

# بررسی کیفیت داده‌های مکانی داوطلبانه از منظر شاخص، با تاکید بر شاخص برازندگی استفاده

زهرا پاشائی<sup>۱</sup>، محمدرضا ملک<sup>۲\*</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه

نصیرالدین طوسی

z\_pashaei69@yahoo.com

<sup>۲</sup> دانشیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

mrmalek@kntu.ac.ir

(تاریخ دریافت آذر ۱۳۹۵، تاریخ تصویب اردیبهشت ۱۳۹۶)

## چکیده

اطلاعات مکانی داوطلبانه گونه‌ای از اطلاعات مکانی است که به صورت اختیاری، توسط افراد عادی، بدون نیاز به آموزش علمی و با توجه به دانش محلی تهیه می‌شود. اطلاعات مکانی داوطلبانه در مقایسه با داده‌های مکانی استاندارد، دارای مزایای فراوانی از جمله دسترسی آسان به اطلاعات و قابلیت به‌روزرسانی سریع است، ولی کیفیت این داده‌ها یکی از موانع استفاده گسترده از این داده‌ها و محل بحث و تحقیق بسیاری از محققان بوده است. کیفیت داده‌های مکانی داوطلبانه نقش مهمی در به‌کارگیری آن برای کاربردهای مختلف دارد. استفاده از ابزارها و روش‌های مناسب برای برآورد کیفیت داده‌های مکانی داوطلبانه و نیز بهبود بخشیدن به آن امری مهم می‌باشد. این مهم موجب افزایش میزان استفاده از این گونه داده‌ها می‌شود. یکی از روش‌های ممکن برای ارزیابی کیفیت، به‌ویژه هنگامی که نتوان از داده مرجع سود جست، تعیین شاخص‌ها و برآورد آن‌ها است. در این مقاله سعی شده است در ابتدا مروری بر کیفیت اطلاعات مکانی صورت گیرد و سپس کیفیت اطلاعات مکانی داوطلبانه از منظر شاخص‌ها و روش‌های ارزیابی‌کننده آن بررسی گردد. همچنین شاخص برازندگی ژرفناکانه تشریح شده است.

**واژگان کلیدی:** کیفیت اطلاعات مکانی، اطلاعات مکانی داوطلبانه، شاخص، برازندگی استفاده

## ۱- مقدمه

با پیشرفت فناوری اطلاعات، گسترش اینترنت و فراگیر شدن استفاده از تجهیزات همراه مجهز به امکانات تعیین موقعیت جهانی، انگیزه‌های زیادی برای تولید اطلاعات مکانی توسط مشارکت‌کنندگان ایجاد شد. این امر موجب گردید نقش کاربران از مصرف‌کننده صرف، به تولیدکننده-مصرف کننده تغییر یابد. این نوع رویکرد در تولید اطلاعات مکانی باعث ایجاد مفهوم جدیدی به نام اطلاعات مکانی داوطلبانه شد که برای اولین بار توسط Goodchild (۲۰۰۷) مطرح گردید. او داده‌های مکانی داوطلبانه را این‌گونه تعریف کرده است: "توانایی کاربران وب برای تولید محتوا که بعداً یکپارچه می‌شوند و بر روی سایت‌ها قرار می‌گیرند." [۱۵]

اطلاعات مکانی داوطلبانه یک منبع اطلاعاتی جدید و در حال رشد به شمار می‌آید که به کاربران اجازه می‌دهد تا دنیا را بر اساس درک و چشم‌انداز خود تصویر کنند [۱،۲۸]. این اطلاعات، از این جهت اهمیت یافتند که مزایای بسیاری از جمله دسترسی آسان به اطلاعات، قابلیت به‌روزرسانی سریع، امکان اصلاح خطای احتمالی نسبت به داده‌های استاندارد که پس از انتشار در بازه زمانی خاص به‌روز می‌شوند، دارند؛ اما از آنجایی که افراد غیرمتخصص تولیدکننده این داده‌ها هستند، هیچ نظارتی بر نحوه تولید و یا منبع مورداستفاده در تولید آن‌ها وجود ندارد؛ بنابراین کیفیت داده‌های تولیدشده مشخص نیست و از ابتدای پیدایش این داده‌ها محل بحث و تحقیق بسیاری از محققان بوده است [۲۸].

در آغاز تمام زمینه‌های تحقیقاتی جدید معمولاً مشکلاتی پیش‌رو است، اطلاعات مکانی داوطلبانه نیز به‌عنوان یک زمینه تحقیقاتی جدید با مسائل و مشکلاتی از جمله کیفیت مواجه است. کیفیت داده‌های مکانی داوطلبانه نقش مهمی در به‌کارگیری آن در کاربردهای مختلف دارد. تعریف و استفاده از ابزارها و روش‌های مناسب برای برآورد کیفیت داده‌های داوطلبانه و نیز بهبود آن در افزایش میزان استفاده از این داده‌ها مؤثر است به‌خصوص در مواردی که دسترسی به داده‌های مرجع سخت و یا هزینه‌بر باشد [۳۰].

معمولاً ارزیابی کیفیت از طریق برآورد شاخص‌های آن امکان‌پذیر است، به‌طوری‌که پس از مشخص شدن کیفیت یک مجموعه داده از طریق شاخص‌های<sup>۱</sup> آن، می‌توان در

رابطه با "برازندگی استفاده"<sup>۲</sup> به‌عنوان شاخص کلیدی ارزیابی کیفیت اطلاعات مکانی بحث کرد. در واقع اغلب شاخص‌های کیفیت در قالب برازندگی استفاده معنا پیدا می‌کنند. برازندگی استفاده به نیاز کاربران و هدف در نظر گرفته‌شده برای پروژه بستگی دارد و از پروژه‌های به پروژه‌های دیگر و کاربردی به کاربرد دیگر متغیر است [۳۷]. همان‌طور که Longhorn و Blakemore [۲۷] بیان کرده‌اند، ارزش یک مجموعه داده معین برای اشخاص و سازمان‌های مختلف به علت تفاوت در نیازها و اهداف متفاوت است. این مورد را به خصوصیات و ویژگی‌های بیرونی اطلاعات مکانی ارجاع داده‌اند و ارزش نسبی داده‌ها را به پتانسیل رفع نیاز و متناسب بودن با اهداف و انتظارات کاربران، مربوط دانستند. در واقع عبارت برازندگی استفاده میزان تناسب پایگاه داده‌ی مشخص را برای یک کاربرد خاص بیان می‌کند و در نظر گرفتن این موضوع در هنگام رویارویی با تمام مجموعه داده مکانی داوطلبانه امری ضروری به حساب می‌آید [۲۸].

