Perspective

Utilization of Structured Data to Improve Visibility of Digital Health Websites

Corresponding Author: Ms Suhanee Mitragotri
Co-Authors: Mr David Chen, Dr Jarone Lee

Edited By: Dr Shuhan He

CONDUCTSCIENCE
ACADEMIC PUBLISHING

Corresponding Author: Ms Suhanee Mitragotri
Co-Authors: Mr David Chen, Dr Jarone Lee
Type: Perspective
Keywords: structured data, search engine optimization, digital health, website, hospital
Edited By: Dr Shuhan He
Received: 2024-04-16
Published: 2024-12-24

Utilization of Structured Data to Improve Visibility of Digital Health Websites

ABSTRACT

Objective

This article proposes a framework to design, organize, and implement structured data, a standardized format to embed metadata into websites, to improve visibility of information from digital health websites.

Methods

We described the utility of structured data to improve visibility of websites and a proposed framework to design and evaluate context-specific structured data using existing structured data schemas.

Results

To motivate the increased use of structured data of health websites, we 1) highlight the considerations of implementing structured data in websites, 2) outline a framework and case study to build structured data, and 3) describe practical methods to implement structured data into healthcare websites.

Discussion

The use of structured data in healthcare websites can increase the visibility of health resources through search engine optimization (SEO), show rich, personalized results to target users, and allow the website to serve as a living digital directory of resource information for patients.

Conclusion

Further quality improvement and implementation science research is needed to promote the systematic uptake of structured data as a tool to deliver personalized healthcare information to patients.

INTRODUCTION

One of the key problems that currently faces healthcare is equity and accessibility. Patients' first access point to healthcare resource information is often through the online websites of hospitals and healthcare organizations.[1]https://doi.org/10.2196/27750 Furthermore, 83% of online users seek out health information utilizing general search engines, motivating the need to optimize the visibility of health facility and resource websites to connect patients with accurate information from reliable sources.[2]https://doi.org/10.3390/healthcare9121740 Patient users find it easier to use websites that have an organized layout, interactive web features, and identifiable sources and authorship of information to assess the credibility of the health information. [3]https://doi.org/10.2196/jmir.7579

At a national level in the US, developing an updated repository of healthcare resources is a primary goal of the Centers for Medicare and Medicaid Services (CMS).[4] To complement public-funded initiatives, private not-for-profit organizations such as Health On the Net Foundation have been independently monitoring health information websites to ensure that they maintain high quality, accurate health information.[5]https://doi.org/10.1016/s0010-4825(98)00037-7 Taken together, these entities survey online health resources to ensure that they are reliable and up-to-date to effectively address the needs of the communities that they are designed to serve.[6] The limitation of maintaining an updated, centralized repository of healthcare resources is the need for manual self-reporting surveys, such as surveys that report the prevalence of specific hospital departments and resource distribution across the U.S.[7]https://doi.org/10.1111/acem.13633

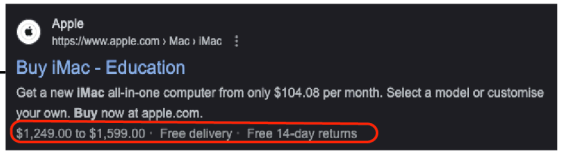
To extend these efforts of national healthcare resource mapping, health websites of healthcare facilities and organizations should report accessible information about the services they offer and accurate health information directly to patient users. Structured data of websites in schemas can be used to facilitate retrieval of pre-processed, explicit metadata of websites by search engines. Designing structured data as part of healthcare websites ensures that modern web crawlers can readily collect website metadata to connect search queries of online patients with personalized results in the age of digital medicine.

STRUCTURED DATA AND SCHEMAS

Structured data is the machine-readable metadata of websites that explicitly represents the informational content and metadata of websites. Structured data reports key information to search engines as part of a standardized schema with specified property types and offers the ability to enrich search results to provide additional multimedia elements. For instance, these enriched results may include text, pictures, and videos that are more engaging than what a traditional text-only search result may yield.[8] These are known as "rich features" and are offered by structured data beyond improving search engine optimization (Figure 1).

