arXiv:2303.01223v1 [cs.CY] 2 Mar 2023

# BikeDNA：自行车基础设施数据与网络评估工具

期刊标题 XX(X)：2–22©作者 2022 转载及许可信息：
sagepub.co.uk/journalsPermissions.nav 数字对象标识符：
10.1177/ToBeAssigned
www.sagepub.com/SAGE

**Ane Rahbek Vierø[1], Anastassia Vybornova[1], Michael Szell[1,2,3]**

**摘要**
高质量的现有自行车基础设施数据是基于证据进行自行车网络规划的必要条件，而此类规划有助于人类出行的绿色转型。然而，这一条件很少得到满足：政府机构或 OpenStreetMap 等众包项目提供的数据，往往存在质量未知、参差不齐或较低等问题。目前可用的道路网络数据质量评估工具，通常未能考虑网络拓扑、空间异质性以及自行车专属数据特征。为弥补这些不足，我们推出 BikeDNA——一款开源工具，专为自行车基础设施数据打造，可复现地开展质量评估，并重点关注网络结构和连通性。BikeDNA既可以独立分析单个数据集，也可以对 OpenStreetMap 与参考数据集进行比较分析，其中包括要素匹配。数据质量指标既针对整个研究区域进行全局评估，也在网格单元层级进行局部评估，从而揭示数据质量的空间变化。工具还会生成交互式地图和 HTML/PDF报告，便于直观地探索和交流结果。BikeDNA支持多种应用场景下的自行车基础设施数据质量评估，涵盖城市规划、OpenStreetMap 数据改进，以及面向可持续出行的网络研究。

**关键词**
自行车网络、空间数据质量、OpenStreetMap、志愿地理信息、免费开源软件

## 引言

全球各地的城市都在努力使其交通运输系统更具环境和社会可持续性。实现这一目标最具成本效益的方法之一是促进骑行，例如巴黎或波哥大所采取的做法´（C40城市 2019；巴黎市 2021年），这也符合欧盟委员会和 IPCC 的建议（欧洲委员会 2021年；Jaramillo 等，2022年）。然而，要实现骑行方式的大幅转变仍然面临挑战。自行车基础设施通常不足，相关数据的收集与提供也同样如此，而这些数据对于发挥数据驱动的自行车网络规划的潜力至关重要。近期研究表明，在网络层面对自行车基础设施进行定量分析，有助于规划决策，并能显著提升规划投资的成效（Natera Orozco 等2020年；Olmos等，2020年；Steinacker 等，2022年；Szell等，2022年；Vybornova 等，2022年），但这需要完整且高质量的基础设施数据。此外，即使能够获取此类数据，人们对其完整性和质量也往往知之甚少。行政数据和众包数据（例如 OpenStreetMap，OSM）都存在这一问题（Nelson 等，2021年；Rambøll 2022年）。

为弥补这一空白，并帮助研究人员、规划人员及该领域的其他人士评估自行车基础设施数据的质量，我们推出了 BikeDNA （自行车基础设施数据与网络评估）。BikeDNA 是一款使用 Python 编写的计算工具，可对一个或两个自行车基础设施数据集进行可复现的质量评估（见图 1）。BikeDNA 将“自行车基础设施”定义为道路网络中专供骑行的任何部分；用户可以修改此定义。BikeDNA 可对 OSM 或行政数据进行独立分析。如果 OSM 和行政数据都有，则 BikeDNA 会比较这两个数据集之间的差异。据我们所知，目前还没有其他专为自行车基础设施数据的特定性质量身打造的开源工具可用于数据质量评估。

## 关于空间数据质量的既有研究

虽然 BikeDNA 的设计兼容自行车基础设施的众包数据和行政数据，但鉴于 OSM 在自行车网络研究中被广泛使用，OSM 数据的质量评估是该工具的一项核心功能（Ferster

[1]哥本哈根信息技术大学，丹麦[2]ISI基金会[3]维也纳复杂性科学中心

**通讯作者：** Ane Rahbek Vierø，哥本哈根信息技术大学计算机科学系，电子邮箱：anev@itu.dk

使用以下工具编写*sagej.cls*

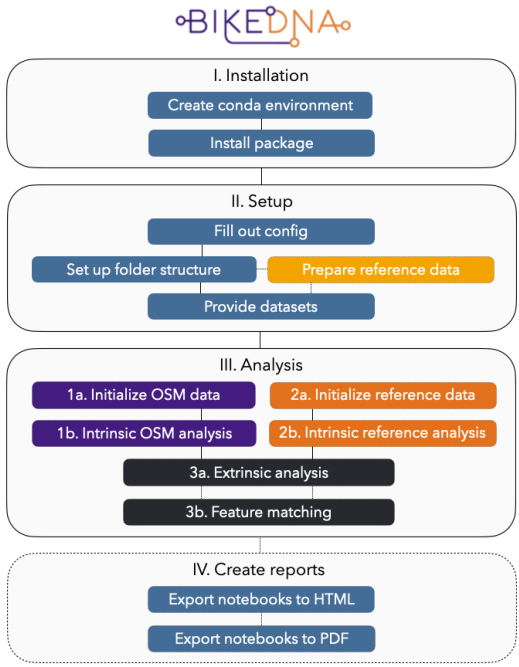


**图 1**。 BikeDNA 的流程。分析（III）分为三个部分：1）OSM 自行车网络数据的内在分析（紫色）；2）参考自行车网络数据的内在分析（橙色）；以及 3）OSM 与参考数据的比较（黑色），包括要素匹配。虚线部分为可选项。

等，2020 年；Nelson 等，2021年）。此外，近年来许多关于空间数据质量的研究都聚焦于众包或志愿者提供的数据质量，试图让人们信任来自非官方来源的数据集。因此，我们对既有研究的综述聚焦于众包道路网络数据的质量评估。关于志愿地理信息（VGI）质量评估方法更全面、深入的综述，请参阅Fonte (2017)、Senaratne 等(2017）和Degrossi 等(2018）。关于更广泛的空间数据质量解决方案概述，请参阅Medeiros 和 Holanda (2019）。

内部分析与外部分析 空间数据质量评估方法通常分为“内部”（独立）或“外部”（比较）两类，BikeDNA 也采用这种分类。内部方法研究数据集本身固有的属性，或数据集的创建方式；外部方法则利用外部参考数据集进行比较。早期关于志愿地理信息数据质量的研究主要采用外部评估，着重评估完整性和空间精度，使用局部指标衡量某个区域内制图网络的长度（Haklay 2010），以及更先进的要素匹配方法（Koukoletsos 等人

2012）。数据完整性和准确性虽是相关的数据质量指标，却无法反映拓扑质量，即网络要素在结构上如何连接；然而，这对道路网络数据的许多应用至关重要（见图 2）。因此，外部评估方法很快扩展到涵盖道路网络拓扑、路径规划及其他网络属性（Girres 和 Touya 2010；Mondzech 和 Sester 2011；Neis 等，2012；Zielstra 和 Hochmair 2012；Graser 等，2014）。当缺少参考数据，或为避免采用外部比较中一些计算量较大的方法时，人们便开发了志愿地理信息数据质量的内部评估方法。内部评估方法分析数据的空间和结构维度（Neis 等，2012；Barron 等，2014），以及贡献者数量和编辑次数（Keßler 等，2011；Neis 等，2013；Barron 等，2014；Gr¨ochenig 等，2014）。外部和内部方法常结合使用，有时还会纳入更复杂的空间数据质量评估框架（Barron 等，2014；Ballatore 和 Zipf 2015），BikeDNA 也采用了这种做法。

