

UEH Digital Repository

Book Chapter

2021

Big data cho mục tiêu phát triển bền vững 2030: kinh nghiệm quốc tế

TS. Phạm Thái Bình

UEH University

Citation:

TS. Phạm Thái B. (2021), "Big data cho mục tiêu phát triển bền vững 2030: kinh nghiệm quốc tế", Thông tin và Truyền thông

Available at <https://digital.lib.ueh.edu.vn/handle/UEH/62515>

This item is protected by copyright and made available here for research and educational purposes. The author(s) retains copyright ownership of this item. Permission to reuse, publish, or reproduce the object beyond the bounds of Vietnam Intellectual Property Law (2005, 2009 and 2022) or other exemptions to the law must be obtained from the author(s).

BIG DATA CHO MỤC TIÊU PHÁT TRIỂN BỀN VỮNG 2030: KINH NGHIỆM QUỐC TẾ

TS. Phạm Thái Bình *

Viện Tài chính Bền vững UEH, Trường Đại học Kinh tế TpHCM

TÓM TẮT:

Chương trình Nghị sự 2030, được 193 quốc gia và lãnh thổ thành viên Liên Hiệp Quốc cam kết thực hiện năm 2015, đã xây dựng 17 mục tiêu phát triển bền vững (SDGs) được cụ thể hóa thành 169 mục tiêu thực hiện và 231 chỉ tiêu thống kê. Dữ liệu lớn (Big data) được xem như là yếu tố quan trọng hỗ trợ thực hiện, theo dõi và giám sát các tiến trình thực hiện các SDGs. Với mục đích đánh giá vai trò của Big data trong thực hiện SDGs thông qua kinh nghiệm tích lũy từ các dự án được triển khai ở các quốc gia đang phát triển để đưa ra những khuyến nghị chính sách ứng dụng Big data cho Việt Nam, báo cáo khoa học này phân tích sự tiến hóa mô hình ứng dụng Big data từ 3V sang 3C, mô hình sau còn được gọi là Big data cho phát triển. Kế tiếp, những tình huống và kinh nghiệm ứng dụng Big data ở các quốc gia như China, Colombia, Indonesia, Nepal, Philippines, Sri Lanka, Singapore và châu Âu được trình bày. Trên cơ sở đó, chúng tôi gợi ý những chính sách ứng dụng Big data cho thực hiện mục tiêu phát triển bền vững 2030 ở Việt Nam.

Từ khóa: Big data, Nghị sự 2030, SDGs, phát triển bền vững

“No one will be left behind – Không ai bị bỏ lại ở phía sau”

1. MỞ ĐẦU

Tháng 9 năm 2015 đánh dấu cột mốc quan trọng trong hợp tác quốc tế giữa 193 quốc gia và lãnh thổ thành viên Liên Hiệp Quốc (LHQ) khi tất cả cùng thông qua Chương trình Nghị sự 2030 cho các mục tiêu phát triển bền vững (The 2030 Agenda for Sustainable Development Goals, hay viết tắt SDGs). Các mục tiêu phát triển bền vững được xem là những mục tiêu phổ quát được thiết kế như là sự tiếp nối của Mục tiêu Phát Triển Thiên Niên kỷ

* Liên hệ: binhpt@ueh.edu.vn

(MDGs) nhằm tiến tới chấm dứt đói nghèo dưới mọi hình thức, bảo vệ trái đất và đảm bảo hoà bình thịnh vượng cho mọi người dân ở mọi quốc gia vào năm 2030 (UN Desa, 2016).

SUSTAINABLE DEVELOPMENT GOALS



Hình 1: 17 Mục tiêu phát triển bền vững

Hình 1 giới thiệu 17 mục tiêu phát triển bền vững bao gồm:⁷⁸ xoá nghèo (SDG1), không còn nạn đói (SDG2), sức khoẻ và chất lượng sống tốt (SDG3), giáo dục có chất lượng (SDG4), bình đẳng giới (SDG5), nước sạch và vệ sinh (SDG6), năng lượng sạch với giá thành hợp lý (SDG7), công việc tốt và tăng trưởng kinh tế (SDG8), công nghiệp, sáng tạo, và phát triển hạ tầng (SDG9), thu hẹp bất bình đẳng (SDG10), đô thị và cộng đồng bền vững (SDG11), tiêu thụ và sản xuất có trách nhiệm (SDG12), hành động về khí hậu (SDG13), tài nguyên và môi trường biển (SDG14), tài nguyên và môi trường trên đất liền (SDG15), hoà bình, công lý và các thể chế mạnh mẽ (SDG16), quan hệ đối tác vì các mục tiêu (SDG17).⁷⁹ Để cụ thể hoá việc thực hiện, từ 17 mục tiêu tổng quát (goals), LHQ đã xây dựng thành 169 mục tiêu cụ thể (targets) với 231 chỉ tiêu thống kê (indicators). Tuy nhiên, các quốc gia có thể bản địa hoá các chỉ tiêu thống kê phát triển bền vững (SDIs) tương thích với SDGs của Chương trình nghị sự 2030. Một vài ví dụ điển hình như Liên minh châu Âu (EU) có bộ SDIs gồm 134 chỉ tiêu

⁷⁸ Xem chi tiết ở Phụ lục.

⁷⁹ Xem chú giải chi tiết các mục tiêu ở Phụ lục.

cho 10 lĩnh vực tương thích với 17 SDGs,⁸⁰ tương tự như vậy Anh quốc có 66 chỉ tiêu và 3 lĩnh vực (xem chi tiết Jones và cộng sự, 2016).

Đối với Việt Nam, Thủ tướng Chính phủ đã ban hành quyết định 622/QĐ-TTg ngày 10/05/2017 bản địa hoá 115 mục tiêu cụ thể, tương thích hoàn toàn với 17 SDGs, phù hợp với năng lực và điều kiện nước ta.⁸¹ Và Bộ Kế hoạch Đầu tư đã ban hành thông tư 03/2019/TT-BKHDT ngày 22/01/2019 quy định 158 chỉ tiêu thống kê phát triển bền vững cùng với lộ trình thực hiện từ 2019 đến 2025. Có thể thấy, dù Chính phủ đã hành động rất quyết liệt từ năm 2017 nhưng việc triển khai đo lường và giám sát tiến trình thực hiện các SDGs ở Việt Nam còn chậm trễ và tiềm ẩn rủi ro không đạt được các mục tiêu cụ thể cốt lõi (Baum, 2017).

Tuy nhiên những khó khăn trong việc thu thập dữ liệu SDGs Việt Nam đang gặp phải cũng là vấn đề chung của thế giới. Nghiên cứu của LHQ do Campbell (2019) thực hiện cho thấy rằng chỉ có 22% chỉ tiêu SDGs có liên quan đến môi trường được cập nhật liên tục và 68% chỉ tiêu thống kê khác không có dữ liệu để đánh giá tiến trình thực hiện. Campbell nhận định rằng các chính phủ đã không đầu tư đúng mức cho mục tiêu phát triển bền vững vì thiếu dữ liệu chúng ta không biết làm gì để đạt mục tiêu.

Một trở ngại khác của việc xây dựng cơ sở dữ liệu SDIs cho các SDGs là chi phí đầu tư rất lớn cho việc thu thập dữ liệu về môi trường và xây dựng hạ tầng thống kê có đủ năng lực đo lường và hỗ trợ thực hiện các SDGs (Castell và cộng sự, 2017). Schmidt-Traub và cộng sự (2015) đã ước tính cần đầu tư 1 tỉ USD mỗi năm để nâng cấp năng lực thống kê ở 77 quốc gia có thu nhập thấp, và 45 – 50 tỉ USD cho dữ liệu SDGs cần giám sát ở cấp độ I (Tier I).