با وجود اینکه محققان بسیاری به اهمیت این شاخص کیفیت اشاره کرده‌اند، اما تحقیقات اندکی تفاوت‌های موردنیاز برای کیفیت در کاربردهای مختلف را آنالیز و بررسی کرده‌اند [۳۵]. بنابراین در این مقاله در نظر است بعد از مروری بر مفهوم کیفیت، شاخص‌ها و روش‌های ارزیابی‌کننده آن‌ها، کیفیت از دیدگاه برازندگی استفاده بر پایه مقالات و تحقیقات صورت گرفته موردبررسی قرار گیرد.

ساختار مطالبی که در تحقیق حاضر بیان گردیده به این شرح است که در بخش دوم به کیفیت داده‌های مکانی، مفهوم کیفیت و تعریف تعدادی از شاخص‌ها پرداخته شده است. در بخش سوم، کیفیت داده‌های مکانی داوطلبانه، طبقه‌بندی کیفیت، روش‌های ارزیابی کیفیت و به‌طورکلی روش‌های موجود برای ارزیابی شاخص‌ها گردآوری شده و در ادامه کیفیت از دیدگاه برازندگی استفاده بررسی شده است. همچنین بخش آخر به ارائه جمع‌بندی و پیشنهاد اختصاص پیدا کرده است.

## ۲- کیفیت داده‌های مکانی

موضوع درک کیفیت پایگاه داده‌های مکانی، چندین سال پیش شناخته شد و توجه بسیاری از نقشه‌برداران، کارتوگراف‌ها و جغرافی‌دان‌ها را به خود جلب کرد. کارهایی

<sup>۲</sup> Fitness for Use

<sup>۱</sup> Elements

علاوه بر عناصر مذکور، سه عنصر هدف<sup>۷</sup>، کاربرد<sup>۸</sup> و پیشینه<sup>۹</sup> به‌عنوان معیارهای غیر کمی کیفیت برای موارد موردنیاز معرفی شده است [۲۶]؛ اما در زمینه کیفیت اطلاعات مکانی مفهوم "برازندگی استفاده" برای اولین بار توسط Jakobsson و Tsoulos [۲۳] مطرح شد. این مفهوم بیانگر این است که یک مجموعه داده ممکن است با وجود داشتن کیفیت مناسب برای یک کاربرد، برای کاربردی دیگر، کیفیت کافی را نداشته باشد. بنابراین مسئله کیفیت به‌صورت نسبی و از کاربردی به کاربرد دیگر تغییرپذیر است. و به طور کلی می‌توان نتیجه گرفت که "برازندگی استفاده" به نیاز کاربران و هدف در نظر گرفته شده برای پروژه وابسته است [۱۷]. از طرفی واحدی [۳۷] بیان می‌کند: "بیشتر شاخص‌های کیفیت در قالب برازندگی استفاده معنا می‌یابند." به طور مثال ممکن است برای کاربرد مسیریابی، کامل بودن و دقت هندسی راه‌ها اهمیت به‌سزایی داشته، درحالی‌که در دیگر کاربردها اهمیت چندانی نداشته باشد. بنابراین بهتر است کاربران در هنگام تولید داده‌های مکانی، کیفیت داده‌ها را در قالب شاخص‌های کیفیت ارائه کنند تا دیگر کاربران در هنگام استفاده، از میزان مناسب بودن داده‌ها برای اهداف خود اطمینان حاصل نمایند.

### ۳- کیفیت داده‌های مکانی داوطلبانه

اهمیت داده‌های مکانی داوطلبانه مبتنی بر این حقیقت است که مردم به خصوص ساکنین یک محل بیشترین شناخت نسبت به توصیف منطقه زندگی خود دارند و عموماً این افراد بدون دانش نقشه‌برداری هستند و به جای فراداده، روش‌های بصری همانند تغییر در رنگ و ضخامت خطوط نقشه را برای نمایش کیفیت داده‌ها بکار می‌برند. درحالی‌که در روش‌های استاندارد، علاوه بر نظارت افراد متخصص بر نحوه تولید و کیفیت داده‌ها، از فراداده به منظور توصیفات تکمیلی داده‌ها استفاده می‌گردد. بنابراین کیفیت در محیط‌های اطلاعات مکانی داوطلبانه نامعلوم و محل بحث و تحقیق بسیاری از محققان بوده است [۱]. در واقع کیفیت اطلاعات مکانی داوطلبانه یک ویژگی ترکیبی است به‌طوری‌که نه تنها معیارهایی که با ویژگی‌های داده سروکار دارد را در نظر می‌گیرد، بلکه ویژگی‌های تولیدکنندگان داده را نیز در برمی‌گیرد [۳۸].

<sup>۷</sup> Purpose  
<sup>۸</sup> Usage  
<sup>۹</sup> Lineage

که در زمینه کیفیت اطلاعات مکانی صورت گرفته به اواخر دهه ۱۹۶۰ و اوایل دهه ۱۹۷۰ برمی‌گردد. با ظهور سیستم‌های اطلاعات مکانی<sup>۱</sup> در دهه ۱۹۸۰، این زمینه تحقیقاتی به‌سرعت رشد کرد به‌طوری‌که توجه بسیاری از محققان را مجذوب خود ساخت [۳۸].