质量标准与适用性 对空间数据质量现有方法的另一种分类方式，是区分空间数据质量标准的制定与“适用性”理念。以前一种方法为例，ISO 19157 将空间数据质量定义为由以下方面构成：准确性（位置、主题和时间）；数据完整性；内部逻辑一致性；以及可用性（Fonte 2017）。这些概念很有用，在 BikeDNA 的质量评估中也始终有所采用，但正式数据质量标准所要求的一致性，难以应用于 OSM 这类异质的志愿者提供的数据（Hashemi andAbbaspour 2015），对于某些 OSM 应用场景也可能不太相关。因此，Barronet al.（2014）建议改为根据适用性来评估 OSM 数据，即依据数据是否满足特定用例的要求，来评估数据集的质量和适用性。BikeDNA 也采用了这种方法。因此，BikeDNA 并非参照通用标准，而是旨在帮助判断数据集是否足够好，能否用于预期用例。

网络级自行车数据评估与工具方面的不足 专门针对自行车基础设施数据质量评估的研究寥寥无几，其中 Hochmair 等 (2015) 和 Ferster 等 (2020) 是最具代表性的例子。这两项研究将专用自行车基础设施的 OSM 数据与外部参考数据集进行比较，评估数据的总体完整性以及不同类型自行车基础设施数据的完整性。这一点尤其值得关注，因为关于 OSM 总体完整性的研究表明，自行车道和路径往往是较晚才被绘制到地图上的要素（Neis 等人，2012；Barron 等，2014）。这些研究为了解 OSM 中自行车基础设施数据的完整性提供了重要见解，但并未探讨更复杂的空间数据质量指标，例如网络拓扑。这是一个非常值得关注的研究空白，因为与机动车交通数据相比，行人和骑行者基础设施数据更容易出现拓扑错误（Neis 等，2012）。此外，用于评估汽车道路网络的许多假设，例如将不连通分量和边之间的小间隙视为错误，也不适用于自行车网络，因为自行车网络

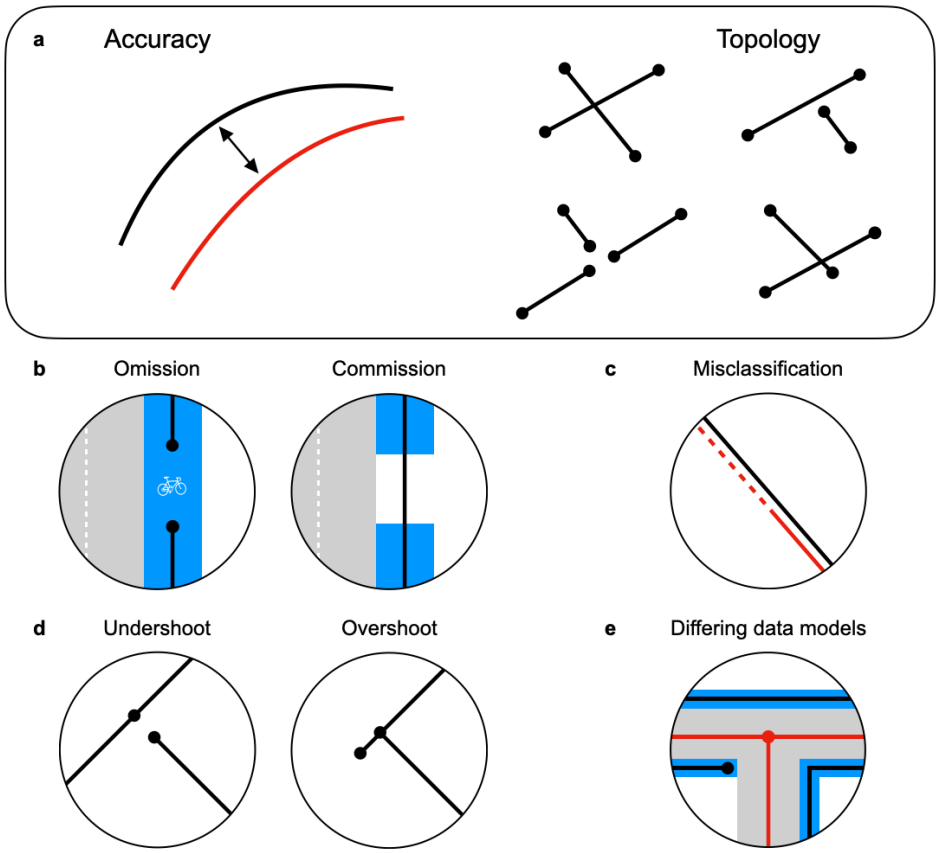


**图 2.** 自行车基础设施数据中已知的质量问题。a）数据质量评估的不同方面：准确性（左）与拓扑（右）。早期的空间数据质量研究侧重于准确性，而 BikeDNA 特别强调拓扑。改编自 Haklay 等人 (2010) 和 Neis 等人 (2012）。b）遗漏错误（左）或误报错误（右）会导致数据完整性不同。c）错误分类会使两个对应的数据集（红色与黑色）中的自行车基础设施类型不同。d）边缘不足（左）和边缘过冲（右）。e）将自行车基础设施映射到道路中心线（红色）或道路一侧（黑色）。在此例中，中心线制图以牺牲准确性为代价，生成了可路由网络数据。

通常以零散、分段的方式建成（Natera Orozco 等 2020；Szell等，2022年）。

总体而言，志愿地理信息和 OSM 数据的质量已得到充分研究，尤其是关于汽车道路网络的数据（Fonte 2017），而且研究通常表明，与其他数据源相比，OSM 数据的质量较高（Neis 等，2012；Zielstra 和 Hochmair 2012；Hochmair 等人 2015；Ferster 等 2020；Zhang 等人 2021）。然而，数据错误

和不一致现象在 OSM 中并非随机分布，而是与较低的人口密度及社会经济变量相关，并且因国家而异（Haklay2010；Ferster 等 2020）。研究还表明，行政数据的质量也参差不齐，例如受到地方数据维护者的作用及不同制图实践的影响（Hvingel 和 Jensen 2023）。因此，在某一特定地点进行的质量评估无法推广到其他地点。此外，关于骑行者基础设施的数据质量往往落后于其他道路网络数据，无论是在完整性还是一致性方面。尽管这些发现对使用 OSM 数据的自行车研究具有深远影响，但迄今只有少数研究专门评估 OSM 中自行车基础设施数据的质量。简言之，缺乏评估自行车基础设施制图质量的研究和工具，严重阻碍了开放街图数据的采用和数据驱动的自行车网络规划。

## 用 *BikeDNA* 填补空白

为弥补相关知识和质量评估可用方法方面的空白，我们开发了 BikeDNA。该工具提供了系统且易于遵循的自行车基础设施数据质量评估。这对于自行车路径规划（Murphy 和 Owen 2019）、连通性分析（Natera Orozco 等 2020；Vybornova 等 2022）、网络质量评估（Dill 2004）以及可达性研究（PeopleForBikes 2023）等应用场景都具有重要意义。BikeDNA 实现并可视化了多种现有的空间志愿地理信息和 OSM 数据质量指标，但专门针对自行车基础设施数据进行了定制。BikeDNA 的主要功能包括：