Nhận thức được tầm quan trọng của dữ liệu SDGs, từ năm 2013-2015, một nhóm công tác chuyên trách của LHQ⁸² (Maarroof, 2015) đã khảo sát tính khả thi, tiềm năng và thách thức khi ứng dụng dữ liệu lớn (sau đây gọi là Big data) cho việc đo lường và quan trắc các chỉ tiêu thống kê và hỗ trợ thực hiện mục tiêu phát triển bền vững của các quốc gia và toàn cầu. Rất đáng ngạc nhiên, vào thời điểm khảo sát, chỉ có 2% quốc gia phản hồi có ứng dụng Big data cho các chỉ tiêu thống kê SDGs, gần 30% chỉ sử dụng

⁸⁰ Cập nhật EU SDIs tại <https://ec.europa.eu/eurostat/web/sdi/indicators>

⁸¹ Xem chi tiết ở Phụ lục I Quyết định 622/QĐ-TTg.

⁸² Nhóm công tác có sự tham gia của 5 trường đại học và viện nghiên cứu, 9 tổ chức quốc tế và 3 quốc gia.

Big data cho thông kê giá cả hàng hoá. Dù vậy, có đến 60% phản hồi nhận thấy rất cần một hướng dẫn ứng dụng Big data cho các chỉ số SDGs.

Các nghiên cứu tiếp theo của Van Halderen và cộng sự (2021) đã ghi nhận có 65 chỉ tiêu thống kê gắn với 16 SDGs (ngoại trừ mục tiêu SDG16: hoà bình, công lý và các thể chế mạnh mẽ) có thể có lợi ích trực tiếp hoặc gián tiếp từ các nguồn Big data, trong đó các mục tiêu SDG1 (xóa nghèo), SDG3 (sức khoẻ và chất lượng sống tốt), SDG6 (nước sạch và vệ sinh) và SDG11 (đô thị và cộng đồng bền vững) nhận được nhiều lợi ích nhất. Dữ liệu không gian địa lý và dữ liệu được tạo ra bởi công dân được nhóm chuyên gia đánh giá là nguồn dữ liệu lớn tốt nhất. Trong khi đó, các nguồn dữ liệu lớn khác như dữ liệu sử dụng di động hay dữ liệu trực tuyến về giá và lao động chỉ đóng vai trò gián tiếp cho một số ít chỉ tiêu SDGs.

Báo cáo khoa học này tìm hiểu vai trò, tầm quan trọng và những kết quả ứng dụng thực tiễn của Big data cho các mục tiêu phát triển bền vững 2030 ở một số nền kinh tế đang phát triển nói chung và khu vực ASEAN nói riêng để tìm kiếm giải pháp phù hợp với điều kiện phát triển kinh tế - xã hội và công nghệ của Việt Nam cho 2 vấn đề sau: (i) Việt Nam cần làm gì để có nguồn Big data đáng tin cậy đáp ứng cho các SDGs?; và (ii) Làm thế nào để khai thác hiệu quả các nguồn Big data để phục vụ cho việc thực hiện các SDGs và hoạch định chính sách ở nước ta?.

Mục tiếp theo trình bày về vai trò và tầm quan trọng của Big data đối với các SDGs và sau đó, phần 3, thảo luận về các kết quả triển khai ứng dụng Big data ở một số quốc gia đang phát triển. Các khuyến nghị chính sách cho Việt Nam được nêu chi tiết ở phần 4, và phần 5 kết luận các vấn đề được nghiên cứu.

2. BIG DATA VÀ VAI TRÒ CỦA BIG DATA ĐỐI VỚI CÁC MỤC TIÊU PHÁT TRIỂN BỀN VỮNG (SDGs)

2.1 Big data (dữ liệu lớn) là gì?

Big data, dữ liệu lớn, là một từ thông dụng đề cập đến một khái niệm bao trùm và sự thực không có một định nghĩa thống nhất về Big data dù chúng ta có thể hiểu theo nghĩa hẹp Big data là lượng rất lớn dữ liệu cấu trúc và phi cấu trúc, truyền thống và phi truyền thống được tạo ra hàng ngày bởi nhân loại như là dữ liệu vệ tinh, các cảm biến quan trắc khí hậu, nhật ký các cuộc gọi điện thoại, thông tin tương tác trên web,... Big data theo nghĩa rộng không chỉ là dữ liệu mà còn bao hàm các phương pháp và mô hình phân tích – dự báo để “*thấu hiểu*” dữ liệu hỗ trợ quá trình ra quyết định hay

thực hiện một vấn đề cụ thể. Nghiên cứu của Faravetto và cộng sự (2020) cho thấy các nhà nghiên cứu trong lĩnh vực khoa học xã hội cơ bản thống nhất Big data cần được định nghĩa thực tế hơn, liên kết nó với các quá trình như thu thập dữ liệu và xử lý dữ liệu.

Vậy, Big data cho phát triển là gì? Đây là một khái niệm mới được khởi xướng bởi kinh tế gia Letouzé (2015) đề cập tới các nguồn dữ liệu tin cậy phù hợp cho việc xây dựng chính sách và hoạch định các chương trình phát triển bền vững (Pulse, 2013). Cụ thể hơn, chúng ta có thể nói đến các ứng dụng Big data cho mục đích thống kê kinh tế, dân số và xã hội, hoạch định chính sách, hay dữ liệu thông tin địa không gian phục vụ quan trắc khí hậu, quản lý và chống phá rừng, cháy rừng, dữ liệu do người dùng mạng xã hội tạo ra, v.v. phục vụ hay hỗ trợ cho việc thực hiện các mục tiêu phát triển bền vững (xem thêm Hassani và cộng sự, 2021).