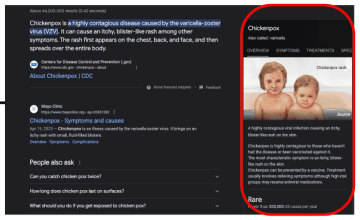
Rich snippets

A summary of key points from a website displayed below each website link in a list of search results



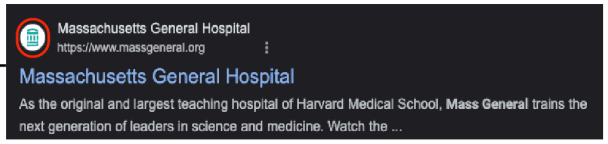
Knowledge graph panels

Information boxes that appear on the browser that provide extra information about a topic that a



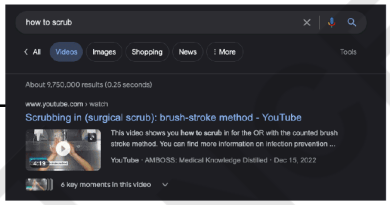
Logo

Your organization's logo in search results and knowledge panels



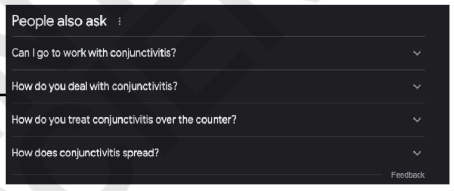
Video

Videos are displayed as results to an engine search



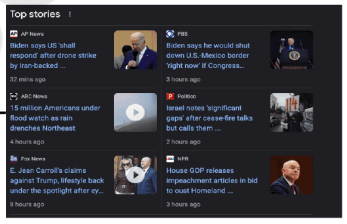
FAQ

Consists of questions and answers pertaining to a specific topic



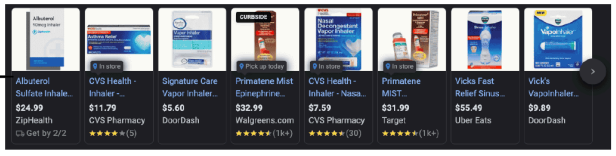
Article

An article that is displayed using rich result features upon an Internet search



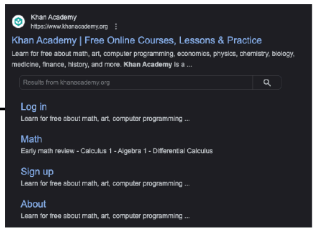
Carousel

Rich results that display in a list or gallery from a particular site



Course

Rich results that display information from an educational website



Breadcrumbs

A breakcrumb trail on a website shows where a particular page is located in a website's hierarchy