در این بخش کیفیت داده‌های مکانی به‌عنوان یکی از مهم‌ترین مسائل مربوط به داده‌ها، مورد بررسی قرار می‌گیرد.

### ۲-۱- مفهوم کیفیت

برداشت‌های مختلفی از مفهوم کیفیت داده وجود دارد، به‌طوری‌که برای برخی از کاربران، داده باکیفیت بالا، داده‌ای است که هیچ خطایی نداشته باشد، در صورتی‌که برای عده‌ای دیگر، برآورده کردن انتظار مصرف‌کننده کافی است [۱۰، ۳۱]. استاندارد ISO 19113 [۲۱] کیفیت را این‌گونه تعریف می‌کند: "یک مجموعه داده دربرگیرنده خصوصیات و ویژگی‌هایی است که بتواند نیازهای روشن و ضمنی کاربران را برآورده سازد." به مجموعه ویژگی‌های یک محصول شاخص یا معیار گفته می‌شود [۳۷].

### ۲-۲- شاخص‌های کیفیت

Kress و Fadaie [۲۶] معیارهای کیفیت را بر اساس استاندارد تدوین شده توسط ISO19113 و ISO19114 در قالب پنج عنصر تمامیت، سازگاری منطقی، دقت هندسی، دقت زمانی، دقت موضوعی بررسی می‌کند:

تمامیت<sup>۲</sup>: نشان‌دهنده‌ی وجود و عدم وجود عوارض، اطلاعات توصیفی و روابط آن‌هاست.

سازگاری منطقی<sup>۳</sup>: نشان‌دهنده‌ی میزان مطابقت با قوانین منطقی ساختار داده، اطلاعات توصیفی و روابط است.

دقت مکانی<sup>۴</sup>: نشان‌دهنده‌ی دقت موقعیت عوارض مکانی است.

دقت زمانی<sup>۵</sup>: نشان‌دهنده‌ی دقت روابط و توصیف زمانی عوارض است.

دقت موضوعی<sup>۶</sup>: نشان‌دهنده‌ی توصیف کمی و غیر کمی و طبقه‌بندی و روابط عوارض است.

<sup>۱</sup> GIS  
<sup>۲</sup> Completeness  
<sup>۳</sup> Logical Consistency  
<sup>۴</sup> Positional Accuracy  
<sup>۵</sup> Temporal Accuracy  
<sup>۶</sup> Thematic Accuracy

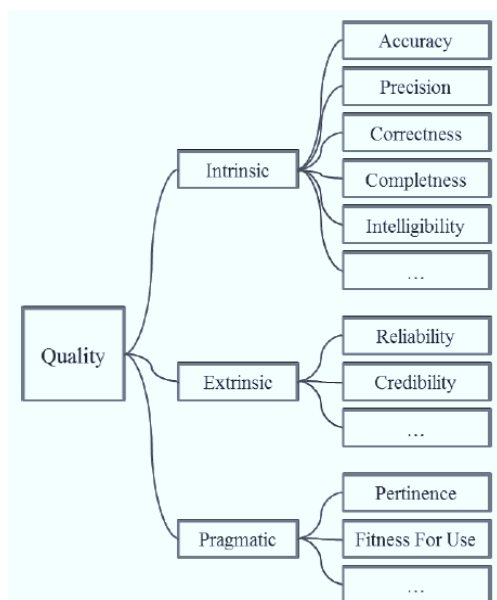
### ۳-۱- طبقه‌بندی کیفیت داده‌های مکانی داوطلبانه

ISO 19113 [۲۱] کیفیت داده‌های مکانی داوطلبانه را در دودسته درونی<sup>۱</sup> و بیرونی<sup>۲</sup> تعریف می‌کند. دسته درونی به ویژگی‌های درونی اطلاعات همانند دقت مکانی، دقت زمانی، دقت معنایی و غیره مربوط می‌شود درحالی‌که دسته دوم با "برازندگی استفاده" اطلاعات سروکار دارد. شکل ۱ اولین طبقه‌بندی از کیفیت داده‌های مکانی داوطلبانه را نشان می‌دهد.

Intrinsic quality	Extrinsic quality
Positional Accuracy	<b>Fitness For Use</b>
Temporal Accuracy	
Thematic Accuracy	
Semantic Accuracy	
Completeness	
Lineage, ...	

شکل ۱- طبقه‌بندی کیفیت داده‌های مکانی داوطلبانه از دو بعد درونی و بیرونی [۲۱]

بر اساس نظر Criscuolo و همکاران [۸] این دسته‌بندی برای ارزیابی کیفیت بخشی از داده و یا کل مجموعه داده ضروری است اما بدلیل اینکه سومین معیار کیفیت یعنی قابلیت اعتماد<sup>۳</sup> را پوشش نمی‌دهد، کامل نمی‌باشد. بدین منظور کیفیت این داده‌ها را مطابق شکل ۲ توصیف کردند.



شکل ۲- طبقه‌بندی کیفیت از سه بعد: درونی، بیرونی، عملی [۸]

همانطور که در شکل ۲ دیده می‌شود؛ کیفیت درونی به ویژگی‌ها و محتویات اطلاعات بستگی دارد، کیفیت بیرونی به قابلیت اطمینان و اعتبار تولیدکنندگان وابسته است و کیفیت عملی<sup>۴</sup> به قابلیت برآورده کردن نیازهای یک کاربر یا یک کاربرد بستگی دارد و به عبارتی بیانگر برازندگی استفاده است [۸].

### ۳-۲- ارزیابی کیفیت داده‌های مکانی داوطلبانه

OpenStreetMap به‌عنوان یکی از پرکاربردترین پروژه‌های VGI<sup>۵</sup> مبنا محسوب می‌شود که هدف آن به‌طور خاص تهیه و تولید اطلاعات مکانی رایگان مانند نقشه خیابان‌ها برای عموم مردم است. داده‌های سایت OSM، به‌عنوان یک پروژه VGI مبنا، به‌منظور ارزیابی کیفیت این داده‌ها در اکثر تحقیقات استفاده شده است.

