jaad.2010.01.051](https://doi.org/10.1016/j.jaad.2010.01.051)] [Medline: [21236514](https://pubmed.ncbi.nlm.nih.gov/21236514/)]
94. Hoffman J, Salzman C, Garbaccio C, Burns SP, Crane D, Bombardier C. Use of on-demand video to provide patient education on spinal cord injury. *J Spinal Cord Med* 2011;34(4):404-409 [FREE Full text] [doi: [10.1179/2045772311Y.0000000015](https://doi.org/10.1179/2045772311Y.0000000015)] [Medline: [21903014](https://pubmed.ncbi.nlm.nih.gov/21903014/)]
95. Zhao J, Han H, Zhong B, Xie W, Chen Y, Zhi M. Health information on social media helps mitigate Crohn's disease symptoms and improves patients' clinical course. *Comput Hum Behav* 2021 Feb;115:106588. [doi: [10.1016/j.chb.2020.106588](https://doi.org/10.1016/j.chb.2020.106588)]
96. Belle A, Hobson R, Najarian K. A physiological signal processing system for optimal engagement and attention detection. In: *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. 2011 Presented at: IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW); Nov. 12-15, 2011; Atlanta, GA, USA. [doi: [10.1109/bibmw.2011.6112429](https://doi.org/10.1109/bibmw.2011.6112429)]

97. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med* 2019;2:88 [FREE Full text] [doi: [10.1038/s41746-019-0166-1](https://doi.org/10.1038/s41746-019-0166-1)] [Medline: [31508498](https://pubmed.ncbi.nlm.nih.gov/31508498/)]
98. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annu Rev Clin Psychol* 2017 May 08;13:23-47. [doi: [10.1146/annurev-clinpsy-032816-044949](https://doi.org/10.1146/annurev-clinpsy-032816-044949)] [Medline: [28375728](https://pubmed.ncbi.nlm.nih.gov/28375728/)]
99. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. *JAMA Netw Open* 2020 Jul 01;3(7):e2010791 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.10791](https://doi.org/10.1001/jamanetworkopen.2020.10791)] [Medline: [32678450](https://pubmed.ncbi.nlm.nih.gov/32678450/)]
100. Char DS, Abràmoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth* 2020 Nov;20(11):7-17. [doi: [10.1080/15265161.2020.1819469](https://doi.org/10.1080/15265161.2020.1819469)] [Medline: [33103967](https://pubmed.ncbi.nlm.nih.gov/33103967/)]
101. The future of healthcare: our vision for digital, data and technology in health and care. Department of Health & Social Care. 2018. URL: <https://www.gov.uk/government/publications/the-future-of-healthcare-our-vision-for-digital-data-and-technology-in-health-and-care> [accessed 2021-01-17]
102. Exchange of electronic health records across the EU. European Commission. URL: <https://ec.europa.eu/digital-single-market/en/exchange-electronic-health-records-across-eu> [accessed 2021-01-11]
103. Gray EA, Thorpe JH. Comparative effectiveness research and big data: balancing potential with legal and ethical considerations. *J Comp Eff Res* 2015 Jan;4(1):61-74 [FREE Full text] [doi: [10.2217/ce.14.51](https://doi.org/10.2217/ce.14.51)] [Medline: [25565069](https://pubmed.ncbi.nlm.nih.gov/25565069/)]
104. Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019 Jan;25(1):37-43 [FREE Full text] [doi: [10.1038/s41591-018-0272-7](https://doi.org/10.1038/s41591-018-0272-7)] [Medline: [30617331](https://pubmed.ncbi.nlm.nih.gov/30617331/)]
105. No authors listed. Time to discuss consent in digital-data studies. *Nature* 2019 Aug;572(7767):5. [doi: [10.1038/d41586-019-02322-z](https://doi.org/10.1038/d41586-019-02322-z)] [Medline: [31367033](https://pubmed.ncbi.nlm.nih.gov/31367033/)]
106. Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. *J Big Data* 2018 Jan 9;5(1):1. [doi: [10.1186/s40537-017-0110-7](https://doi.org/10.1186/s40537-017-0110-7)]
107. 2016 Breach Report. Cynergistik. 2016. URL: <https://insights.cynergistik.com/reports/2016-breach-report> [accessed 2020-10-15]

## Abbreviations

**API:** application programming interface

**ML:** machine learning

*Edited by R Kukafka; submitted 04.11.20; peer-reviewed by P Dattathreya, A Teles; comments to author 03.01.21; revised version received 19.01.21; accepted 16.03.21; published 19.10.21*

*Please cite as:*

Guni A, Normahani P, Davies A, Jaffer U

Harnessing Machine Learning to Personalize Web-Based Health Care Content

*J Med Internet Res* 2021;23(10):e25497

URL: <https://www.jmir.org/2021/10/e25497>

doi: [10.2196/25497](https://doi.org/10.2196/25497)

PMID:

©Ahmad Guni, Pasha Normahani, Alun Davies, Usman Jaffer. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 19.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.