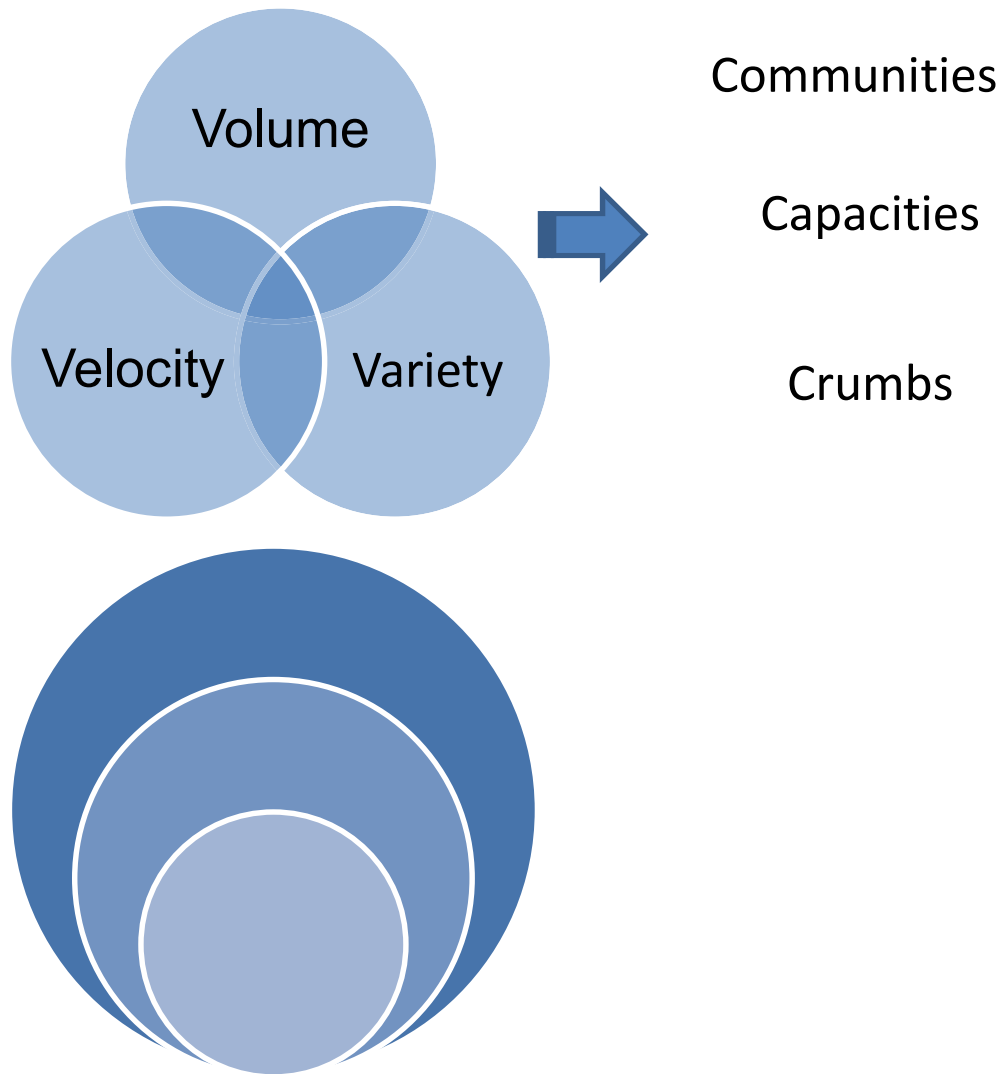
2.2 Từ mô hình 3V đến 3C của Big data

Big data có các đặc trưng 3V sau: dung lượng dữ liệu cực kỳ lớn (Volume) tính bằng petabytes hay exabytes, đa dạng về kiểu loại và nội dung (Variety) và yêu cầu về tốc độ, tần suất xử lý đến mức thời gian thực (Velocity). Vì thế cần thiết có phương pháp và công nghệ tiên tiến để thu thập, lưu trữ, phân phối, quản trị và phân tích Big data (Maarroof, 2015).

Kinh tế gia Letouzé (2015), King (2013) cùng nhiều chuyên gia khác của World Bank và nhóm công tác Big data cho SDGs của LHQ (Letouzé và Jütting, 2015) cho rằng Big Data⁸³ phải được tiếp cận như là một hệ sinh thái mới đặc trưng bởi sự hợp phần và chứa đựng của 3 khía cạnh (3C): Crumbs (dữ liệu thô), Capacities (khả năng phân tích, hỗ trợ ra quyết định bằng các mô hình dữ liệu, thuật toán và sức mạnh tính toán) và Communités (sự tham gia của các bên tham gia tương tác với hệ thống Big data). Theo đó, các đặc tính của Big data được mô tả chính xác hơn, ít gây bối rối, nhằm lẫn hơn quan điểm 3V của giai đoạn 2010 – 2012.

Hình 2 mô tả sự khác biệt chính yếu giữa 2 mô hình 3V và 3C của Big data. Chữ C đầu tiên (Crumbs) đề cập đến những mảnh vụn, rời rạc dữ liệu số (digital breadcrumbs) với hàm ý rằng dữ liệu thô tự nó không được tạo ra cho mục đích suy luận thống kê, thay vào đó đa số dữ liệu ngày nay được tạo ra bởi hoạt động đời sống của nhân loại mà 5 – 10 năm trước chưa hề có loại dữ liệu đó.

⁸³ Ở đây Data được Letouzé viết hoa chữ D hàm ý Big Data là một danh từ riêng nói về 1 hệ sinh thái.



Hình 2: Từ mô hình 3V đến 3C của Big data

Nếu dữ liệu cấu tạo nên hạt nhân của hệ sinh thái Big data thì chữ **C** thứ hai (Capacities) đại diện cho những công cụ, phương pháp phân tích, phần mềm và phần cứng thiết bị như Giáo sư Gary King của đại học Havard từng nhận định “Big Data không chỉ là về Big data (dữ liệu lớn)”. Như vậy năng lực ở đây chính là các hệ thống siêu máy tính, hệ thống dữ liệu phân bố và tính toán song song, và các thuật toán máy học và trí tuệ nhân tạo v.v. được sử dụng để phân tích sắc bén và thấu hiểu dữ liệu.

Cuối cùng, chữ **C** thứ 3 (Communitites) là các cộng đồng hàm ý các bên tham gia (stakeholders, như là người dùng cuối, người tạo ra dữ liệu, nhà nghiên cứu hay người ra quyết định chính sách), đối tượng ra quyết định bằng cách sử dụng hệ sinh thái Big Data để đạt một mục đích cụ thể và

Big Data là một hệ sinh thái phức tạp với vòng lặp thông tin phản hồi. Trong hệ sinh thái đó, phương pháp mới sẽ tạo ra dữ liệu mới; dữ liệu mới mang lại cho người dùng những ý tưởng khởi nghiệp mới, v.v. Vì thế làm thế nào tương tác với hệ sinh thái Big Data để có thể hưởng lợi và đổi mới từ nó là một trong những thách thức nhận được sự quan tâm gần đây từ giới học thuật và hoạch định chính sách.

Trên quan điểm tiếp cận 3C, dữ liệu được tạo ra từ các hoạt động sống của nhân loại có thể phân làm 3 nhóm chính: dữ liệu truy vết (Exhaust), dữ liệu nội dung số (Digital Content), và dữ liệu viễn thám và cảm biến (Sensing). Bảng 1 mô tả chi tiết một số phân loại dữ liệu và mục đích sử dụng cho phát triển bền vững (Alliance, 2017).

	Kiểu / Loại	Dữ liệu	Cơ hội sử dụng
Exhaust	Dữ liệu di động	<ul style="list-style-type: none"> - Bản ghi chi tiết cuộc gọi (CDRs) - Dữ liệu định vị toàn cầu (GPS) 	Ước lượng phân bố dân số và trạng thái kinh tế xã hội
	Giao dịch tài chính	<ul style="list-style-type: none"> - Định danh giao dịch - Hợp đồng điện tử - Thẻ tín dụng - Thẻ sử dụng dịch vụ vận tải 	Cung cấp dữ liệu quan trọng về tình hình di chuyển của dân cư và hành vi cư dân sau thảm họa
	Giao thông – vận tải	<ul style="list-style-type: none"> - Dữ liệu định vị toàn cầu (GPS) - Dữ liệu thu phí cầu đường 	Cung cấp cảnh báo sớm về các hư hại hạ tầng do thiên tai
	Truy vết trực tuyến	<ul style="list-style-type: none"> - Cookies - Địa chỉ IP 	Giảm thiểu tác động của bệnh truyền nhiễm thông qua giám sát kịp thời nhật ký truy cập Wikipedia
Digital Content	Truyền thông xã hội	<ul style="list-style-type: none"> - Tweets - Check-ins - Nội dung Facebook 	Cung cấp cảnh báo sớm cho các mối đe dọa từ dịch bệnh đến an toàn thực phẩm

		- Nội dung Youtube	
	Dữ liệu trực tuyến có nguồn gốc cộng đồng	- Bản đồ trực tuyến - Đánh dấu vị trí - Các báo cáo / cảnh báo (U-Report)	Hỗ trợ tình nguyện viên trong công tác xác thực ở hiện trường
Sensing	Vật lý	- Thiết bị đo thông minh - Thiết bị quan trắc tốc độ, khối lượng, thư từ,...	Phân tích dữ liệu quản lý vận chuyển hàng hóa để suy luận hình mẫu và khuynh hướng thương mại
	Từ xa	- Hình ảnh vệ tinh - Thiết bị không người lái	Hình ảnh vệ tinh có thể tiết lộ những thay đổi về chất lượng đất và nguồn nước để có giải pháp nông nghiệp kịp thời