تاکنون تحقیقات زیادی به ارزیابی کیفیت داده‌های مکانی داوطلبانه پرداخته‌اند و روش‌های متفاوتی از جمله مقایسه داده‌های VGI با داده‌های استاندارد، استفاده از شاخص‌های درونی، استفاده از نظرات کاربران و سایر روش‌ها را بکار برده‌اند. در این قسمت به برخی از تحقیقات انجام‌شده، اشاره‌ای خواهد شد.

### ۳-۲-۱- مقایسه داده‌های VGI با داده‌های استاندارد

مطالعات بسیاری در زمینه اندازه‌گیری کمی کیفیت داده‌های داوطلبانه برای مناطق مختلف، از طریق مقایسه آن‌ها با مجموعه داده‌های با دقت بالا، انجام شده است. این تحقیقات، چشم‌اندازی از کیفیت داده‌های داوطلبانه را برای مناطق مختلف فراهم می‌سازد.

محققان بسیاری برای ارزیابی دقت مکانی، دقت موضوعی و تمامیت، داده‌های استاندارد را برای مقایسه با داده‌های VGI بکار برده‌اند [۴، ۱۲، ۱۴، ۱۹، ۳۶، ۲۴، ۳۰]. به‌طور مثال Broveli و همکاران [۶]، به بررسی کیفیت ساختمان‌های موجود در نقشه OSM شهر میلان واقع در شمال ایتالیا، پرداختند. در این تحقیق برخلاف بسیاری از تحقیقات علاوه بر بررسی تمامیت، دقت مکانی نیز مورد ارزیابی قرار گرفت. Fram [۱۳] به بررسی کیفیت عوارض مسطحاتی در برخی از شهرهای انگلستان پرداخت. برای

۴ Practical Quality  
 ۵ Volunteer Geographic Information

۱ Intrinsic Quality  
 ۲ Extrinsic Quality  
 ۳ Trustworthiness

تصاویر استریو دقت مکانی عوارض راه OSM را مورد بررسی قرار دادند.

نکته قابل توجه در مورد روش‌های ذکر شده این است که هرکدام مزایا و معایبی دارند و نمی‌توان با قاطعیت دسته‌ای را از دسته‌ای دیگر برتر دانست. به‌طور مثال در مورد دسته اول همیشه نیاز به داده استاندارد وجود دارد و نتایج به‌دست‌آمده به داده‌های ورودی وابسته‌اند. از طرف دیگر در بسیاری از موارد داده‌های سیستم‌های داوطلبانه کامل‌تر از داده‌های استاندارد هستند، بنابراین در این‌گونه موارد مقایسه با داده استاندارد برآورد نادرستی از کیفیت را ارائه می‌دهد. ولی مزیت این روش ارزیابی کیفیت، این است که در مناطقی که داده‌ها همپوشانی دارند، نتایج بسیار دقیق خواهند بود. در مورد دسته دوم تحقیقات کمی انجام شده و این روش تا حدی ناپخته به نظر می‌رسد. از معایب دسته سوم نیز می‌توان به کمبود داده اشاره کرد.

به نظر می‌رسد مطلوب‌ترین حالت این است که ترکیبی از این روش‌ها برای ارزیابی کیفیت استفاده گردد. به‌طور مثال، در تحقیق صورت گرفته توسط Mohammadi و Malek، ترکیبی از روش اول و دوم برای بهبود سازگاری منطقی [۳۰،۳۱] استفاده شده است. در این تحقیق ابتدا با استفاده از روش تناظرایی نقطه مبنا، عوارض استاندارد و داوطلبانه متناظر در دو مجموعه داده مشخص گردید. سپس پارامترهای ذاتی، مکانی، زمانی و کاربر برای کیفیت هندسی معرفی شدند. پس‌از آن، با استفاده از یک روش ترکیبی بر مبنای هوش مصنوعی، رابطه میان هر یک از این پارامترها با شاخص دقت هندسی مشخص گردید. در نهایت با استفاده از این روابط، دقت هندسی داده‌های داوطلبانه بدون متناظر، محاسبه شد. در جدول ۱، شاخص‌ها، زیرشاخص‌ها و روش ارزیابی آن‌ها که می‌توانند برای ارزیابی دقت کلی مجموعه داده VGI در نظر گرفته شوند، جمع‌بندی شده است.

این منظور داده‌های OSM با داده‌های استاندارد تهیه شده توسط OS<sup>۱</sup> مقایسه و تمامیت به‌عنوان شاخص برآورد کیفیت ارزیابی گردید. نتایج این تحقیق، تمامیت متفاوت داده‌های OSM بین شهرهای انگلستان را نشان می‌دهد.

### ۳-۲-۲- استفاده از شاخص‌های درونی مجموعه داده

تحقیقات انجام شده در این زمینه باوجود معرفی مفاهیم، پارامترها و شاخص‌های مختلف برای اندازه‌گیری کیفیت، مدل مشخصی از تأثیر آن‌ها بر کیفیت داده و چگونگی استفاده از آن‌ها برای ارزیابی کیفیت ارائه ندادند. به‌طور مثال Ciepluch و همکاران [۹] استفاده از شاخص‌هایی مانند توزیع مکانی عوارض و تعداد عوارض و نقاط را برای اندازه‌گیری کیفیت داده‌های OSM، پیشنهاد دادند. Barron و همکاران [۵] مجموعه‌ای از شاخص‌ها را برای ارزیابی و تحلیل کیفیت داده‌های OSM، بدون نیاز به داده‌های استاندارد، معرفی کردند.