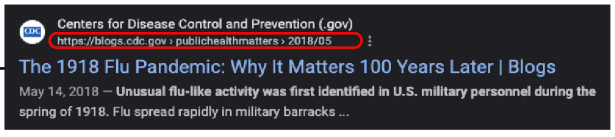
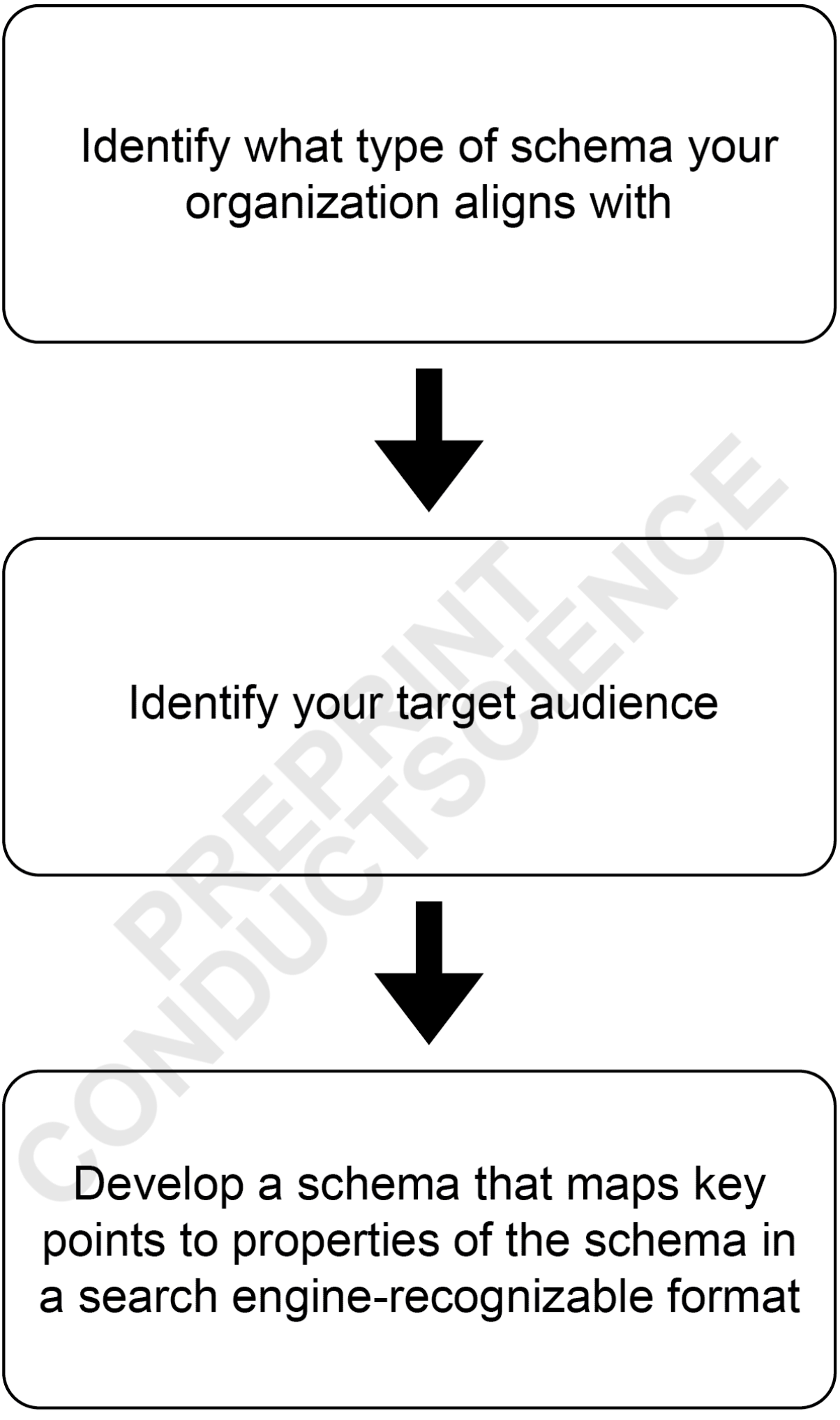


Figure 1: Types of rich features that can be implemented using website structured data.

Optimizing search engine visibility through the implementation and evaluation of structured data can help address the need to improve the visibility of information from healthcare facilities and organizations to their target demographics.[9] Evaluation of structured data of websites is commonly implemented using data formats such as JavaScript Object Notation for Linked Data (JSON-LD), Resource Description Framework (RDF), and Microdata.