- 完整性、网络密度和 OSM 标签分析
- 网络拓扑和连通性分析
- OSM 数据与参考数据之间的要素匹配
- 生成详细的 PDF 和交互式 HTML 报告

因此，BikeDNA 同时评估数据完整性（即每个数据集包含多少信息）、数据一致性和准确性。此外，BikeDNA 还考虑了不同的数据可用性情景，可用于仅有 OSM、仅有参考数据或两者兼有的情况。所有数据质量指标（如适用）均针对整个研究区域进行全局计算，并针对覆盖研究区域的网格中的每个单元进行局部计算，从而能够检测并直观呈现数据质量的局部差异。通过详尽的文档、各步骤中的解读辅助，以及自动绘制静态地图和交互式地图，用户可以轻松获取这些质量指标。BikeDNA 可以通过指出要素和位置中存在缺口、错误或不一致之处，帮助提升数据集的质量。潜在用户包括规划人员、研究人员、数据维护者，以及所有需要了解与自行车应用相关的数据质量的人。

本文其余部分安排如下：下一节将简要讨论 BikeDNA 的概念与分析框架。随后，我们将对该工具进行描述性概述：首先介绍如何使用 BikeDNA，接着展示 BikeDNA 对 OSM 和开放政府数据进行质量评估的示例输出

'GeoDanmark'在丹麦大哥本哈根的一个示范区域的数据。最后，我们讨论其局限性以及未来开展自行车数据质量研究的必要性。此外，我们还将概述 BikeDNA 及质量评估的一般改进方向。BikeDNA 的完整说明和技术规格可在 GitHub 上查看：https://github.com/anerv/BikeDNA。

## 自行车基础设施数据评估

在以下部分中，我们将介绍概念与分析框架，概述 BikeDNA 考虑的不同数据可用性情景和数据源，描述自行车基础设施数据中已知的数据质量问题，并探讨如何在没有地面真实数据的情况下开展和解读质量评估。

### 数据源

BikeDNA 考虑了自行车网络数据可用性的三种不同情景：数据可从 OSM 获取；数据可从行政数据源（例如当地市政府）获取；或两者皆可。OSM 是自行车研究和路径规划应用中使用的主要数据源。据我们所知，行政数据主要由政府机构使用。OSM 数据与行政数据并非必然相互冲突，因为 OSM 数据有时部分基于行政数据源（Zielstra 等 2013；Rambøll 2022）。

在许多情况下，OSM 中的自行车基础设施数据在完整性等方面的质量与行政数据相当，甚至更高（Hochmair 等 2015；Ferster 等 2020）。这部分是由于政府资源不足，也因为许多辖区在数据收集和维护方面对主动出行重视不够（Rambøll 2022）。然而，志愿地理信息数据本身具有异质性，加上缺乏关于数据质量参数的元数据，阻碍了这类数据的应用，例如在公共交通规划中（Hashemi 和 Abbaspour 2015）。众包数据缺乏标准化质量保证，因此需要开展本地数据质量评估，但这类方法往往并不存在。BikeDNA 致力于弥补这一缺口。

### 已知的质量问题

在适用性方法中，某些数据错误和不一致是否构成问题，取决于具体用例。例如，只要数据拓扑表示正确（见图2a），较小的对象位移等空间不准确性通常不会影响基于网络的自行车研究。不过，有几类错误和不一致会对大多数数据应用造成问题，下面将对此作简要说明。

遗漏错误/误报错误"遗漏错误"指缺失数据，而"误报错误"指数据集中不应存在的要素（见图 2b）。在有关自行车基础设施的 OSM 和行政数据集中，这两种错误都很常见（Hochmair 等，2015；Ferster 等，2020）。

**错误分类** 当自行车基础设施被归为其他类别，或非自行车基础设施被归为自行车基础设施时，就会发生错误分类，分别导致遗漏错误和误报错误（见图 2c）。例如，未受保护的自行车基础设施可能被错误分类为受保护的自行车基础设施，这会影响交通压力等级的评估（Mekuria 等，2012）。以往关于 OSM 自行车基础设施数据质量的研究记录了多种错误分类（Hochmair 等，2015；Ferster 等，2020）。

**拓扑错误** 拓扑错误是指网络边在交叉口未能正确连接的情况，其原因可能是交叉口缺失节点，也可能是出现了“悬挂线段”（一条或多条边过短，未能在交叉口处连接）或“过伸”（一条或多条边在交叉口处过长，从而形成带有悬挂节点的小边），如图 2d 所示。在 OSM 中，与其他类型的道路网络相比，这类问题在自行车基础设施中更为突出（Neis 等，2012）。

**不一致的制图流程** 虽然这在技术上并非错误，但不同或不一致的制图方法会使数据完整性难以评估。在某些方法中，自行车道被数字化为道路中心线，无论道路只有一侧还是两侧设有自行车道。其他方法则将自行车道按其独立几何图形进行数字化；在这种情况下，同一道路两侧的自行车道会在数据集中分别表示（示意见图 2e）。这会严重扭曲网络密度和网络长度的测量结果，也使得基于不同制图方法的数据集难以比较。其他研究通过在长度计算中仅纳入一条车道（Hochmair 等，2015），或将平行线概化为中心线（Ferster 等 2020），来应对这一挑战。

## 没有地面真实数据的质量评估

缺少自行车基础设施的地面真实数据或权威数据集，是一个挑战——这不仅对自行车规划人员和研究人员如此，对任何试图评估数据质量的人来说也是如此。没有验证数据，数据质量评估就无法完全自动化。此外，鉴于许多实际自行车网络具有碎片化且质量低下的特征，通常无法通过自动化方式区分：数据质量指标表现不佳，究竟是由于数据质量较低，还是由于基础设施本身质量较低。基于这些原因，BikeDNA不会对所分析数据集的质量给出任何最终结论。相反，它可以帮助探索数据的不同特征；结合地方知识，这些特征有助于对数据质量以及数据可用性的潜在局限性作出更加充分的评估。正如Brovelli 等(2017)和Ferster 等(2020所指出的，掌握实地状况和 OSM制图规范方面的地方知识，对于质量评估的验证与解读至关重要。因此，BikeDNA强调，在任何规划或研究过程中，地方知识以及对相关区域的熟悉程度都具有重要意义。

## BikeDNA 的工作原理

这里简要说明如何使用 BikeDNA。该流程包括四个主要步骤：I. 安装、II. 设置、III. 分析，以及 IV. 报告生成（可选），参见图 1。关于规格、安装和设置的所有详细信息，请参阅 `README` 以及笔记本中的说明，这些内容均位于 GitHub代码库中。BikeDNA 完全通过 Jupyter 笔记本在 Python 中运行。因此，使用 BikeDNA 需要具备操作 Python Jupyter 笔记本的知识，但不要求具备编程能力。BikeDNA 广泛使用开源 Python 库，其中最重要的是 `GeoPandas` (Jordahl 等 2021年)、`OSMnx` (Boeing2017)、`NetworkX` (Hagberg 等 2008)，以及 `Momepy`(Fleischmann 2019)。

创建 Python 环境（步骤 I）后，填写配置文件并准备可能的参考数据集（步骤 II），即可通过运行 Jupyter 笔记本（步骤 III）完成整个分析。笔记本的结构如下：OSM 数据的准备和内在分析（*1a*、*1b*）；参考数据的准备和内在分析（*2a*、*2b*）；OSM 和参考数据的外在分析（*3a*）；最后是 OSM 与参考数据之间的要素匹配（*3b*）。OSM 和参考数据的内在分析可以独立运行，而外在分析则需要预先完成两个数据集的内在分析。

可选的最后一步 IV 是基于已执行的笔记本生成 BikeDNA 结果的详细 HTML 报告（包括交互式地图）或 PDF 报告。BikeDNA 的所有输出（处理后的数据、图对象、图表、静态和交互式地图、被识别为不一致或错误的数据）都会在笔记本执行过程中自动保存到相应的子文件夹中。