### ۳-۲-۳- استفاده از نظرات کاربران

دسته سوم تحقیقات، به ارزیابی اعتبار کاربران و مشارکت‌کنندگان در پروژه‌های داوطلبانه می‌پردازد و براساس کیفیت اطلاعات وارد شده از طرف هر کاربر، سایر کاربران به او امتیاز می‌دهند و برآیند این امتیازها، کیفیت کار هر کاربر را نشان می‌دهد. به‌طور مثال Adler و Alfaro [۲] یک سیستم اعتبار برای سایت Wikimapia طراحی کردند که بر اساس تعداد ویرایش‌های انجام شده روی اطلاعات وارد شده از سوی یک کاربر به او امتیاز می‌دهد.

### ۳-۲-۴- سایر روش‌ها

در این دسته از تحقیقات، از روش‌های گوناگونی برای بررسی کیفیت استفاده شده است. به‌طور مثال Haklay [۱۸] قانون لینوس<sup>۲</sup> را برای بررسی کیفیت داده‌های داوطلبانه بکار برد و نشان داد هر چه تعداد مشارکت‌کنندگان در واحد زمانی افزایش یابد، کیفیت این نوع داده‌ها نیز بهبود می‌یابد. Canavosio-Zuzelski و همکاران [۷]، با استفاده از مدل سرشکنی برداری و

<sup>۱</sup> Ordnance Survey  
<sup>۲</sup> Linus

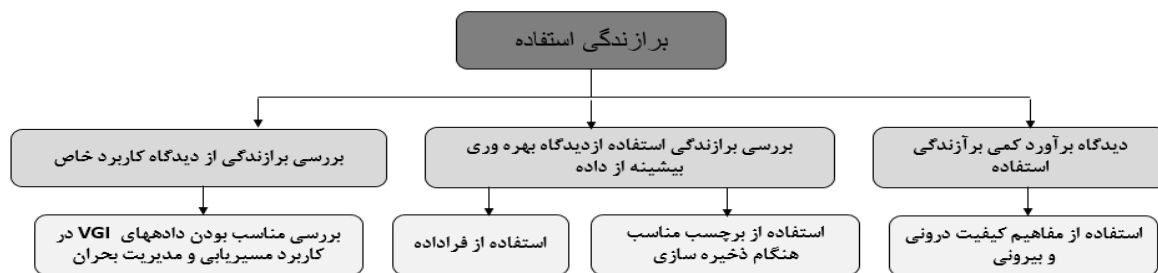
جدول ۱- شاخص‌ها درونی کیفیت، زیر شاخص‌ها و روش ارزیابی [۲۱، ۲۲]

شاخص‌های کیفیت	زیر شاخص‌های کیفیت	روش ارزیابی
تمامیت	خطای مازاد	مقایسه تعداد آیتم‌های مجموعه داده VGI با تعداد آیتم‌ها در مجموعه استاندارد
	خطای حذف	
سازگاری منطقی	سازگاری مفهومی	شمارش تعداد عوارض و روابطی که شمای مفهومی را نقض کند
	سازگاری فرمت	مقایسه ساختار ثبت‌شده برای همه موارد با تعریف صحیح
	سازگاری توپولوژی	چک کردن محدوده‌ها از نظر بهم‌رسیدگی و تکرار، شمارش تعداد تناقضات
دقت هندسی	دقت مطلق	محاسبه خطای مطلق/ نسبی فاصله برای هر نقطه از طریق مقایسه مقادیر مختصات مطلق/ نسبی هر نقطه در دو مجموعه VGI و استاندارد، محاسبه RMS از خطای فاصله
	دقت نسبی	
دقت زمانی	دقت اندازه‌گیری زمانی	اندازه‌گیری تفاوت بین وقوع در دو مجموعه داده VGI و استاندارد، محاسبه RMS از تفاوت زمانی
	اعتبار زمانی	بررسی و تأیید صحیح بودن زمان اکتساب
	سازگاری زمانی	تأیید ترتیب زمانی
دقت موضوعی	صحت طبقه‌بندی	مقایسه طبقه‌بندی آیتم با طبقه صحیح با استفاده از ضریب کاپا
	صحت توصیفات غیر کمی	مقایسه توصیفات غیر کمی با کل مجموعه
	صحت توصیفات کمی	اندازه‌گیری تفاوت بین توصیفات کمی نسبت به کل مجموعه، محاسبه RMS از تفاوت

### ۳-۳- کیفیت از دیدگاه برازندگی استفاده

امروزه، موضوع تعیین کیفیت به منزله اندازه‌گیری خطای یک مجموعه داده عنوان می‌گردد و کمتر به این بحث پرداخته می‌شود که یک مجموعه داده با یک کیفیت مشخص، مناسب چه نوع کاربردی است [۱۱]. بنابراین برازندگی استفاده، تنها مشخصه کیفیت از این نظر است. به منظور درک بهتر معیار برازندگی استفاده، در این قسمت، دیدگاه‌های موجود به سه دسته مطابق شکل ۳ طبقه‌بندی و مورد بررسی قرار می‌گیرند.

و اما شاخص "برازندگی استفاده"، مطابق با تعریف ISO1957 [۲۰]، به نیازهای کاربران و هدف در نظر گرفته‌شده برای پروژه وابسته است و به‌وسیله سایر شاخص‌های کیفیت قابل ارزیابی است. در برخی از تحقیقات اظهارشده است که ارزیابی داده‌های مکانی برای سنجش تناسب آن برای یک کاربرد خاص، اهمیت فراوانی دارد [۲۸]. بدلیل اهمیت بالای این شاخص، در ادامه کیفیت از دیدگاه برازندگی استفاده بحث خواهد شد.



شکل ۳- طبقه‌بندی برازندگی استفاده از سه دید

عملی داشت. در این راستا، تلاش‌هایی هم برای کمی‌سازی کیفیت عملی با استفاده از این مفهوم، انجام شده است. به طور مثال، واحدی [۳۷] تعدادی شاخص سطح متوسط و سطح پایین برای محاسبه کیفیت درونی و کیفیت بیرونی معرفی نمود و سپس با اعمال عملگر OWA<sup>۱</sup> روی این شاخص‌ها، مقدار کمی کیفیت عملی را محاسبه کرد.