To implement structured data in standardized vocabulary, open-source schema dictionaries such as those hosted by Schema.org, a collective initiative founded by Google, Microsoft, Yahoo, and Yandex, are commonly used to ensure that search engines can accurately parse and meaningfully interpret standardized structured data. Implementation of structured data can increase a website's click-through rate (CTR), a quantitative measure of how often a person clicks on a link when it is displayed to them.[9] Likewise, implementation of structured data markup can be used to increase SEO by accurately connecting patient queries with target information. [10]https://doi.org/10.12955/cbup.v3.645[11]https://doi.org/10.21083/partnership.v10i1.3328[12] [13][14]

The Unified Medical Language System (UMLS) was created to offer cohesive access to numerous biomedical resources, which was achieved by consolidating the vocabularies utilized to access those resources.[15] Schema can build on this already existing framework by creating a vocabulary that is specific to online health resources. In fact, biomedical ontology is available under Schema.org and can be used to develop structured data according to the MedicalOrganization schema for hospital websites to improve visibility and accessibility of health services and information.[14] For example, hospital websites that explicitly state that they have a pediatric emergency department will make it easy for search engines to map pediatric emergency resources in a local area. Through utilizing structured data in standardized schemas, external search engines can automatically extract relevant content from hospital websites and offer them as a living, updated repository of resources to online users seeking out health information and services. To achieve this aim, we propose a framework to design, organize, and implement structured data of digital health websites (Figure 2).



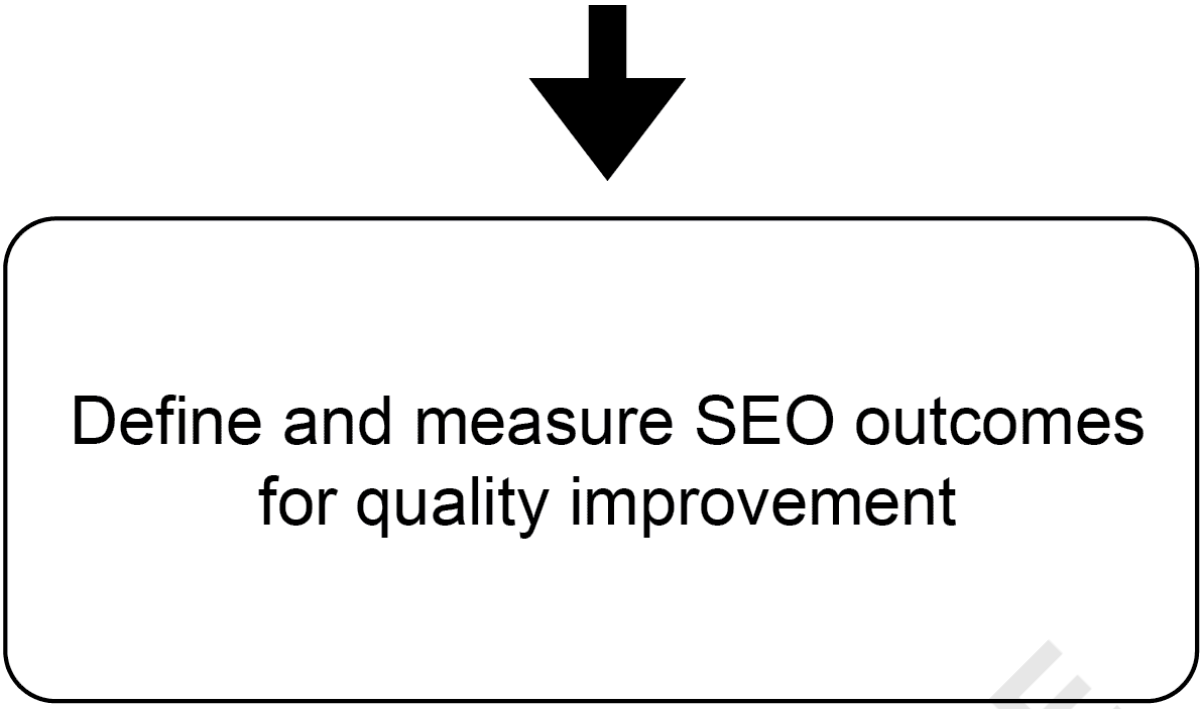


Figure 2: Framework to guide the development of improved digital health websites using structured data

Framework

Identify what type of schema your organization aligns with

The first step is to understand what types of schemas and rich features make the most sense to present health information for your organization. For example, a Primary Stroke Center may find it useful to explicitly state their stroke center credential and feature rich multimedia about the services they provide for the prevention and treatment of strokes. The decision to choose a specific schema template for structured data is dependent on what services, features, and information you want to make publicly visible for your organization.