BikeDNA 需要用户提供以下几项输入：

1. **研究区域**。用于界定待分析研究区域的多边形。2. **OSM 查询**。 BikeDNA 提供了用于根据 OSM 数据界定自行车网络的默认查询，这些数据会作为 OSM 分析的一部分自动下载。如果要执行 OSM 分析，应调整这些查询，以适应当地标记实践的差异，以及对自行车网络所包含的基础设施类型的当地定义。3. **参考数据**。 如果提供参考数据集，数据必须按照数据集要求（SI 文件 2）进行预处理。此外，还必须在配置文件中提供用于解析参考数据属性的设置。

除了这些必需输入外，BikeDNA 还允许用户选择性地配置分析参数，例如用于识别过度/不足延伸的距离阈值、缺失标签分析中要纳入的 OSM 标签，以及用于要素匹配的距离阈值。

## BikeDNA 的功能

本节将以大哥本哈根的一个地区为例，使用来自 GeoDanmark 的本地数据集（GeoDanmark

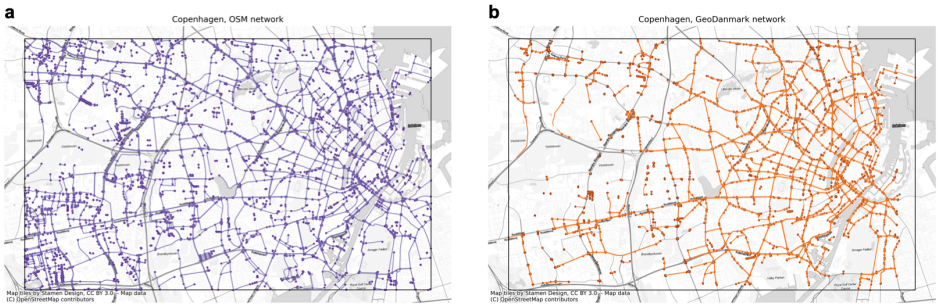


**图3.** 测试区域的输入数据：a) OSM 和 b) GeoDanmark。地图使用 BikeDNA v.1.0.0 创建。

2023）作为参考数据。方法论的详细阐述和结果解读见分析笔记本（参见补充信息文件1）。所有代码和结果也可在 GitHub 代码库的“GeoDanmark”分支中找到。

## 全局与局部分析

BikeDNA 分两个层面进行分析：全局和局部。在全局分析步骤中，会计算整个研究区域的聚合指标。有些结果（例如网络总长度）只有在全局尺度上才有意义。然而，对于某些数据不一致性，我们不仅想知道它们是否存在，还想知道它们发生在哪里。由于空间异质性，以及高、低数据质量区域的位置往往呈现非随机模式，这些不一致性的位置尤为重要；这些特征通常见于众包数据（Haklay2010）。因此，在局部分析步骤中，BikeDNA 会针对覆盖研究区域的可自定义方形网格，分别计算每个网格单元的指标。

## 网络密度

交通网络的密度定义为每平方千米的边长度或节点数量，这是能够反映数据完整性的最基本描述性统计量。由于制图方法和数据模型各不相同，基于网络密度比较两个自行车基础设施数据集的完整性向来十分困难（见图 2e）。如上文制图流程不一致一节所述，OSM 采用两种方法来绘制自行车基础设施。因此，BikeDNA 根据基础设施长度计算边密度，而非几何边长度。例如，一条长 100 米的双向路径会计为 200 米的自行车基础设施。这样便可以比较采用不同制图方法的数据集之间的数据完整性。

分析需要使用简化网络，以便在计算网络密度时不将中间网络节点（即不代表交叉口或死胡同的节点）计入其中。因此，OSM 和参考网络均使用

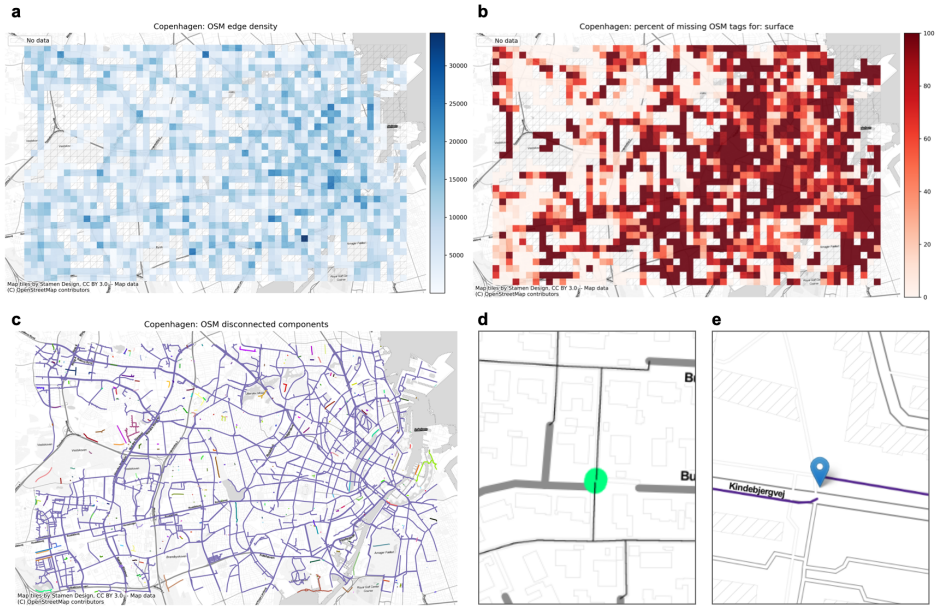


**图 4.** OSM 和 GeoDanmark 数据内在分析的结果：a）OSM 边密度。b）缺失的 OSM 标签。c）不连通分量。d）悬挂线段示例。e）来自两个不连通分量的边，它们之间的距离小于指定的距离阈值。地图使用 BikeDNA v.1.0.0 创建。

一个经修改的 OSMnx 函数（Sebastiao 2022）进行网络简化，该函数仅保留交叉口和死胡同处的节点，以及重要属性值发生变化的位置处的节点。

在内在分析中，如果网络密度模式明显偏离预期模式，网络密度值可以指示研究区域中测绘不足或过度测绘的区域。在外在分析中，对比 OSM 与参考数据集的网络密度值，可用于比较数据完整性（Haklay 2010）。

## *OSM* 标签的一致性

OSM 自行车基础设施数据的一个重要特征是标签（例如宽度、限速、路灯或骑行者关注的其他特征）分布高度异质，这给评估骑行适宜性等指标带来了障碍（Wasserman 等人，2019）。同样，OSM 标注不受限制，可能导致标签相互冲突，从而影响对自行车骑行条件的评估。因此，BikeDNA 可根据 OSM 的数据结构，检查 OSM 标签的一致性。BikeDNA 通过三种方式分析 OSM 标签：识别并可视化用户定义的 OSM 标签缺少信息的位置；突出显示哪些边被标记了两个或更多个被用户定义为互相矛盾的标签；以及