Bảng 1: Phân loại các nguồn Big data cho phát triển bền vững

2.3 Vai trò của Big data đối với việc thực hiện các mục tiêu phát triển bền vững (SDGs)

Như đã đề cập ở trên, Chương trình Nghị sự 2030 nhận thấy để đạt được các SDGs rất cần dữ liệu chất lượng cao (xác đáng, tin cậy, có thể so sánh và kịp thời) để xác định độ rộng và sâu của những vấn đề chính sách và nhận dạng được đối tượng cư dân bị ảnh hưởng để định hướng xây dựng chính sách và ra quyết định đầu tư, giám sát tiến trình thực hiện các SDGs và đánh giá tác động của chính sách (Guo, 2017; Malhotra và cộng sự, 2018; Xu và cộng sự, 2020). Hơn nữa, như Van Halderen và cộng sự (2021) đã chỉ ra, có khoảng 65 chỉ tiêu thống kê gắn với 16 SDGs nhận được lợi ích thiết thực từ hệ sinh thái Big Data. Điều này có hàm ý rằng khoảng trống dữ liệu còn rất lớn. Tuy nhiên với sự phát triển mạnh mẽ của điện toán đám mây kết hợp với Big data, bức tranh về dữ liệu cho SDGs đang thay đổi theo chiều hướng mở ra nhiều cơ hội thu thập dữ liệu mới, từ đó ảnh hưởng tích cực đến các tiến trình thu thập dữ liệu truyền thống và hoạt động của các chương trình chính sách được nó giám sát (Wu và cộng sự, 2020). Bảng 2 dưới đây tóm tắt một số đóng góp trực tiếp của Big data vào công tác đo lường và theo dõi kết quả thực hiện các mục tiêu phát triển bền vững, trong khi đó những thực tiễn ứng dụng Big data thành công ở một số quốc gia đang phát triển điển hình sẽ được trình bày ở mục kế tiếp (Mục 3).

Mục tiêu	Đóng góp của Big data vào mục tiêu phát triển bền vững
SDG1: xóa nghèo	Dữ liệu chi tiêu cho sử dụng các dịch vụ điện thoại di động có thể là chỉ số đại diện cho mức thu nhập.
SDG2: không còn nạn đói	Dữ liệu về giá thực phẩm từ các nguồn cộng đồng hay các trang trực tuyến có thể giúp giám sát an ninh lương thực theo thời gian thực.
SDG3: sức khỏe và chất lượng sống tốt	Lập bản đồ dữ liệu di chuyển của người dùng di động có thể giúp dự báo mức độ lây lan của dịch bệnh.
SDG4: giáo dục có chất lượng	Báo cáo / phản hồi của công dân có thể tiết lộ lý do sinh viên / học sinh bỏ học.
SDG5: Bình đẳng giới	Phân tích dữ liệu giao dịch tài chính có thể tiết lộ hình mẫu chi tiêu và tác động khác nhau của cú sốc kinh tế lên phụ nữ và đàn ông.
SDG6: nước sạch và vệ sinh	Các cảm biến kết nối vào bơm nước có thể theo dõi tiếp cận nguồn nước sạch.
SDG7: năng lượng sạch với giá thành hợp lý	Thiết bị đo thông minh cho phép công ty năng lượng tăng cường hay hạn chế lưu lượng sử dụng để kéo giảm lãng phí và đảm bảo cung ứng vào các thời điểm cao điểm.
SDG8: công việc tốt và tăng trưởng kinh tế	Hình mẫu dữ liệu giao dịch thư tín toàn cầu có thể cung cấp những chỉ số thống kê về tăng trưởng kinh tế, kiều hối, thương mại và GDP.
SDG9: công nghiệp, sáng tạo, và phát triển hạ tầng	Dữ liệu từ các thiết bị định vị toàn cầu (GPS) có thể được sử dụng để điều khiển giao thông và cải thiện vận tải công cộng.
SDG10: thu hẹp bất bình đẳng	Phân tích dữ liệu giọng nói thành văn bản từ các nguồn radio địa phương có thể tiết lộ các quan ngại về phân biệt đối xử và hỗ trợ cho các phản ứng chính sách
SDG11: đô thị và cộng đồng bền vững	Dữ liệu quan trắc viễn thám có thể theo dõi sự xâm phạm đất công hay không gian công cộng như công viên và rừng. Dữ liệu người dùng điện thoại ẩn danh có thể giúp quy hoạch hiệu quả không gian đô thị và giao thông công cộng.

SDG12: tiêu thụ và sản xuất có trách nhiệm	Dữ liệu từ khóa tìm kiếm hay giao dịch thương mại điện tử có thể tiết lộ tốc độ chuyển đổi sử dụng các sản phẩm tiết kiệm năng lượng.
SDG13: hành động đối với biến đổi khí hậu	Kết hợp dữ liệu vệ tinh, dữ liệu mở và các bằng chứng từ cộng đồng có thể giúp theo dõi tình trạng phá rừng. Phân tích dữ liệu thời tiết và thu hoạch nông nghiệp có thể giúp dự báo sản lượng và điều kiện canh tác.
SDG14: tài nguyên và môi trường biển	Dữ liệu theo dõi tàu biển có thể tiết lộ các hoạt động đánh bắt bất hợp pháp, không được kiểm soát và không được báo cáo.
SDG15: tài nguyên và môi trường trên đất liền	Giám sát mạng xã hội có thể hỗ trợ quản lý thiên tai với thông tin thời gian thực về vị trí nạn nhân, ảnh hưởng và mức độ của cháy rừng hoặc khói mù.
SDG16: hoà bình, công lý và các thể chế mạnh mẽ	Phân tích dữ liệu xúc cảm trên kênh truyền thông xã hội có thể tiết lộ ý kiến của công chúng về hiệu quả quản lý nhà nước, dịch vụ công hoặc quyền con người.
SDG17: quan hệ đối tác vì các mục tiêu	Quan hệ đối tác để cho phép kết hợp số liệu thống kê, dữ liệu di động và internet có thể cung cấp hiểu biết về một thế giới siêu kết nối tốt hơn trong thời gian thực.

Bảng 2: Một số đóng góp của Big data vào thực hiện các mục tiêu bền vững.⁸⁴

3. MỘT SỐ KINH NGHIỆM TRIỂN KHAI BIG DATA CHO SDGs

Trên Thế Giới

3.1 Big data cho SDGs

Với mục đích tìm hiểu những bài học thành công của các quốc gia có trình độ phát triển tương đồng với Việt Nam trong triển khai các ứng dụng Big data cho mục tiêu phát triển bền vững, trong bài viết này, chúng tôi tổng hợp một số kết quả chính được phân tích bởi các tác giả Curry và cộng sự (2021), ElMassah và Mohieldin (2020), Giest (2017), Maarroof (2015), Malhotra và cộng sự (2018), Liengpunsakul (2020), Vinuesa và cộng sự

⁸⁴ Nguồn: UNGlobalPulse, 2017. Liên kết <http://www.unglobalpulse.org>. Truy cập ngày 04/06/2021.

(2020), Wu và cộng sự (2020), và Van Halderen và cộng sự (2021). Bên cạnh đó, ở tiểu mục 3.2 của phần này, chúng tôi thảo luận thêm về “*làm thế nào Big data hỗ trợ kiểm soát đại dịch COVID-19*”. Bảng 3 dưới đây trích lược một số dự án Big data dưới sự hỗ trợ của LHQ.