Identify your target audience and needs

The second step is to identify the target demographics of online users for your website and adapt the structured data of your website to match the accessible needs of the demographics. For example, if your hospital is located in an area that consists predominantly of people who may not speak English, ensure that the structured data of your website includes metadata in the languages of your target demographic and states the services that may be offered in different languages. Through user-centric quality improvement research, informational elements should be tailored to the unique needs of your target audience.

Develop a schema that maps key points to properties of the schema in a search engine-recognizable format

The third step is to convert those informational elements to a set of structured data properties and rich results to embed in the website. When you create a schema for your website, you want to ensure that it is capturing the most important parts of your website and efficiently mapping that information to the search engine.

Define and measure outcomes for quality improvement

The fourth step is to conduct quality improvement research to evaluate the effect of different structured data implementations on patient-reported outcomes and quantitative SEO measures of

the target audience. We highlight a set of common SEO measures to assess the effect of structured data implementation in Table 1.

Table 1: Common SEO Metrics and Lay Interpretation

Metric Name	Metric Definition	How to interpret metric
Page views	The number of times users view your website[16]https://doi.org/10.2196/jmir.2632	Google Analytics can provide information on page views, which can inform you on if there is a need to improve the structured data within your website or modify rich features
Session duration	The amount of time a user spends on your website[16]https://doi.org/10.2196/jmir.2632	Google Analytics can provide information on session duration, which can inform you on if there is a need to improve the structured data within your website or modify rich features
Bounce rate	The number of times a user leave your site after viewing only one page[16]https://doi.org/10.2196/jmir.2632	Google Analytics' Audience Overview report provides the bounce rate for your whole website and the All Pages report provides the bounce rate for each individual page

CASE STUDY: STRUCTURED DATA FOR THE MGH EMERGENCY DEPARTMENT WEBSITE

To demonstrate a proof-of-concept of developing website structured data in practice, we conducted a case study to generate structured data for a custom schema tailored to the Massachusetts General Hospital (MGH) Emergency Medicine department website (https://www.massgeneral.org/emergency-medicine). First, we selected data elements and their associated structured data properties that were relevant to emergency medicine care at MGH based on a consensus review from study investigators. Each of the data properties have been categorized within one of the following larger schema classes from Schema.org: Organization, Hospital, Civic Structure, Local Business, Place, or Medical Organization.

Within the Organization schema class, we included data properties that provide key general information about the hospital, including its physical address, credentials, and diversity policy. Within the Hospital schema class, we included information about the specific services that the Massachusetts General Hospital emergency department offers, which is useful for patient users to confirm if the facility offers the care that they need. Within the Civic Structure schema class, we included information about the hours of operation of the facility, which can be helpful for patients

who are seeking care at times outside of regular daytime working hours. Within the schema class of Local Business, we included information about modes of payment offered by the facility, which is particularly useful for patients seeking care from outside of the United States who typically pay using a currency other than the US dollar. Within the schema class of Place, we included administrative information on the hospital, including its policies, maps, and contact information, which a patient may use in order to initiate contact and access care. Within the Medical Organization schema class, we included a data property that states if the facility is accepting new patients.

We showcase the proposed outline of structured data properties as part of the custom schema of the Massachusetts General Hospital (MGH) emergency department website in Table 2. The proposed schema would be implemented directly in a structured format with standardized schema data properties within the website HTML.