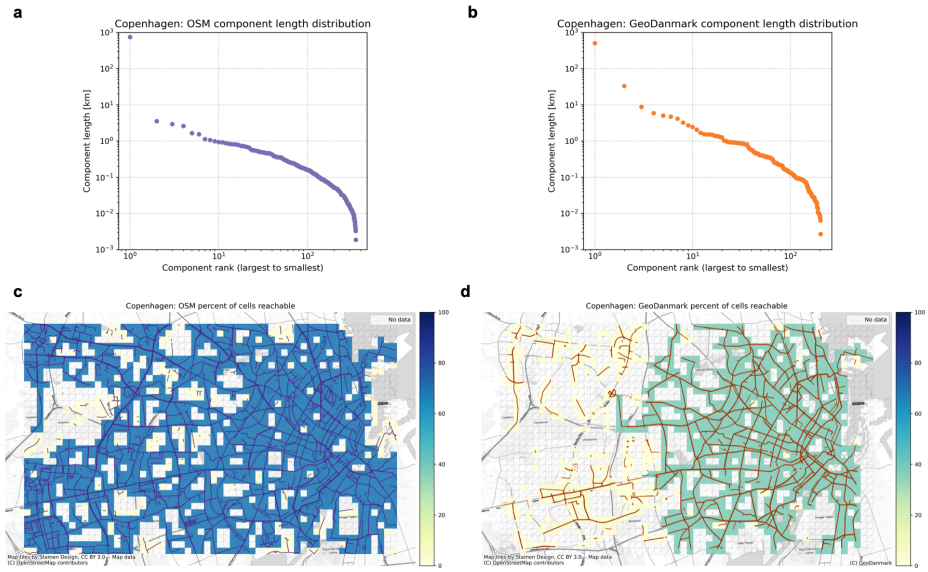


**图5.** 基于内在分析结果比较的外在分析：a) OSM 和 b) GeoDanmark 的分量长度分布齐夫图。通过 c) OSM 和 d) 参考网络可到达的单元格百分比。使用 BikeDNA v.1.0.0 创建的地图。

可视化标记模式，即用于描述 OSM 中自行车基础设施的标签在空间上的变化。此外，BikeDNA 利用 OSM 标签识别缺失的交叉口节点：当两条边相交却没有在交点处设置节点，且两者均未标记为桥梁或隧道时，这显然表明存在拓扑错误（Neis 等，2012；Barron 等，2014）。

### 网络拓扑：悬挂线段/过冲与悬挂节点

BikeDNA 提供两种方法，用于检查交叉口和死胡同处网络拓扑的一致性：分析欠伸和过伸（参见图 2），以及分析悬挂节点。悬挂线段/过冲通常由数据数字化错误造成，但也可能准确反映网络状况，例如受保护的自行车道在无法为骑行者提供保护的交叉口处终止。过冲和欠伸的存在会扭曲网络中节点与边的比例，进而影响网络指标计算。此外，自行车基础设施中的欠伸会妨碍骑行者在网络上正确进行路径规划。BikeDNA 会查找所有处于用户定义距离阈值内的过冲和欠伸，并将其显示出来以供进一步分析。

悬挂节点是度为一的节点，即只连接一条边的节点。大多数基础设施网络在实际的死胡同处或某些要素的端点处自然会包含一些悬挂节点，例如当

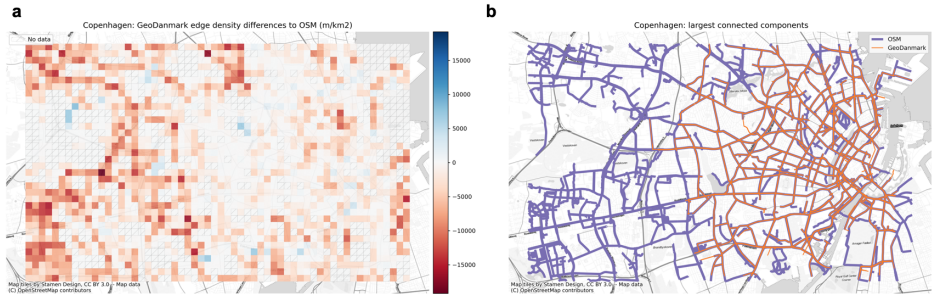


**图 6**。 OSM 与参考数据差异的外在分析：a）每个网格单元的边密度比较。b）最大连通分量。使用 BikeDNA v.1.0.0 创建的地图。

自行车道在道路中间终止时。然而，悬挂节点也可能由悬挂线段/过冲或数据遗漏导致。了解悬挂节点是由实际的死胡同还是数字化错误造成的非常重要。因此，BikeDNA 会在网格单元层级上可视化单个悬挂节点及其密度，以便调查其空间异质性。

## 网络拓扑：不连通分量

网络分量是由路径相互连接的节点构成的最大集合。换言之，一个分量内的所有节点都可以相互到达，但无法到达网络其余部分中的任何节点。大多数现实世界的自行车基础设施网络都由许多不连通分量组成（Natera Orozco等，2020年）。不过，两个彼此非常接近的不连通分量，可能表明存在真实的“缺失连接”（Vybornova 等，2022年），也可能是类似悬挂线段的数字化错误。BikeDNA 可识别并可视化输入网络中的所有不连通分量，并在齐普夫图上绘制所有网络分量长度的分布；该图按降序排列所有分量的长度。当齐普夫图在双对数尺度上呈直线时，意味着找到小型不连通分量的概率远高于传统指数分布所预期的概率（Clauset 等，2009年）。这可能意味着网络尚未进行整合，其元素只是以片段化或随机的方式添加的（Szell等，2022年）。然而，也可能出现这样的情况：最大连通分量（见图 5a 中排名 $10^0 = 1$ 处最左侧的标记）明显偏离其他分量，而图的其余部分呈现不同的形状。这可能意味着网络已经完成整合。在对同一区域进行比较时，如图 5 所示，如果一个数据集中的最大连通分量明显偏离，而另一个数据集没有这种情况，则可能表明前一个数据集更加完整。在这一特定案例中，OSM 数据集（图 5a）可能比参考数据（图 5b）更加完整。

当然，如果考虑整个道路网络，自行车基础设施的两个不连通分量也可能通过其他类型的基础设施相连。然而，




**图 7.** 要素匹配：a）匹配（蓝色）和未匹配（红色）的 OSM 要素；b）两个数据集中保护等级相同（蓝色）和不同（红色）的匹配要素。地图由 BikeDNA v.1.0.0 创建。

专用自行车基础设施各组件之间错误地缺少连接，会给自行车路径规划带来问题，并在评估自行车可达性和交通压力等级时导致误导性或不理想的结果（Murphy 和 Owen 2019；Wasserman 等人 2019）。因此，BikeDNA 还会通过查找并直观突出显示彼此间距离在用户定义的距离阈值以内、但属于不同不连通分量的边，来识别可能缺失的连接。此外，BikeDNA 还会在局部（网格单元）层面进行组件连通性分析，并将从各网格单元可达的单元格数量差异可视化。这项指标对于比较和量化两个数据集之间的网络可达范围和连通性尤其有用（见图 5c、d）。

## 要素匹配

要素匹配是指识别两个不同数据集中哪些要素对应同一个现实对象。例如，同一条道路在两个不同的数据集中可能以略有差异的几何形状表示。要比较单个要素而非要素密度，或将属性从一个数据集转移到另一个数据集（例如用于补全缺失的属性值），都需要进行要素匹配。BikeDNA 包含一种要素匹配算法，可识别参考数据集与 OSM 数据集中的对应要素（见图 7）。该方法会先将所有网络边转换为长度统一的较短线段，再寻找潜在匹配项。该匹配算法接受用户定义的输入参数（Koukoletsos 等，2012；Will，2014），首先识别给定边的最大搜索距离范围内的所有线段，然后将与给定边的无向豪斯多夫距离最小且夹角最小的线段确定为最佳匹配项。

### 比较 *OSM* 与参考数据

如果提供了参考数据集，BikeDNA 会将其与同一区域的 OSM 数据进行比较，突出显示两个数据集的差异及其所在位置，也就是说，比较两个数据集中绘制了多少自行车基础设施，以及如何绘制这些基础设施。比较分析不会预先假定哪个数据集的质量更高。因此，BikeDNA 不会自动得出结论，而是需要对差异进行解读，例如，差异要素是遗漏错误还是多绘错误造成的，以及哪个数据集更适合特定用途。数据集之间的差异会在全局和局部层面计算并呈现，并在视觉上适宜时进行叠加（见图 5 和图 6）。识别出的差异可用于评估 OSM 数据集和参考数据集的质量，并为选择使用哪个数据集提供依据。