<sup>۱</sup> Ordered Weighted Averaging

### ۳-۳-۱- محاسبه برازندگی استفاده از منظر کیفیت عملی

همانطور که در بخش ۳-۱ بیان شد، کیفیت عملی، بیانگر شاخص برازندگی استفاده است. از طرفی اغلب شاخص‌های کیفیت، چه شاخص‌های درونی و چه بیرونی، در قالب شاخص برازندگی استفاده معنا پیدا می‌کنند. بنابراین با بهره‌جستن از این مفاهیم، می‌توان برآورد کمی از کیفیت

### ۳-۳-۲- برزندگی استفاده از دیدگاه بهره‌وری بیشینه از داده

با توجه به جریان عظیمی که در تولید اطلاعات مکانی داوطلبانه بوجود آمده است، ضرورت بهترین استفاده از این داده‌ها اجتناب ناپذیر است. برای این منظور نیاز به مدیریت و بهره‌برداری از این جریان داده احساس می‌شود که از جمله آن می‌توان به ذخیره‌سازی و پردازش بهینه داده‌ها اشاره کرد [۱۱]. به‌طور مثال بعضی از پروژه‌های VGI همانند OSM، مشارکت‌کنندگان را به استفاده از اصطلاحات فنی از قبل تعریف شده، به‌منظور بهبود ذخیره‌سازی تشویق می‌کنند [۳۴].

فراداده<sup>۱</sup> به‌عنوان یک شاخص کلیدی برای بهره‌بری بیشینه از داده بکار می‌رود. نمونه آن را می‌توان در تحقیق انجام شده توسط Kalantari [۲۵] مشاهده کرد. در این تحقیق، به‌منظور غنی‌سازی توصیفات داده، فراداده داده‌های VGI به دو صورت ضمنی<sup>۲</sup> و صریح<sup>۳</sup> ایجاد شده‌است. در حالت ضمنی مشارکت‌کنندگان هیچ اطلاعی از مشارکت خود ندارند درحالی‌که در حالت صریح مشارکت‌کنندگان با آگاهی کامل اقدام به ایجاد فراداده و بهبود آن می‌کنند.

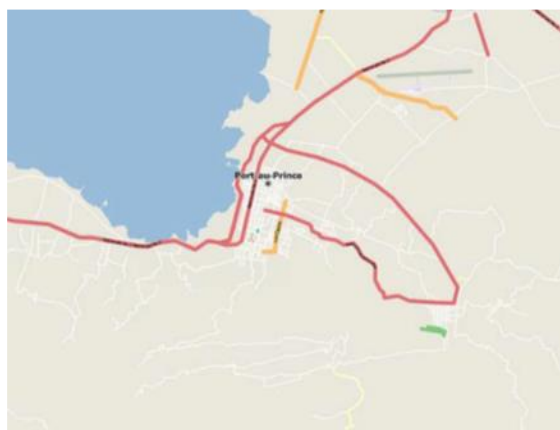
استفاده از برچسب مناسب در هنگام ذخیره‌سازی داده نیز می‌تواند در بهره‌برداری بیشینه از داده نقش مهمی داشته باشد [۴۰]. به‌همین منظور Devillers و Vandecasteele [۳۹] یک سیستم توصیه‌گر به‌نام OSMantic برای وب‌سایت OSM طراحی کردند. وظیفه این توصیه‌گر این است که کاربر را به استفاده از برچسب مناسب برای داده ورودی هدایت می‌کند. بعلاوه عشقی [۱۱] یک مدل مفهومی طراحی کرد که با استفاده از آن، داده‌های ورودی به پایگاه داده داوطلبانه، به صورت صحیح و بهینه ذخیره‌سازی می‌شوند، اما از آنجایی که ممکن است در هنگام استفاده، کاربرد متفاوتی از داده مدنظر کاربران باشد، سعی شده است با اعمال تمهیداتی در فرآیند ذخیره‌سازی، فرصتی برای بهره‌بری بیشینه از داده‌ها فراهم آید.

### ۳-۳-۳- برزندگی استفاده از دیدگاه کاربردی خاص

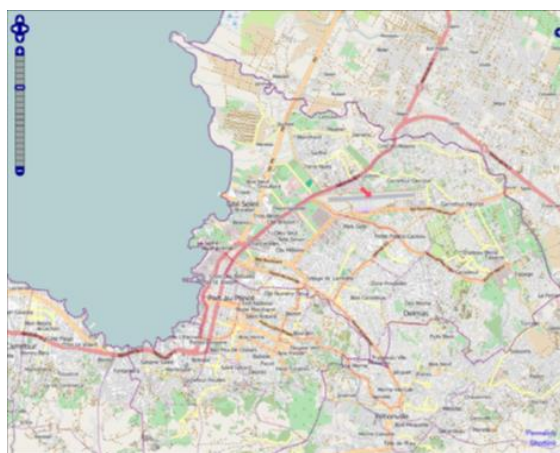
همانطور که پیش‌تر اشاره شد، اطلاعات مکانی داوطلبانه یک منبع اطلاعاتی جدید و در حال رشد برای

تهیه نقشه‌هایی دقیق از چشم‌اندازها به‌شمار می‌آید. در این قسمت، نقش داده‌های مکانی داوطلبانه و تناسب این داده‌ها برای کاربردهای مدیریت بحران و مسیریابی بیان می‌گردد.