Quốc gia	Tình huống ứng dụng Big data cho SDGs
Nepal	<p>Tình huống: Thảm họa thiên nhiên là một trong những mối đe dọa lớn nhất đối với phát triển bền vững. Nepal, một quốc gia nghèo gánh chịu nhiều thiệt hại do động đất, ưu tiên phát triển những chính sách và công nghệ thực tiễn giúp tăng cường khả năng phòng chống và phục hồi sau thảm họa tự nhiên để giảm thiểu tác động của thiên tai đến đối nghèo và nỗ lực phát triển bền vững.</p> <p>Đơn vị công nghệ Flowminder triển khai công nghệ xác định vị trí địa lý của thẻ SIM điện thoại mỗi khi thực hiện cuộc gọi từ đó dựng nên bản đồ phân bố và đặc tính của những nhóm dân cư dễ bị tổn thương bởi thiên tai ở các nước có thu nhập trung bình và thấp. Ngày 25/04/2015 đã xảy ra thảm họa động đất 7.8 độ Richter gây thương vong hơn 32.000 người ở Nepal. Ngay lúc đó, Flowminder đã trợ giúp chính phủ Nepal, các cơ quan LHQ và tổ chức nhân đạo bằng những phân tích thời gian thực về sự dịch chuyển của người dân sử dụng dữ liệu ẩn danh người dùng di động được cung cấp bởi nhà mạng NCell. Thông qua các phân tích đó các tổ chức đã lên kế hoạch hỗ trợ nhân đạo một cách hiệu quả. Những phân tích sau đó của Flowminder cho thấy sau khi loại trừ yếu tố xuất cư tự nhiên thì đã có hơn nửa triệu người Nepal đã rời khỏi thung lũng Kathmandu (nơi xảy ra động đất) chỉ trong 2 tuần sau thảm họa.</p> <p>Kinh nghiệm: Công nghệ Big data có thể hỗ trợ rất hiệu quả cho công tác nhân đạo và khắc phục hậu quả thiên tai (SDG1, 2, 3, 11, 15). Phân tích sâu sắc lượng lớn dữ liệu người dùng di động có thể giúp nhà hoạch định chính sách hiểu hơn về hành vi con người về 3 khía cạnh: tính linh động, tương tác xã hội và hoạt động kinh tế.</p>

Sri Lanka	<p><u>Tình huống:</u> Những công nghệ Big data mới đang tạo ra diện mạo thông minh mới cho các đô thị và một hình mẫu công dân mới qua đó khuyến khích hoạt động xã hội và sự tham gia của người dân nhiều hơn vào hệ thống quản lý chính quyền của các đô thị.</p> <p>Tại Sri Lanka, LIRNEasia, một “think-tank”, đã phân tích dữ liệu lịch sử người dùng mạng di động (<i>đã được ẩn danh</i>) ở thành phố Colombo để hiểu biết sâu sắc về những hình mẫu và tần suất di chuyển của cư dân thành phố. Kết quả phân tích cho ranh giới hữu hình giữa các đơn vị hành chính cấp quận đã không còn tồn tại. Dưới góc độ chính sách giao thông đô thị, điều này gợi ý cần thiết lập những hành lang giao thông công cộng sử dụng những phương tiện vận chuyển có sức chứa lớn. Thêm vào đó, LIRNEasia cũng chỉ ra không gian địa lý của trung tâm kinh tế năng động của thành phố Colombo đang được mở rộng.</p> <p><u>Kinh nghiệm:</u> Dữ liệu người dùng điện thoại <i>ẩn danh</i> có thể được xem là mỏ dầu thông tin nếu khai thác hiệu quả sẽ mang lại nhiều thông tin quý giá giúp cho quy hoạch không gian và giao thông đô thị và điều tiết giao thông công cộng hiệu quả hơn (SDG 11). Bên cạnh đó, công nghệ Big data như của LIRNEasia nếu được triển khai ở khu vực / thành phố đông dân và có mật độ sử dụng điện thoại di động cao như TpHCM hay Hà Nội có thể là giải pháp thay thế tốt hơn, chi phí thấp hơn so với các phương pháp phân tích truyền thống dựa trên tổng điều tra hay khảo sát dân cư.</p>
Colombia	<p><u>Tình huống:</u> 57% dân số Colombia truy cập internet và đất nước này xếp hạng 55 thế giới về mức độ sẵn sàng công nghệ. Colombia đã thực hiện chuyển đổi số mạnh mẽ và triển khai nhiều ứng dụng Big data cho các SDGs. Cụ thể, trung tâm nghiên cứu nhiệt đới quốc tế (CIAT) đã phát triển một ứng dụng Big data hỗ trợ nông dân trồng lúa bằng cách kết hợp dữ liệu thu hoạch và thời tiết. Bằng những thuật toán máy học như <i>random forest</i> và <i>clustering</i> ứng dụng Big data xác định mối tương quan giữa các yếu tố khí hậu và sự biến đổi về sản</p>

	<p>lượng lương thực từ đó dự báo được những điều kiện thuận lợi và không thuận lợi cho mùa màng ở một số khu vực cụ thể. Toàn bộ những thông tin trợ giúp mùa màng được cập nhật đến mọi nông dân một cách dễ dàng qua ứng dụng trên hệ điều hành Android hay iOS.</p> <p><u>Kinh nghiệm:</u> Sáng kiến triển khai ứng dụng Big data hỗ trợ ngành nông nghiệp ở trực tiếp tác động đến tiến trình hiện thực SDG1 (xóa nghèo) và SDG2 (không còn nạn đói) đồng thời liên quan đến SDG13. Thành công của sáng kiến không chỉ ở khía cạnh ứng dụng Big data mà còn ở cung cấp khả năng tiếp cận thông tin dễ dàng cho đối tượng liên quan hưởng lợi từ kết quả phân tích – dự báo.</p>
Indonesia	<p><u>Tình huống:</u> Công nghệ Big data mang lại một sự đổi mới trong cách phân tích sắc bén dữ liệu về hành vi cư dân trong thời gian thực. Tại Indonesia, đơn vị nghiên cứu UN Global Pulse Lab đã khai phá dữ liệu mạng xã hội, cụ thể hơn là dữ liệu tweet của người dùng Twitter, để tìm hiểu thông tin về phân biệt đối xử trong môi trường làm việc (SDG 5) được xem hiệu quả hơn rất nhiều so với phương pháp truyền thống dựa trên các khảo sát.</p> <p><u>Kinh nghiệm:</u> Phân tích dữ liệu thời gian thực từ các nguồn truyền thông xã hội như Twitter, Facebook hay Zalo (ở Việt Nam) có thể hỗ trợ chính phủ và cộng đồng quốc tế hiểu hơn về động cơ phân biệt đối xử ở nơi làm việc để từ đó phát triển chính sách phù hợp nhằm ngăn ngừa những sự cố nhằm vào phụ nữ ở nơi công sở và đảm bảo môi trường làm việc tốt, công bằng với tất cả mọi người lao động. Tuy nhiên, để khai thác hiệu quả những dữ liệu này cần đảm bảo sự tuân thủ bảo mật dữ liệu bao gồm <i>án danh người dùng</i> và sự tuân thủ trách nhiệm xã hội của các doanh nghiệp tư nhân.</p>
Philippines	<p><u>Tình huống:</u> Philippines là quốc gia có mức độ phi tập trung hóa rất cao với 43.6% dân số truy cập internet và xếp hạng 55 về mức độ sẵn sàng công nghệ. Chính phủ Philippines có chiến lược thực hiện một cách hiệu quả SDG9 (công nghiệp, sáng tạo, và phát triển hạ tầng) và SDG11 (đô thị và cộng</p>