## 讨论

在下文中，我们总结 BikeDNA 在大哥本哈根示范区域的研究发现，并介绍 BikeDNA 的一般性贡献。本节最后将讨论 BikeDNA 的局限性，并提出对未来工作的建议。

### 数据质量问题依然严重

大哥本哈根地区的用例表明，数据质量问题可能会导致对实际自行车骑行条件的呈现产生误导，影响自行车网络的范围和连通性。尽管丹麦以相对浓厚的自行车文化而闻名（Agervig 和 Ebert，2012），我们的分析发现自行车基础设施数据质量存在若干严重问题，尤其是行政参考数据集存在大片区域明显未能充分绘制的情况。此外，OSM 和参考数据均存在缺失连接和欠伸，导致数据集比实际自行车网络更加零散。重要的是，分析显示缺失或误导性数据存在很大的空间差异，这凸显了开展局部数据评估的重要性。

### *BikeDNA*：面向自行车研究人员、规划人员和骑行者的工具 *s*

BikeDNA通过支持对数据质量的空间异质性进行直观探索，并采用超越传统自行车基础设施数据质量指标的基于网络的测量指标，为自行车基础设施数据质量带来了全新的见解。这些创新功能带来了丰富的应用场景，例如：

- 自行车基础设施的城市与区域规划，尤其需要将整个网络作为一个整体加以考量（Szell 等人，2022）。
- 维护和改进 OSM 与行政数据。
- 关于主动出行以及多模式网络的交通研究要求所有交通层（Alessandretti等，2022年）具备高质量的数据。

- 改进用于交通规划和研究工具的自行车网络数据，例如用于交通规划的“Propensity to Cycle”（Lovelace等，2017年）、用于交通仿真的“A/B Street”（Carlino等，2023年），或PeopleForBikes的自行车网络分析（PeopleForBikes，2023年）。

- 公民科学和行动主义通过让公众参与贡献可靠且经过质量评估的数据集，发挥作用；这与其他支持公众收集数据、推动更可持续交通的项目类似，例如 OpenBikeSensor (2023) 和StreetComplete (2023)。

## 局限性

尽管我们尝试涵盖了自行车网络数据质量的主要相关方面，但 BikeDNA 的设计仍存在一些局限性。

在数据建模方面，出于简化考虑，BikeDNA 使用了无向简化网络。这意味着不考虑允许的行驶方向和转向限制，假设所有边均可双向通行，并且同一道路上的不同通行方向不会表示为独立的边。因此，该工具目前尚未使用网络路径规划。不过，在未来的迭代中，纳入行驶方向和底层道路网络，或许有助于进行准确的路径计算。

另一项局限涉及该工具的核心用途及其可生成的结果类型：由于我们不会将某个数据集视为地面真实数据，并据此评估另一个数据集，因此在发现差异时，无法判断错误出在哪里。要成功应用 BikeDNA，用户需要对 OSM 数据结构和标记约定有所了解，同时也要充分熟悉研究区域，以便正确解读结果。

最后，我们并未直接评估 OSM 或参考数据的位置准确性——尽管可以通过要素匹配推断出一定程度的相对位置准确性。

## 未来工作

未来仍有大量潜在工作有待开展——不仅是针对自行车基础设施数据总体质量评估的工作，也包括针对 BikeDNA 的工作。

首先，尚待确定所分析的指标中，是否有任何一项能通过与其他质量指标的相关性，作为数据质量的通用预测指标。这将使评估更快速、更简单，并免去与参考数据进行比较以及使用更复杂分析工具的必要。其次，对 OSM 和其他数据源中的自行车基础设施数据质量进行评估时，应将其与更广泛的自行车基础设施制图标准化工作联系起来。比较不同数据集的一项挑战，是需要在不同的自行车基础设施类型体系之间进行转换。在 BikeDNA 中，我们通过仅区分受保护和未受保护的自行车基础设施，解决了这一歧义；不过，借助更完善的分类方案，这种做法还可以进一步改进。为了

支持更好的规划和研究，还需要进一步探讨如何在数据中最佳地呈现自行车出行条件。

对于 BikeDNA 而言，实现便捷比较两个非 OSM 数据集并进行要素匹配的功能仍有待完成。此外，对于 OSM 数据的内在分析，未来纳入有关贡献者和编辑的历史数据可以提供一种补充视角，因为这些元数据可能是数据质量的潜在指标（Neis 等，2012；Gr¨ochenig 等，2014）。我们期待听到社区分享未来的案例研究，也欢迎您对当前工具可能的修改和改进提出建议。

## 结论

BikeDNA 是一款用于对自行车基础设施数据进行可复现质量评估的开源工具。BikeDNA 支持基于“适用性”理念的定制化评估，并采用可配置的设计，可根据不同的数据可用性场景进行调整，同时支持独立分析和比较分析。该工具计算了广泛的质量指标，并在以往自行车基础设施质量研究的基础上，进一步关注数据拓扑和结构属性。通过比较 OSM 与丹麦行政数据这一 BikeDNA 应用示例可以看出，BikeDNA 能够揭示重要的空间差异，例如缺失数据和拓扑错误方面的差异；示例分析揭示了许多不一致之处和错误，也说明了对自行车基础设施数据进行质量评估的必要性。关于高质量自行车基础设施数据的定义和收集，无论是对 BikeDNA 还是更广泛的相关领域而言，仍有一些开放性问题。缺乏高质量的官方数据以及公认的自行车基础设施类型体系，反映出主动出行方式在历史上一直受到忽视。对于世界各地致力于利用数据驱动的规划最新进展来扩展自行车网络的城市、地区和国家而言，迫切需要有关现有自行车骑行条件的更优质数据。

### 致谢

感谢所有 OSM 贡献者，他们的努力让空间数据得以开放共享；感谢开发并贡献 Python 库的所有人，是他们让这项工作得以实现；也感谢 Cl´ement Sebastiao 开发了用于网络简化的修改版 OSMnx 函数。我们感谢丹麦交通部提供的支持。

### 参考文献

Agervig CT 和 Ebert AK（2012）《北欧自行车文化：从“黄金时代”到“复兴”》。载于：Parkin J（编）《自行车骑行与可持续性》，《交通与可持续性》，第1卷。Emerald Group Publishing Limited。国际标准书号 978-1-78052-299-9 978-1-78052-298-2，第23–58页。数字对象标识符：10.1108/S2044-9941(2012)0000001004。网址 https://doi.org/10.1108/S2044-9941(2012)0000001004。