Table 2: Outline of structured data properties for custom schema of the MGH Emergency Department website

Schema Class	Property	Expected Data Type	Description
Organization	url	URL	https://www.massgeneral.org/emergency-medicine
Organization	address	PostalAddress or Text	55 Fruit St, Boston, MA 02114[17]
Organization	areaServed	AdministrativeArea or GeoShape or Place or Text	National[17]
Organization	award	Text	U.S. News and World Report, The Centers for Medicare and Medicaid Services rated 5/5, The American Nurses Credentialing Center awarded highest honor available for nursing excellence, The Human Rights Campaign Foundation,

			DiversityInc named MGH to Top Hospitals and Health Systems, The American Heart Association awarded Mass General with the Get With the Guidelines Stroke Gold Plus with Honor Roll, Advanced Therapy and Target: Type 2 Diabetes Honor Roll achievement awards[18]https://www.massgeneral.org/quality-and-safety/awards
Organization	contactPoint	ContactPoint	617-724-4100 (MGH Emergency Department Front Desk)[19]
Organization	diversityPolicy	CreativeWork or URL	https://www.massgeneral.org/careers/commitment-to-diversity
Organization	description	Text or TextObject	The Massachusetts General Hospital Department of Emergency Medicine offers full-service, state-of-the-art emergency care around the clock. Our doctors and nurses are equipped to handle any medical emergency.[17]
Organization	hasCredential	EducationalOccupationalCredential	Level 1 Trauma Center[20]https://www.massgeneral.org/surgery/trauma-center

Organization	legalName	Text	Massachusetts General Hospital Emergency Department
Organization	location	Place or PostalAddress or Text or VirtualLocation	55 Fruit St, Boston, MA 02114[17]
Organization	url	URL	https://www.massgeneral.org/emergency-medicine
Hospital	availableService	MedicalProcedure or MedicalTest or MedicalTherapy	Disaster Medicine, Neurologic Emergencies, Vascular Emergencies, Critical Care, Emergency Ultrasound, Geriatric Emergency Medicine, Pediatric Emergency Medicine, Wilderness Medicine, Emergency Pharmacy, Freedom Clinic, Emergency Medical Services (EMS)[21]
Hospital	medicalSpecialty	MedicalSpecialty	Emergency Medicine
Civic Structure	openingHours	Text	Monday through Sunday, all day[17]
LocalBusiness	currenciesAccepted	Text	USD[22]
LocalBusiness	paymentAccepted	Text	Credit Card, Check[22]

Place	address	PostalAddress or Text	55 Fruit St, Boston, MA 02114[17]
Place	containedInPlace	Place	Massachusetts General Hospital
Place	event	Event	Life Cycle Assessment Boot Camp: LCA for the Health Sector, Harvard Interdisciplinary Program in Psychedelics, Regenerative Agriculture, Fall Session: Mind Body Program for Individuals at High Risk for Cancer, High Cholesterol, Multiple Sclerosis: Understanding the Condition and Working to Mitigate Its Impact[23]https://www.massgeneral.org/events
Place	faxNumber	Text	617-726-3661[17]
Place	geo	GeoCoordinates or GeoShape	42.36300027491238, -71.06867454806678[24]
Place	hasMap	Map or URL	https://www.massgeneral.org/assets/mgh/pdf/visit/main-campus-map.pdf
Place	keywords	DefinedTerm or Text or	Emergency department (ED), trauma, triage, critical care, acute care, life-

		URL	threatening, emergency medical services (EMS), emergency physician, emergency nurse, intensive care unit (ICU), airway management, shock, cardiac arrest, resuscitation, stroke, seizure, sepsis, fracture, overdose, hypertension, myocardial infarction[17]
Place	latitude	Number or Text	42.36300027491238[24]
Place	logo	ImageObject or URL	https://www.massgeneral.org/
Place	longitude	Number or Text	-71.06867454806678[23]https://www.massgeneral.org/events
Place	maximumAttendeeCapacity	Integer	31 beds[17]
Place	smokingAllowed	Boolean	No[25]
Place	telephone	Text	617-724-4100[17]
Medical Organization	isAcceptingNewPatients	Boolean	Yes[17]

To implement structured data in schema format, the first step is to write structured data in the form of JSON-LD,[30] Microdata,[31] or RDF schema languages.[32] The second step is to test the structured data using tools such as the Google Rich Results Test[27] and the Schema Markup Validator[28]https://validator.schema.org/ to assess if your website's structured data can generate rich results and is implemented in a machine-readable format for general purpose search engines.