	<p>đồng bền vững). Vì thế xuất hiện nhiều sáng kiến khơi gợi tiềm năng của Big data hỗ trợ thực hiện các SDG9 và SDG1. Chẳng hạn, sáng kiến OpenRoads kết hợp dữ liệu video và hình được đánh dấu địa lý từ các nguồn cộng đồng qua các thiết bị di động để theo dõi chất lượng mạng lưới giao thông và các dự án đầu tư hạ tầng quốc gia.</p> <p><u>Kinh nghiệm:</u> Ứng dụng Big data cho các nguồn dữ liệu hình ảnh và video từ điện thoại di động của người dân hỗ trợ chính phủ và người dân cùng giám sát tiến độ triển khai và vòng đời của các dự án hạ tầng cơ sở từ đó khuyến khích địa phương hóa các dự án đầu tư.</p>
Singapore	<p><u>Tình huống:</u> Singapore là quốc gia có trình độ phát triển cao duy nhất của khối ASEAN đã đầu tư rất mạnh vào hạ tầng công nghệ thông tin quốc gia. Mức độ thâm nhập của internet và điện thoại di động ở Singapore lên đến hơn 130% vào năm 2013 với hơn 70% dân số thường xuyên sử dụng internet và xếp hạng chính phủ điện tử của Singapore luôn nằm trong Top 3 trong thập niên vừa qua. Sáng kiến xây dựng nền tảng quốc gia thông minh (the Smart Nation Platform, SNP) nằm trong chiến lược chuyển đổi số của quốc gia tập trung vào 3 trụ cột tăng trưởng: tăng tốc (accelerate), cạnh tranh (compete) và chuyển đổi (transform).</p> <p>SNP dựa trên công nghệ điện toán đám mây kết nối toàn bộ hệ thống thông tin và cảm biến trên toàn bộ lãnh thổ Singapore (Ng, 2018). Nền tảng này cho phép tập trung hóa và chia sẻ nhiều nguồn dữ liệu chính phủ khác nhau đồng thời cung cấp những năng lực tính toán, phân tích mới và tiên tiến. SNP kết nối tất cả các cảm biến thông qua hơn 12000 hộp gộp dữ liệu (Aggregation Gateway Box, AGB) đặt tại các khu vực công cộng và khu dân cư. Mỗi hộp gộp dữ liệu cung cấp kết nối đến đèn chiếu sáng công cộng, lưới mạng không dây để đảm bảo các cảm biến có thể dễ dàng cắm và nhúng vào hệ thống mạng quốc gia. Chính phủ Singapore không cấp phép cho tư nhân sở hữu hay khai thác các hộp gộp dữ liệu này.</p> <p>Về cơ bản, các AGB truyền tải toàn bộ dữ liệu quan trắc và nhạy cảm được thu thập bởi các cảm biến về trung tâm dữ</p>

	<p>liệu. Toàn bộ dữ liệu được bảo mật ở cấp độ cao nhất. Các cơ quan chính phủ sẽ được cấp quyền khai thác dữ liệu để tạo ra các phân tích sắc bén chuyên sâu phục vụ công tác dự báo, hoạch định chính sách, dịch vụ công ích và hoạt động của chính phủ điện tử.</p> <p><u>Kinh nghiệm:</u> dự án SNP của Singapore là một ví dụ điển hình của tham vọng phát huy sức mạnh điện toán đám mây và phân tích dữ liệu lớn phục vụ cho hoạt động chính phủ điện tử và hoạch định chính sách. Đất nước Singapore có diện tích chưa tới 1/3 diện tích của TpHCM và dân số chỉ khoảng hơn 5 triệu dân nên các thành phố trực thuộc trung ương như Hà Nội, TpHCM, Hải Phòng, Đà Nẵng và Cần Thơ hoặc tỉnh Bình Dương có thể học hỏi và xây dựng một nền tảng quản lý đô thị thông minh và phát triển bền vững tương tự.</p>
China	<p><u>Tình huống:</u> Trung Quốc là một trong những quốc gia đi đầu thế giới ứng dụng Big data và Trí tuệ nhân tạo ở cả hai khu vực công và tư. Trung Quốc đã cam kết thực thi toàn diện mục tiêu phát triển bền vững 2030 và nhanh chóng thu được những những kết quả tích cực ban đầu.</p> <p>Học viện Khoa học Trung Quốc (Chinese Academy of Sciences, CAS) đã triển khai Chương trình CAS Big Earth Data Science Engineering (CASEarth) từ năm 2018, một nỗ lực đẩy nhanh quá trình chuyển đổi hệ thống dữ liệu Trái đất sang Trái đất kỹ thuật số. Dự án thúc đẩy việc chia sẻ dữ liệu, kiến thức, và kinh nghiệm trên khắp thế giới, đồng thời hỗ trợ khám phá khoa học, đổi mới công nghệ và hoạch định chính sách, và trực tiếp hỗ trợ việc thực hiện SDG2, 6, 11, 13, 14 và 15 tại Trung Quốc. Vì dự án rất lớn, bao quát đa lĩnh vực nên trong khuôn khổ báo cáo khoa học chúng tôi chỉ nêu một tình huống triển khai điển hình và đề nghị đọc giả tham khảo thêm 26 tình huống ứng dụng khác ở Huadong (2021).</p> <p>Một nghiên cứu của chương trình CASEarth phân tích dữ liệu nhiệt độ cao nhất trong ngày thu được từ 754 trạm quan trắc trong khoảng thời gian từ 1960 đến 2018 để xác định khả năng tái hiện các sóng nhiệt và hiện tượng nhiệt độ khí hậu cao bất thường ở Trung Quốc. Kết quả nghiên cứu cung cấp một sự</p>

	<p>hiểu biết tốt hơn về tần suất xuất hiện các sóng nhiệt và hiện tượng nhiệt cao bất thường và chỉ ra cực đoan thời tiết xảy ra thường xuyên hơn trong giai đoạn sau năm 2000 (theo chu kỳ ngắn 5 năm hay vì 10 hoặc 20 năm như trước đây), và cường độ mạnh hơn ở khu vực phía bắc hơn phía nam.</p> <p>Đồng hành với sự phát triển vượt bậc của nền kinh tế Trung Quốc trong hơn 3 thập niên qua là làn sóng di cư khổng lồ từ nông thôn ra thành phố đã tạo nên các siêu đô thị và thu hẹp đáng kể diện tích mặt nước tự nhiên. Vì thế dự báo chính xác hơn chu kỳ xảy ra của các sóng nhiệt và hiện tượng nhiệt cực đoan và đánh giá chính xác hơn tác động của nó lên hoạt động kinh tế, đời sống nhân dân và hệ sinh thái tự nhiên là một nhiệm vụ của Chương trình CASEarth hỗ trợ thực hiện SDG 13.</p> <p><u>Kinh nghiệm:</u> trong tầm 10 năm gần đây, nhiều đợt nắng nóng, hạn hán hay hiện tượng thời tiết cực đoan xảy ra trên nhiều vùng ở nước ta gây nhiều thiệt hại và ảnh hưởng đến sản xuất nông nghiệp và sức khỏe người dân. Vì thế mở hóa các nguồn dữ liệu quan trắc và thu hút nghiên cứu ứng dụng Big data phân tích dữ liệu và đánh giá tác động kinh tế - xã hội của các hiện tượng thời tiết cực đoan sẽ hỗ trợ xây dựng các kịch bản hành động ứng phó biến đổi khí hậu tốt hơn.</p>
Europe (EU)	<p><u>Tình huống:</u> Phát triển bền vững gắn chặt vào các Hiệp ước châu Âu và luôn là trọng tâm chính sách của Châu Âu trong một thời gian dài. Đối với Chương trình Nghị sự 2030 về phát triển bền vững, Liên minh châu Âu (EU) cam kết đóng một vai trò tích cực trong việc tối đa hóa tiến độ hướng tới các SDGs.</p> <p>Có rất nhiều dự án về Big data phục vụ cho mục tiêu phát triển bền vững ở châu Âu ở cấp độ quốc gia và châu lục. Trong giai đoạn 2008 – 2018, Ủy ban châu Âu đã xây dựng một chiến lược về dữ liệu lấy con người làm yếu tố trọng tâm trong phát triển công nghệ, đồng thời bảo vệ và thúc đẩy các giá trị và quyền của châu Âu trong thế giới số, từ đó hình thành thị trường chung châu Âu về kỹ thuật số (the European Digital Single Market) với 3 trụ cột:</p>