نقش داده‌های مکانی داوطلبانه در جریان کمک‌رسانی به زلزله‌هاییتی در ژانویه ۲۰۱۰ به‌طور گسترده و با موفقیت دیده شد. به‌طوریکه تکنولوژی OSM توجه رسانه‌ها را بیش از پیش به خود معطوف ساخت [۴۱]. در جریان زلزله‌هاییتی، بسیاری از ارگان‌های امدادی بدلیل عدم دسترسی به اطلاعات بهنگام و دقیق، از اطلاعات مکانی که توسط مردم گردآوری شده‌بود برای کمک‌رسانی به نواحی آسیب‌دیده استفاده کردند. این امر به شناسایی بهتر نواحی نیازمند به کمک‌های اضطراری و تخصیص منابع منجر گردید. شکل ۴ حجم پایین داده‌های مکانی در سایت OSM قبل از زلزله‌هاییتی را نشان می‌دهد، درحالی‌که شکل ۵ بیانگر افزایش حجم داده‌های وارد شده توسط داوطلبان بعد از زلزله می‌باشد [۴۱].



شکل ۴- میزان اطلاعات درج شده در OSM قبل از زلزله [۴۱]



شکل ۵- اطلاعات درج شده بعد از زلزله در OSM [۴۱]

<sup>۱</sup> Metadata

<sup>۲</sup> Implicit

<sup>۳</sup> Explicit

تحقیق نشان داد که مدلسازی خسارات با استفاده از داده‌هایی که از طریق مردم گردآوری شده به نتایج مدلسازی هیدرولیکی نزدیک تر است. بنابراین می‌توان نتیجه گرفت، موضوع برازندگی استفاده داده‌های مکانی داوطلبانه در این نوع از موقعیت‌ها و در جاییکه هیچ داده استاندارد موجود نیست، بسیار کاربردی بنظر می‌رسد.

#### ۴- جمع‌بندی و پیشنهادات

در این مقاله کیفیت به‌عنوان یک از چالش‌های پیشروی اطلاعات مکانی داوطلبانه از جنبه‌های مختلف موردبررسی قرار گرفت. نخست به بیان مفاهیم و شاخص‌های کیفیت و طبقه‌بندی‌های مختلف شاخص‌ها پرداخته شد. در ادامه روش‌های ارزیابی این شاخص‌ها جمع‌بندی شد. سپس وجه قالب کیفیت یعنی برازندگی استفاده موردبحث قرار گرفت. با بررسی و جمع‌بندی تحقیقات صورت گرفته در این زمینه، می‌توان گفت که کمترین توجه به مسئله برازندگی استفاده در میان تحقیقات مشاهده می‌شود. بنابراین پرداختن به کیفیت از این دیدگاه و ارائه راهکارهایی برای بهره‌وری بیشینه از داده‌های مکانی داوطلبانه همانند شناسایی نیازهای داده‌ای یک سازمان و یا یک کاربر، کمک شایانی به این موضوع خواهد کرد.

اگرچه استفاده از این اطلاعات فواید زیادی دارد، اما بهتر است جنبه‌های منفی نیز درنظر گرفته شود [۳۲]، [۳۳]. برای مثال درصورتیکه از این اطلاعات برای تعیین موقعیت مصدومین استفاده شود و این اطلاعات نادرست باشند، منابع مالی و زمانی بسیاری به‌هدر خواهد رفت. برخلاف وجود این‌گونه عدم قطعیت‌ها، برخی از محققان اطلاعات مکانی داوطلبانه را در جریان بحران، به‌منظور تهیه اطلاعات مکانی و زمانی نسبتاً دقیق، کاملاً سودمند تلقی می‌کنند.

در تحقیق انجام شده توسط Mondzecz و Sester [۲۹] که به‌منظور تعیین مناسب‌ترین پایگاه داده در کاربرد مسیریابی صورت پذیرفت، دو پایگاه داده OSM و ATKIS، با اعمال الگوریتم کوتاهترین مسیر، مقایسه شدند. همچنین در مبحث مسیریابی، داده‌های VGI مبنا در سایت‌هایی همچون OpenRouteService.org و MappingRide.com برای سرویس‌دهی مسیریابی به‌کار می‌روند.

Poser و همکاران [۳۳] میزان مناسب بودن داده‌های مربوط به سیل گردآوری شده توسط مردم را با داده‌های اندازه‌گیری شده برای سیل سال ۲۰۰۲ در ایالت متحده Eilenburg مقایسه کردند. در این پژوهش، این دو مجموعه داده به عنوان ورودی، وارد مدل "تخمین خسارت سیل" شده تا میزان خسارت وارد بر ساختمان‌های مسکونی برآورد گردد. سپس نتایج تخمین خسارات این مدل با نتایج بدست آمده از مدلسازی هیدرولیکی مقایسه شدند. نتایج این

#### مراجع

- [1] محمدی نازیلا، ملک محمدرضا. محیط‌های اطلاعات مکانی مردم‌گستر: ویژگی‌ها و چالش‌ها. نشریه علمی ترویجی مهندسی نقشه‌برداری و اطلاعات مکانی. ۱۳۹۲؛ ۳(۴):۲۳-۳۴
- [2] Adler BT, De Alfaro L. A content-driven reputation system for the Wikipedia. In Proceedings of the 16th international conference on World Wide Web 2007 May 8 (pp. 261-270). ACM.
- [3] Antoniou V, Skopeliti A. Measures and indicators of VGI quality: An overview. ISPRS Ann. 2015 Aug 19.
- [4] Arsanjani JJ, Mooney P, Zipf A, Schauss A. Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets. In OpenStreetMap in GIScience 2015 (pp. 37-58). Springer International Publishing.
- [5] Barron C, Neis P, Zipf A. A comprehensive framework for intrinsic OpenStreetMap quality analysis. Transactions in GIS. 2014 Dec 1;18(6):877-95.
- [6] Brovelli MA, Minghini M, Molinari ME, Zamboni G. Positional Accuracy Assessment of the Openstreetmap Buildings Layer Through Automatic Homologous Pairs Detection: the Method and a Case Study. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2016 Jun:615-20.
- [7] Canavosio-Zuzelski R, Agouris P, Doucette P. A photogrammetric approach for assessing positional accuracy of OpenStreetMap© roads. ISPRS International Journal of Geo-Information. 2013 Apr 2;2(2):276-301.