The third step is to conduct iterative revisions to the structured data on the website in order to ensure the accuracy of the provided information over time and to improve search engine optimization metrics for increased visibility and accessibility of your website. To implement custom, machine-readable structured data in schema format in practice, we highlight the main steps and example tools to facilitate each step in Table 3.

Table 3: Steps to implement structured data in schema format for websites

Task	Example Tool	Tool Description
Write and implement structured data	JSON-LD generator	The JSON-LD Generator[26] can generate Google-validated markups.
Test structured data	Google Rich Results Test and Schema Markup validator	The Google Rich Results Test[27] can assess if your website supports rich results based on the structured data it contains. The Schema Markup validator[28]https://validator.schema.org/ validates all Schema.org-based structured data that's embedded in websites.
Revise structured data	Google Analytics	Google Analytics[29] allows you to assess how many users are accessing your website, how long people stay on your website, click-through rate (CTR), and other metrics that can assess user engagement.

CONCLUSION

Sourcing accurate, up-to-date health information online remains a complex task for technically inexperienced patient users. Patients use hospital websites as their first point of contact to seek out healthcare and medical information, so it is important to improve website visibility, accessibility, and engagement necessary to connect patient users with target information. To address this problem, implementation of structured data in healthcare websites is a promising but overlooked method to improve visibility of healthcare facilities and accurate health information from reputable online websites. Moreover, structured data poses the unique potential to implement rich features in search engine results, which can serve as access points to providing information in engaging mediums for online patient users.

In this article, we propose a general framework to implement structured data in websites by 1) identifying the relevant schema type for your organization, 2) identifying relevant structured data

properties for target audiences, 3) encoding structured data in a machine-readable schema format, and 4) measuring SEO and patient-reported outcomes for quality improvement.

Future quality improvement and implementation science research is needed to promote the uptake of structured data to deliver personalized healthcare information to online patient users.

To promote the broad implementation of structured data in practice, there remains the need for centralized oversight to ensure that the implementation of structured data of healthcare websites is standardized and accurate over time. Further development of tools to enable schema development for non-technical users and implementation of standardized schemas for healthcare facility websites nationally can meaningfully improve the accessibility of healthcare services and information in the age of digital medicine.

REFERENCES

1 Gale JJ, Black KC, Calvano JD, et al. An Analysis of US Academic Medical Center Websites: Usability Study. *Journal of medical Internet research* 2021; *23*(12): e27750. https://doi.org/10.2196/27750

2 Jia X, Pang Y, & Liu LS. Online Health Information Seeking Behavior: A Systematic Review. *Healthcare (Basel, Switzerland)* 2021; *9*(12): 1740. https://doi.org/10.3390/healthcare9121740

3 Sbaffi L, & Rowley J. Trust and Credibility in Web-Based Health Information: A Review and Agenda for Future Research. *Journal of medical Internet research* 2017; *19*(6): e218. https://doi.org/10.2196/jmir.7579

4 *The Federal Register*. Federal Register :: Request Access. (n.d.). https://www.federalregister.gov/index/2021/centers-for-medicare-medicaid-services

5 Boyer C, Selby M, Scherrer JR, et al. The Health On the Net Code of Conduct for medical and health Websites. *Computers in biology and medicine* 1998; *28*(5): 603–610. https://doi.org/10.1016/s0010-4825(98)00037-7

6 *Guide to HHS surveys and Data Resources*. ASPE. (n.d.). https://aspe.hhs.gov/reports/guide-hhs-surveys-data-resources

7 Camargo A, Boggs KM, Auerbach M, et al. National Study of Self-reported Pediatric Areas in United States General Emergency Departments. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine* 2018; *25*(12): 1458–1462. https://doi.org/10.1111/acem.13633