	<p>i. Truy cập: cung cấp khả năng truy cập tốt hơn đến hàng hoá và dịch vụ kỹ thuật số cho người tiêu dùng và doanh nghiệp</p> <p>ii. Môi trường: tạo điều kiện thích hợp và sân chơi bình đẳng cho các mạng kỹ thuật số và các dịch vụ sáng tạo phát triển mạnh mẽ;</p> <p>iii. Kinh tế và xã hội: tối đa hoá tiềm năng tăng trưởng của nền kinh tế số.</p> <p>Dưới đây là một số dự án ứng dụng Big data tiêu biểu trong các lĩnh vực y tế, dữ liệu, vận tải, nông nghiệp và môi trường được triển khai trong khuôn khổ chiến lược Digital Single Market của kinh tế số EU phiên bản 2.0 (Zillner và cộng sự, 2021)</p> <p>a) Dự án BigMedilytics (Dữ liệu lớn cho Phân tích Y tế) là sáng kiến lớn nhất do EU tài trợ nhằm chuyển đổi lĩnh vực chăm sóc sức khỏe của khu vực bằng cách sử dụng các công nghệ dữ liệu lớn hiện đại để đạt được năng suất đột phá bằng cách giảm chi phí, cải thiện kết quả điều trị và khả năng tiếp cận tốt hơn đến các cơ sở y tế.</p> <p>BigMedilytics sử dụng hồ sơ sức khỏe của hơn 11 triệu bệnh nhân trên 8 quốc gia thuộc EU, truyền dữ liệu từ các thiết bị được kết nối Internet of Things (IoT) với hơn một triệu bản ghi mỗi giờ và dữ liệu do bệnh nhân tạo từ các ứng dụng di động. Dự án đảm bảo rằng tính bảo mật và quyền riêng tư của dữ liệu cá nhân được bảo vệ và quản lý trong khuôn khổ quy định của quốc gia và EU.</p> <p>Dự án BigMedilytics – được dẫn đầu bởi tập đoàn Philips có sự tham gia của 35 tổ chức y tế, dược phẩm, công nghệ, trường đại học và viện nghiên cứu thuộc 9 quốc gia trong Liên minh châu Âu, United Kingdom, Israel và Serbia – gồm 12 tiểu dự án thí điểm thực hiện những mục tiêu cụ thể sau:</p> <ul style="list-style-type: none"> - Cải thiện kết quả điều trị bệnh mãn tính và ung thư bằng dữ liệu lớn; - Tối ưu hoá quy trình làm việc thông qua công nghiệp hoá các dịch vụ chăm sóc sức khoẻ sử dụng Big data; - Đảm bảo khả năng nhân rộng của các khái niệm Big data cho chăm sóc sức khoẻ;
--	--

- Tăng cường hoạt động thông qua tích hợp dữ liệu;
- Thiết lập các dịch vụ chăm sóc sức khỏe an toàn và riêng tư xuyên biên giới và liên tổ, qua đó củng cố Chiến lược thị trường kỹ thuật số của EU
- Xác định các ứng dụng Big data tốt nhất;
- Cho phép chuyển giao tri thức.

b) Dự án DataPorts: Hàng trăm cảng biển châu Âu hợp tác và trao đổi dữ liệu kỹ thuật số từ nhiều nguồn dữ liệu khác nhau. Tuy nhiên, để đạt được hiệu quả hợp tác và hưởng lợi từ công nghệ dựa trên Trí tuệ nhân tạo, cần phải có một môi trường tích hợp mới. Dự án DataPorts, được EU tài trợ, thiết kế và xây dựng một nền tảng dữ liệu gọi là Cổng nhận thức (Cognitive Ports Data Platform) có vai trò kết nối cơ sở hạ tầng kỹ thuật số hiện có của các cảng biển, thiết lập các quy tắc về chia sẻ và giao dịch dữ liệu an toàn và tin cậy, đồng thời cung cấp các dịch vụ phân tích dữ liệu mạnh mẽ. Dựa trên nền tảng này các ứng dụng thông minh khác nhau sẽ được phát triển đáp ứng nhu cầu liên quan. Trước khi triển khai đầy đủ tại tất cả các cảng ở châu Âu, nền tảng này sẽ được triển khai tại hai cảng biển châu Âu (cảng Valencia ở Tây Ban Nha và cảng Thessaloniki ở Hy Lạp). Nền tảng này sẽ làm cho các cảng biển châu Âu trở nên đáng tin cậy, mạnh mẽ và hội nhập.

c) Dự án DataBio tập trung vào lĩnh vực kinh tế sinh học hướng dữ liệu (Data-driven Bioeconomy), cung cấp dữ liệu chuyên sâu về những nguyên liệu thô có thể sản xuất từ nông nghiệp, lâm nghiệp và ngư nghiệp cho công nghiệp kinh tế sinh học để sản xuất thực phẩm, năng lượng và vật liệu sinh học song song có tính đến các vấn đề khác nhau về trách nhiệm và bền vững. Dự án triển khai nền tảng Big data DataBio hiện đại trên cơ sở hạ tầng và giải pháp sẵn có của đối tác tham gia. Ngoài ra, dự án còn sử dụng công nghệ quan sát Trái đất và truyền thông khác.

Các ứng dụng thí điểm của dự án liên quan đến cải thiện hiệu quả canh tác ô liu, trái cây, nho và rau, cũng như ngũ cốc, sinh khối và cây trồng có chất xơ trong nông nghiệp; lâm nghiệp bao gồm các vấn đề như viễn thám phá hoại rừng, kiểm soát và giám sát các loài ngoại lai xâm hại và lập bản đồ web hồ

trợ chính phủ ra quyết định; lĩnh vực ngư nghiệp tập trung vào dự đoán và lập kế hoạch khai thác cá mặt nước đại dương.

d) Dự án Boost 4.0 là sáng kiến lớn nhất của EU về Big data cho công nghiệp 4.0. Boost 4.0 sẽ dẫn đầu việc xây dựng Không gian dữ liệu công nghiệp châu Âu để cải thiện khả năng cạnh tranh của công nghiệp 4.0 với mô hình nhà máy thông minh 4.0 và hướng dẫn ngành sản xuất Châu Âu trong việc ứng dụng Big data vào sản xuất, cung cấp cho khu vực công nghiệp các công cụ cần thiết để đạt được lợi ích tối đa của Big data.

Mục tiêu của dự án là thiết lập tiêu chuẩn toàn cầu, xây dựng cơ sở hạ tầng kỹ thuật số an toàn, tích hợp các sáng kiến Big data khác vào Không gian dữ liệu công nghiệp châu Âu (European Industrial Data Space, EIDS), xây dựng các nền tảng sản xuất nội dung số và cung cấp chứng nhận cho các thiết bị, dịch vụ Big data hoạt động trong EIDS.

Nhà máy thông minh 4.0 sẽ sử dụng các không gian dữ liệu công nghiệp để nâng cao hiệu quả thông qua việc năng lực xử lý dữ liệu tiên tiến của Trí tuệ nhân tạo; kết nối lực lượng lao động, tài sản và mọi thứ với Internet sẽ cho phép tận dụng bảo trì dự đoán để giảm thời gian ngừng hoạt động của thiết bị; tích hợp với các bộ phận sản xuất cho phép hiểu rõ hơn về kinh doanh, cải tiến sản xuất không lỗi và khả năng điều chỉnh sản xuất trong thời gian thực; cải thiện khả năng trao đổi dữ liệu giữa các doanh nghiệp cho phép các mô hình kinh doanh hợp tác.

Kinh nghiệm: EU có chiến lược và tầm nhìn phát triển ứng dụng Big data cho phát triển bền vững rất dài hạn. Các dự án Big data đã và đang triển khai đều mang lại lợi ích cho công dân và doanh nghiệp theo nhiều cách như: cải thiện chăm sóc sức khỏe, tạo ra các hệ thống giao thông an toàn hơn và sạch hơn, tạo ra các sản phẩm và dịch vụ mới, giảm chi phí dịch vụ công, cải thiện tính bền vững và hiệu quả năng lượng. Việt Nam có thể học hỏi EU kinh nghiệm hoạch định chiến lược kinh tế số, các tiêu chuẩn công nghiệp và kiến thức từ những dự án ứng dụng Big data trong y tế và nông nghiệp như BigMedilytics và DataBio.