Alessandretti L、Natera Orozco LG、Saberi M、Szell M 和 Battiston F（2022）多模式城市交通与多层交通网络。《环境与规划B：城市分析与城市科学》：239980832211081数字对象标识符：10.1177/23998083221108190。网址 http://journals.sagepub.com/doi/10.1177/23998083221108190。Ballatore A 和 Zipf A（2015）志愿地理信息的概念性质量框架。载于：Fabrikant SI、Raubal M、Bertolotto M、Davies C、Freundschuh S 和 Bell S（编）《空间信息理论》，第9368卷。Cham：施普林格国际出版。国际标准书号 978-3-319-23373-4 978-3-319-23374-1，第89–107页。数字对象标识符：10.1007/978-3-319-23374- 1 5。网址 http://link.springer.com/10.1007/978-3-319-23374-1_5。丛书名称：《计算机科学讲义》。Barron C、Neis P 和 Zipf A（2014）OpenStreetMap内在质量分析的综合框架：OpenStreetMap内在质量分析的综合框架。《地理信息系统汇刊》18(6)：877–895。数字对象标识符：10.1111/tgis.12073。网址 https: //onlinelibrary.wiley.com/doi/10.1111/tgis.12073。Boeing G（2017）OSMnx：获取、构建、分析和可视化复杂街道网络的新方法。《计算机、环境与城市系统》65：126–139。数字对象标识符：10.1016/j.compenvurbsys.2017.05.004。网址 https://linkinghub.elsevier.com/retrieve/pii/S0198971516303970。Brovelli MA、Minghini M、Molinari M 和 Mooney P（2017）迈向OpenStreetMap与权威道路数据集的自动化比较。《地理信息系统汇刊》21(2)：191–206。数字对象标识符：10.1111/tgis.12182。网址 https://onlinelibrary.wiley.com/doi/ 10.1111/tgis.12182。C40城市（2019）波哥大自行车网络升级，大幅增加骑行出行次数。网址 https://www.c40.org/case-studies/upgrade-of-the-cycle-network-in-bogota-dramatically-increases-bike-trips/。Carlino D、Li Y 和 Kirk M（2023）A/B Street。网址 https://github.com/a-b-street/ abstreet。原始日期：2018-06-04T00:44:43Z。巴黎市（2021）全新自行车计划：打造100%适宜骑行的城市。网址 https://www.paris.fr/pages/un-nouveau-plan-velo-pour-une-ville-100-cyclable-19554。Clauset A、Shalizi CR 和 Newman MEJ（2009）实证数据中的幂律分布。《SIAM评论》51(4)：661–703。数字对象标识符：10.1137/070710111。网址 https://doi.org/10.1137/ 070710111。Degrossi LC、Porto de Albuquerque J、Santos Rocha Rd 和 Zipf A（2018）志愿和众包地理信息质量评估方法分类。《地理信息系统汇刊》22(2)：542–560。数字对象标识符：10.1111/tgis.12329。网址 https://onlinelibrary.wiley.com/doi/abs/10.1111/tgis.12329。Dill J（2004）测量骑自行车和步行的网络连通性。载于：第83届交通研究委员会年会，第83卷。华盛顿特区，第19页。欧洲委员会（2021）欧盟城市交通新框架。技术报告，欧盟委员会，布鲁塞尔。网址 https://transport.ec.europa.eu/system/files/2021-12/ com_2021_811_the-new-eu-urban-mobility.pdf。

Ferster C、Fischer J、Manaugh K、Nelson T 和 Winters M（2020）利用 OpenStreetMap 清查自行车基础设施：与城市开放数据的比较研究。《国际可持续交通期刊》14(1)：64–73。数字对象标识符：10.1080/15568318.2018.1519746。网址：https://www.tandfonline.com/doi/full/10.1080/15568318.2018.1519746。Fleischmann M（2019）momepy：城市形态测量工具包。《开源软件期刊》4(43)：1807。数字对象标识符：10.21105/joss.01807。网址：https://joss.theoj.org/papers/10.21105/joss.01807。Fonte C（2017）评估志愿地理信息数据质量。《制图与公民传感器》：137–163。数字对象标识符：10.5334/bbf.g。网址：https://www.ubiquitypress.com/site/chapters/10.5334/bbf.g/。GeoDanmark（2023）丹麦地理 - GeoDanmark。网址：https://dataforsyningen.dk/data/3563。Girres JF 和 Touya G（2010）法国开放街图数据集的质量评估：法国开放街图数据集的质量评估。《地理信息系统汇刊》14(4)：435–459。数字对象标识符：10.1111/j.1467-9671.2010.01203.x。网址：https://onlinelibrary.wiley.com/doi/10.1111/j.1467-9671.2010.01203.x。Graser A、Straub M 和 Dragaschnig M（2014）构建用于街道网络比较的开源分析工具箱：比较 OSM 与奥地利官方参考图的指标、工具和结果。《地理信息系统汇刊》18(4)：510–526。数字对象标识符：10.1111/tgis.12061。网址：https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=97351436&site=ehost-live。出版商：Wiley-Blackwell。Gr¨ochenig S、Brunauer R 和 Rehrl K（2014）通过分析不同时期的社区活动估算志愿地理信息数据集的完整性。载于：Huerta J、Schade S 和 Granell C（编），《通过位置与地点连接数字欧洲》，地理信息与制图讲义。Cham：施普林格国际出版。国际标准书号 978-3-319-03611-3，第3–18页。数字对象标识符：10.1007/978-3-319-03611-3 1。网址：https://doi.org/10.1007/978-3-319-03611-3_1。Hagberg AA、Schult DA 和 Swart PJ（2008）使用 NetworkX 探索网络结构、动态和功能。¨ 载于：Varoquaux G、Vaught T 和 Millman J（编），《第七届 Python 科学会议论文集》。美国加利福尼亚州帕萨迪纳，第11–15页。Haklay M（2010）志愿地理信息有多可靠？OpenStreetMap 与英国军械测量局数据集的比较研究。《环境与规划B：规划与设计》37(4)：682–703。数字对象标识符：10.1068/b35097。网址：http://journals.sagepub.com/doi/10.1068/b35097。Haklay MM、Basiouka S、Antoniou V 和 Ather A（2010）要绘制好一个区域，需要多少名志愿者？林纳斯定律对志愿地理信息的适用性。《制图学期刊》47(4)：315–322。数字对象标识符：10.1179/000870410X12911304958827。网址：https://doi.org/10.1179/000870410X12911304958827。Hashemi P 和 Abbaspour RA（2015）基于空间相似性概念评估 OpenStreetMap 的逻辑一致性。载于：Jokar Arsanjani J、Zipf A、Mooney P 和 Helbich M（编），《地理信息科学中的 OpenStreetMap：经验、研究与应用》，地理信息与制图讲义。Cham：施普林格国际出版。