8 Google. (n.d.). Google. https://developers.google.com/search/docs/appearance/structured-data/search-gallery

9 Google. (n.d.-a). Google. https://developers.google.com/search/docs/appearance/structured-data/intro-structured-data

10 Zilincan J. Search engine optimization. *CBU International Conference Proceedings* 2015; *3*: 506–510. https://doi.org/10.12955/cbup.v3.645

11 Scott D. White hat search engine optimization (SEO): Structured web data for Libraries. *Partnership: The Canadian Journal of Library and Information Practice and Research* 2015; *10*(1). https://doi.org/10.21083/partnership.v10i1.3328

12 Li J, & Zaïane OR. Combining Usage, Content, and Structure Data to Improve Web Site Recommendation. In: Bauknecht, K., Bichler, M., Pröll, B. (eds) E-Commerce and Web Technologies. EC-Web 2004. *Lecture Notes in Computer Science, vol 3182* 2004. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30077-9_31

13 *Schema.org: Evolution of structured data on the web*. Schema.org: Evolution of Structured Data on the Web - ACM Queue. (n.d.). https://queue.acm.org/detail.cfm?id=2857276

14 Abbasi BU, Fatima I, Mukhtar H, et al. Autonomous schema markups based on Intelligent Computing for search engine optimization. *PeerJ Computer Science* 2022; *8*. https://doi.org/10.7717/peerj-cs.1163

15 McCray AT. An upper-level ontology for the biomedical domain. *Comp Funct Genomics*. 2003;4(1):80-84. https://doi/org/10.1002/cfg.255

16 Dunne S, Cummins NM, Hannigan A, et al. A method for the design and development of medical or health care information websites to optimize search engine results page rankings on Google. *Journal of medical Internet research* 2013; *15*(8): e183. https://doi.org/10.2196/jmir.2632

17 *Emergency medicine*. Massachusetts General Hospital. (n.d.-b). https://www.massgeneral.org/emergency-medicine

18 *Awards and recognitions*. Massachusetts General Hospital. (n.d.). https://www.massgeneral.org/quality-and-safety/awards

19 *Contact the Department of Emergency Medicine*. Massachusetts General Hospital. (n.d.-b). https://www.massgeneral.org/emergency-medicine/contact

20 *Trauma center*. Massachusetts General Hospital. (n.d.-f). https://www.massgeneral.org/surgery/trauma-center

21 *Divisions, centers and programs*. Massachusetts General Hospital. (n.d.-c). https://www.massgeneral.org/emergency-medicine/divisions-centers-and-programs

22 *Patient care*. Billing | Mass General Brigham. (n.d.). https://www.massgeneralbrigham.org/en/patient-care/patient-visitor-information/billing/

23 *Events*. Massachusetts General Hospital. (n.d.-c). https://www.massgeneral.org/events

24 Google. (n.d.). Google maps. https://www.google.com/maps

25 Visiting Guidelines - Massachusetts General Hospital. (n.d.). https://www.massgeneral.org/assets/mgh/pdf/visit/visit-tips-sheet.pdf

26 *JSON-LD Generator: JSON-LD schema generator tool*. LD. (2019, May 6). https://jsonld.com/json-ld-generator/

27 Google. (n.d.). *Rich results test*. Google Search Console. https://search.google.com/test/rich-results

28 *Schema markup validator*. Schema Markup Validator. (n.d.-a). https://validator.schema.org/

29 Google. (n.d.-b). Google. https://analytics.google.com/analytics/web/

30 *Jsonld.js*. JSON. (n.d.). https://json-ld.org/

31 MozDevNet. (n.d.). *Microdata - HTML: Hypertext markup language: MDN*. HTML: HyperText Markup Language | MDN. https://developer.mozilla.org/en-US/docs/Web/HTML/Microdata

32 *RDF 1.2 schema*. W3C. (n.d.). https://www.w3.org/TR/rdf12-schema/