- [8] Criscuolo L, Carrara P, Bordogna G, Pepe M, Zucca F, Seppi R, Oggioni A, Rampini A. Handling quality i crowdsourced geographic information. *European Handbook of Crowdsourced Geographic Information*.:57.
- [9] Ciepluch B, Mooney P, Winstanley AC. Building generic quality indicators for OpenStreetMap (2011).
- [10] Devillers R, Jeansoulin R. Spatial data quality: concepts. *Fundamentals of Spatial Data Quality*. 2006 Mar:31-42.
- [11] Eshghi M, Alesheikh AA. Introducing a Conceptual Model to Improve the Quality of Storage of Volunteered Geographic Information: In the Field of " Fitness-for-Use" Indicator. *Journal of Geomatics Science and Technology*. 2016 Oct 15;6(1):19-32.
- [12] Fan H, Zipf A, Fu Q, Neis P. Quality assessment for building footprints data on OpenStreetMap. *International Journal of Geographical Information Science*. 2014 Apr 3;28(4):700-19.
- [13] Fram C, Chistopoulou K, Ellul C. Assessing the quality of OpenStreetMap building data and searching for a proxy variable to estimate OSM building data completeness (2015).
- [14] Girres JF, Touya G. Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS*. 2010 Aug 1;14(4):435-59.
- [15] Goodchild MF. Citizens as sensors: the world of volunteered geography. *GeoJournal*. 2007 Aug 1;69(4):211-21.
- [16] Haklay M, Singleton A, Parker C. Web mapping 2.0: The neogeography of the GeoWeb. *Geography Compass*. 2008 Nov 1;2(6):2011-39.
- [17] Haklay M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and planning B: Planning and design*. 2010 Aug 1;37(4):682-703.
- [18] Haklay M, Basiouka S, Antoniou V, Ather A. How many volunteers does it take to map an area well? The validity of Linus' law to volunteered geographic information. *The Cartographic Journal*. 2010 Nov 1;47(4):315-22.
- [19] Helbich M, Amelunxen C, Neis P, Zipf A. Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata. *Proceedings of GI\_Forum*. 2012:24-33.
- [20] International Organisation for Standardisation, 2013. ISO19157:2013 Geographic information - Data quality, Geneva: ISO.
- [21] ISO/IEC IS 19113:2000. 2002. Geographic Information — Quality Principles. Geneva,Switzerland.
- [22] ISO/IEC IS 19114:2003. 2009. Geographic Information — Quality Evaluation Procedures. Geneva, Switzerland.
- [23] Jakobsson A, Tsoulos L. The Role of Quality in Spatial Data Infrastructures. InProceedings of the 23rd International Cartographic Conference, Moscow, Russia, Cd-Rom.–2007.
- [24] Jackson SP, Mullen W, Agouris P, Crooks A, Croitoru A, Stefanidis A. Assessing completeness and spatial error of features in volunteered geographic information. *ISPRS International Journal of Geo-Information*. 2013 Jun 4;2(2):507-30.
- [25] Kalantari M, Rajabifard A, Olfat H, and Williamson I. "Crowd-sourcing metadata for VGI." *AGIL – Leuven*, 2013, May 14-17.
- [26] Kresse W, Fadaie K. ISO standards for geographic information. *Springer Science & Business Media*; 2004 Jan 12.
- [27] Longhorn RA, Blakemore M. Geographic information: value, pricing, production, and consumption. *CRC Press*; 2007 Dec 17.
- [28] Meier JC. "An Analysis of Quality for Volunteered Geographic Information." (2015).
- [29] Mondzech J, Sester M. Quality analysis of OpenStreetMap data based on application needs. *Cartographica: The International Journal for Geographic Information and Geovisualization*. 2011 Jun;46(2):115-25.
- [30] Mohammadi N, Malek M. Artificial intelligence-based solution to estimate the spatial accuracy of volunteered geographic data. *Journal of Spatial Science*. 2015 Jan 2;60(1):119-35.
- [31] Mohammadi N, Malek M. VGI and Reference Data Correspondence Based on Location-Orientation Rotary Descriptor and Segment Matching. *Transactions in GIS*. 2015 Aug 1;19(4):619-39.
- [32] Okolloh O. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. *Participatory learning and action*. 2009 Jun 1;59(1):65-70.
- [33] Poser K, Dransch D. Volunteered geographic information for disaster management with application to rapid flood damage estimation. *Geomatica*. 2010;64(1):89-98.

- [34] Scheider S, Keßler C, Ortmann J, Devaraju A, Trame J, Kauppinen T, Kuhn W. Semantic referencing of geosensor data and volunteered geographic information. In *Geospatial semantics and the semantic web 2011* (pp. 27-59). Springer US.
- [35] Senaratne H, Mobasheri A, Ali AL, Capineri C, Haklay M. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*. 2016 Jun 2:1-29.
- [36] Tenney M. Quality evaluations on Canadian OpenStreetMap data. *Spatial knowledge and information* 2014.
- [37] Vahedi B, Alesheikh AA, Honarparvar S. Quantitative Assessment of Pragmatic Quality of Volunteered Geographic Information Using Fuzzy Linguistic Quantifiers and OWA Operator. *Journal of Geomatics Science and Technology*. 2014 May 15;3(4):65-76.
- [38] Van Oort PA. *Spatial Data Quality: From Description to Application*, Thesis (PhD), Wageningen Universiteit 2006.
- [39] Vandecasteele A, Devillers R. Improving volunteered geographic information quality using a tag recommender system: the case of OpenStreetMap. In *OpenStreetMap in GIScience 2015* (pp. 59-80). Springer International Publishing.
- [40] Zhang X, Ai T. How to Model Roads in OpenStreetMap? A Method for Evaluating the Fitness-for-Use of the Network for Navigation. In *Advances in Spatial Data Handling and Analysis 2015* (pp. 143-162). Springer International Publishing.
- [41] Zook M, Graham M, Shelton T, Gorman S. Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. *World Medical & Health Policy*. 2010 Jul 1;2(2):7-33.