国际标准书号 978-3-319-14280-7，第19–36页。数字对象标识符：10.1007/978-3-319-14280-7 2。网址：https://doi.org/10.1007/978-3-319-14280-7_2。Hochmair HH、Zielstra D 和 Neis P（2015）评估美国 OpenStreetMap 中自行车道和车道要素的完整性：OpenStreetMap 中自行车要素的完整性。《地理信息系统汇刊》19(1)：63–81。数字对象标识符：10.1111/tgis.12081。网址：https://onlinelibrary.wiley.com/doi/10.1111/tgis.12081。Hvingel L 和 Jensen T（2023）人人都能获取优质自行车数据。《Trafik og veje》。网址：https://www.kl.dk/media/53604/artikel_gode_cykeldata_trafik_og_veje_jan_2023_layout.pdf。Jaramillo P、Kahn Ribeiro S、Newman P、Dhar S、Diemuodeke OE、Kajino T、Lee DS、Nugroho SB、Ou X、Hammer Strømman A 和 Whitehead J（2022）交通运输。载于：《气候变化2022：气候变化减缓》。政府间气候变化专门委员会（IPCC），第1049–1160页。网址：doi:10.1017/9781009157926.012。Jordahl K、Bossche JVd、Fleischmann M、McBride J、Wasserman J、Badaracco AG、Gerard J、Snow AD、Tratner J、Perry M、Farmer C、Hjelle GA、Cochran M、Gillies S、Culbertson L、Bartos M、Ward B、Caria G、Taves M、Eubank N、sangarshanan、Flavin J、Richards M、Rey S、maxalbert、Bilogur A、Ren C、Arribas-Bel D、Mesejo-Le´on D 和 Wasser L（2021）geopandas/geopandas：v0.10.2。数字对象标识符：10.5281/zenodo.5573592。网址：https://zenodo.org/record/5573592。Keßler C、Trame J 和 Kauppinen T（2011）追踪志愿地理信息中的编辑过程：以 OpenStreetMap 为例。空间信息理论会议论文集，专题研讨会：识别时空分布数据中的对象、过程和事件：7。Koukoletsos T、Haklay M 和 Ellul C（2012）通过针对线性数据的自动匹配程序评估志愿地理信息的完整性。《地理信息系统汇刊》16(4)：477–498。数字对象标识符：10.1111/j.1467-9671.2012.01304.x。网址：https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9671.2012.01304.x。Lovelace R、Goodman A、Aldred R、Berkoff N、Abbas A 和 Woodcock J（2017）骑行倾向工具：用于可持续交通规划的开源在线系统。《交通与土地利用期刊》10(1)。数字对象标识符：10.5198/jtlu.2016.862。网址：https://www.jtlu.org/index.php/jtlu/article/view/862。期号：1。Medeiros G 和 Holanda M（2019）地理信息系统和志愿地理信息中的数据质量解决方案：系统性文献综述。载于：Rocha A、Adeli H、Reis LP 和 Costanzo S（编），《信息系统与技术新知》，智能系统与计算进展。Cham：施普林格国际出版。国际标准书号 978-3-030-16181-1，第645–654页。数字对象标识符：10.1007/978-3-030-16181-1 61。Mekuria MC、Furth PG 和 Nixon H（2012）低压力骑行与网络连通性。技术报告11-19，米内塔交通运输研究所。网址：https://www.semanticscholar.org/paper/Low-Stress-Bicycling-and-Network-Connectivity-Mekuria-Furth/a50063c06112d3eb6aa752dfd362e1bdbc7f1c7e。Mondzech J 和 Sester M（2011）基于应用需求的 OpenStreetMap 数据质量分析。《Cartographica：地理信息国际期刊

与地理可视化 46(2): 115–125。数字对象标识符:10.3138/carto.46.2.115。网址 https: //www.utpjournals.press/doi/abs/10.3138/carto.46.2.115。出版商：多伦多大学出版社。Murphy B 和 Owen A（2019）在国家可达性评估中实施低压力自行车路线规划。《交通研究记录》2673(5): 240–249。数字对象标识符:10.1177/0361198119837179。网址 https://doi.org/10.1177/0361198119837179。出版商：SAGE 出版公司。Natera Orozco LG、Battiston F、I˜niguez G 和 Szell M（2020）数据驱动的自行车网络优化增长策略。《皇家学会开放科学》7(12): 201130。数字对象标识符:10.1098/rsos. 201130。网址 https://royalsocietypublishing.org/doi/full/10.1098/ rsos.201130。出版商：英国皇家学会。Neis P、Zielstra D 和 Zipf A（2012）众包地图中的街道网络演变：2007–2011 年德国的 OpenStreetMap。《未来互联网》4(1): 1–21。数字对象标识符:10.3390/fi4010001。网址 https://www.mdpi.com/1999-5903/4/1/1。编号：1 出版商：分子多样性保护国际。Neis P、Zielstra D 和 Zipf A（2013）部分世界区域志愿地理信息数据贡献与社区发展的比较。《未来互联网》5(2): 282–300。数字对象标识符:10.3390/fi5020282。网址 https://www.mdpi.com/1999-5903/ 5/2/282。编号：2 出版商：多学科数字出版研究所。Nelson T、Ferster C、Laberee K、Fuller D 和 Winters M（2021）用于自行车研究与实践的众包数据。《交通评论》41(1): 97–114。数字对象标识符:10.1080/01441647.2020. 1806943。网址 https://doi.org/10.1080/01441647.2020.1806943。Olmos LE、Tadeo MS、Vlachogiannis D、Alhasoun F、Espinet Alegre X、Ochoa C、Targa F 和 Gonz´alez MC（2020）用于规划自行车基础设施增长的数据科学框架。《交通研究 C 辑：新兴技术》115: 102640。数字对象标识符：10.1016/j.trc.2020.102640。网址 https://www.sciencedirect.com/science/ article/pii/S0968090X19306436。OpenBikeSensor（2023）OpenBikeSensor。网址 https://openbikesensor.org/en/。PeopleForBikes（2023）BNA 自行车网络分析。网址 https://bna.peopleforbikes.org/#/。Rambøll（2022）步行与骑行数据：实践、挑战、需求与缺口。网址 https://ramboll.com/-/media/files/rgr/documents/markets/transport/walking-cycling-data-gaps-2022.pdf。Sebastiao C（2022）OSMnx 简化的修改版本。GitHub 网址 https:// github.com/anerv/NERDS_osmnx。Senaratne H、Mobasheri A、Ali AL、Capineri C 和 Haklay MM（2017）志愿地理信息质量评估方法综述。《国际地理信息科学杂志》31(1): 139–167。数字对象标识符:10.1080/13658816.2016.1189556。网址 https: //www.tandfonline.com/doi/full/10.1080/13658816.2016.1189556。Steinacker C、Storch DM、Timme M 和 Schr¨oder M（2022）以需求为导向设计自行车基础设施网络，提升城市骑行适宜性。《自然·计算科学》：1– 10数字对象标识符:10.1038/s43588-022-00318-w。网址 https://www.nature.com/articles/s43588-022-00318-w。出版商：自然出版集团。

StreetComplete（2023）StreetComplete。网址https://streetcomplete.app。Szell M、Mimar S、Perlman T、Ghoshal G 和 Sinatra R（2022）城市自行车网络的增长。《科学报告》12(1): 6765。数字对象标识符:10.1038/s41598-022-10783-y。网址
https://www.nature.com/articles/s41598-022-10783-y。编号：1 出版商：自然出版集团。Vybornova A、Cunha T、G¨uhnemann A 和 Szell M（2022）自行车网络缺失连接的自动检测。《地理分析》n/a(n/a)。数字对象标识符:10.1111/gean.12324。网址
https://onlinelibrary.wiley.com/doi/abs/10.1111/gean.12324。Wasserman D、Rixey A、Zhou XE、Levitt D 和 Benjamin M（2019）评估 OpenStreetMap 在交通压力等级分析中的表现潜力。《交通研究记录》2673(4): 284–294。数字对象标识符:10.1177/0361198119836772。网址
https://doi.org/10.1177/0361198119836772。出版商：SAGE 出版公司。Will J（2014）开发一种自动匹配算法，以评估 *OpenStreetMap* 道路网络的质量：瑞典哥¨德堡案例研究。博士论文，隆德大学。网址
h
t
t
p
s://
www.semanticscholar.org/paper/Development-of-an-automated-matching-algorithm-to-%3A-Will/b3b77d579077b967820630db56522bef31654f21。Zhang X、Wang T、Jiao D、Zhou Z、Yu J 和 Cheng X（2021）基于平行车道识别和对称规则，检测众包街道网络中的不一致信息。《国际摄影测量与遥感学会摄影测量与遥感杂志》175: 386–402。数字对象标识符:10. 1016/j.isprsjprs.2021.03.014。网址
https://www.sciencedirect.com/science/article/pii/S092427162100085X。Zielstra D、Hochmair H 和 Neis P（2013）评估数据导入对 OpenStreetMap 完整性的影响——美国案例研究。《地理信息系统汇刊》17。数字对象标识符:10.1111/tgis.12037。Zielstra D 和 Hochmair HH（2012）使用免费数据和专有数据比较街道网络中的最短路径长度，以实现有效的行人路径规划。《交通研究记录》2299(1): 41–47。数字对象标识符:10.3141/2299-05。网址https://doi.org/10.3141/2299-05。出版商：SAGE 出版公司。