Bảng 3: Tình huống và kinh nghiệm ứng dụng Big data ở các quốc gia đang phát triển, Trung Quốc và Châu Âu

Bên cạnh những ứng dụng Big data phân tích dữ liệu hỗ trợ hoạch định và ra quyết định chính sách nêu trên, Big data còn được sử dụng tổng hợp thông tin từ các nguồn đa dạng khác nhau để cung cấp nguồn dữ liệu chất lượng cao, tin cậy, kịp thời và có khả năng phân nhóm đối tượng theo thu nhập, giới tính, chủng tộc, v.v. cấu tạo nên các chỉ tiêu thống kê SDGs. Do giới hạn phạm vi của báo cáo khoa học này, chúng tôi mời độc giả tham khảo Bảng 1 của Van Halderen và cộng sự (2021) về các mục tiêu cụ thể (targets) và chỉ tiêu thống kê SDGs (indicators) có thể đo lường và theo dõi bằng các nguồn và ứng dụng Big data.

3.2 Làm thế nào Big data hỗ trợ kiểm soát đại dịch COVID-19

Đại dịch cúm SARS-CoV-2 hay COVID-19, khởi phát từ tháng 12 năm 2019 ở tỉnh Vũ Hán thuộc Trung Quốc, đã và đang gây rất nhiều khó khăn kinh tế - xã hội và đời sống nhân dân ở hầu hết quốc gia trên thế giới. Tính đến thời điểm 06/08/2021, đã có hơn 200 triệu người nhiễm bệnh và 4,2 triệu ca tử vong trên toàn cầu; Việt Nam, đang trải qua làn sóng COVID-19 thứ 4, có hơn 190.000 bệnh nhân và 3.000 trường hợp tử vong.⁸⁵

Sự phát triển nhanh chóng của dịch bệnh với hàng trăm triệu người nhiễm virus được báo cáo đã tạo ra một thách thức và cơ hội mới trong việc lưu trữ hồ sơ y tế bệnh nhân, thông tin dịch tễ, thông tin điều trị, thử nghiệm lâm sàng và dữ liệu giải mã gen virus, v.v.. Dung lượng dữ liệu rất lớn này, chứa đựng nhiều thông tin quý giá phục vụ cho các nghiên cứu dịch tễ, bào chế thuốc điều trị và vaccine, đòi hỏi giải pháp công nghệ tiên tiến như là Big data cho lưu trữ và phân tích dữ liệu (Haleem và cộng sự, 2020).

Bên cạnh đó, Big data có thể hỗ trợ tích cực giám sát lây nhiễm, bùng phát của dịch bệnh theo thời gian thực. Chẳng hạn, dữ liệu cập nhật về các ca nhiễm mới kết hợp với dữ liệu người dùng điện thoại di động là yếu tố đầu vào hoàn hảo cho các mô hình toán và trí tuệ nhân tạo dự báo sự lây lan của dịch bệnh trong phạm vi thành phố hay quốc gia (Bragazzi và cộng sự, 2020). Công cụ Big data xã hội (social Big data) đã được sử dụng khai thác dữ liệu trên các diễn đàn, mạng xã hội từ đó tái dựng lại thông tin dịch tễ của giai đoạn mới bùng phát dịch bệnh (Sun và cộng sự, 2020).

⁸⁵ Số liệu lấy từ Data Repository của Đại học Johns Hopkins, USA và Bộ Y tế Việt Nam ngày 07/08/2021.

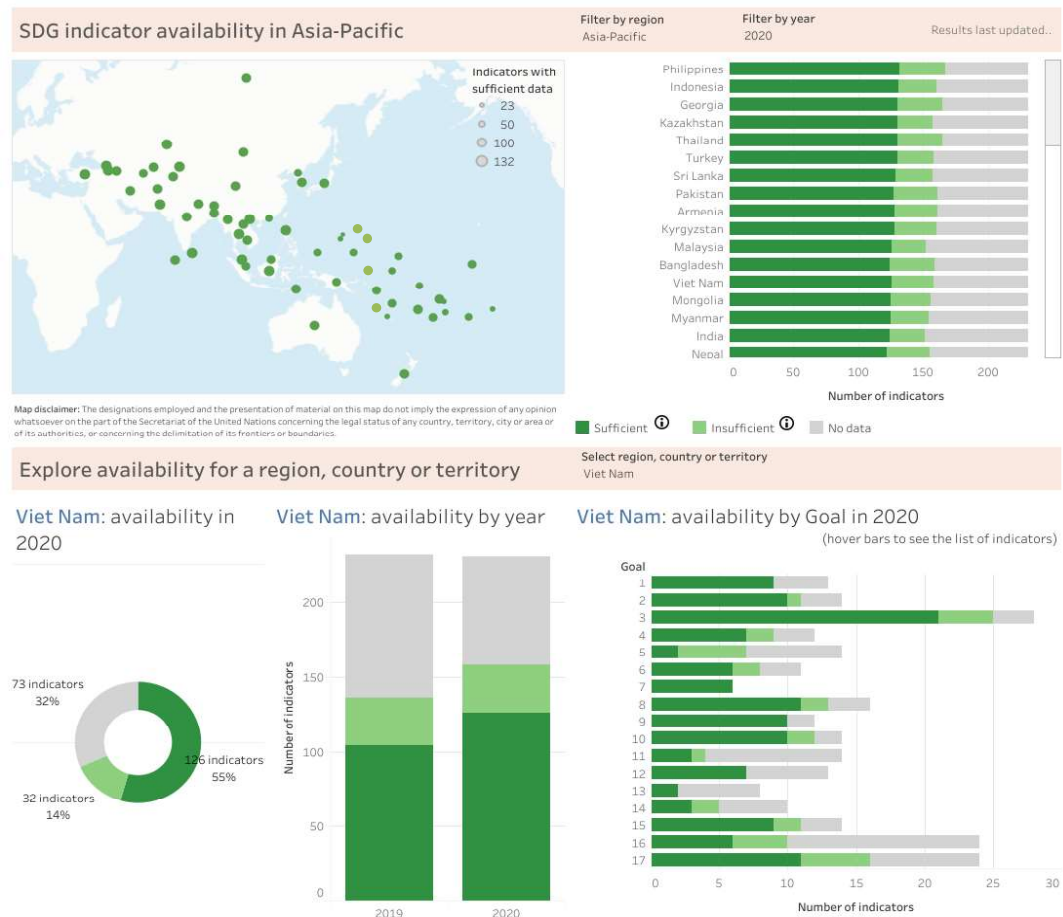
Ting và cộng sự (2020) nhận định rằng, trong đại dịch COVID-19, Big data đã cung cấp cơ hội thực hiện các nghiên cứu mô hình hóa hoạt động virus, hỗ trợ hiệu quả việc công tác kế hoạch giám sát y tế cộng đồng và chính sách ứng phó các đợt bùng phát lây nhiễm ở nhiều quốc gia trên thế giới. Tương tự, tại Việt Nam, dữ liệu người dùng điện thoại đã được sử dụng hỗ trợ truy vết các trường hợp tiếp xúc gần rất hiệu quả tuy nhiên đến nay trong nước vẫn chưa có một giải pháp công nghệ Big data ứng dụng cho y tế cộng đồng được phát triển hoàn chỉnh.

4. NHỮNG KHUYẾN NGHỊ CHO VIỆT NAM

Việt Nam đã sớm có những hành động triển khai và cam kết thực hiện hiệu quả Chương trình Nghị sự 2030 bởi vì các SDGs là sự tiếp nối các mục tiêu phát triển thiên niên kỷ (MDGs) đã nằm trong tiến trình thực hiện từ những năm đầu thế kỷ 21 (Tòng, 2019). Quyết định 622 ban hành Kế Hoạch Hành Động Quốc Gia (KHHĐQG) thực hiện Chương trình Nghị sự 2030 của Thủ tướng Chính phủ năm 2017 đã thể chế hóa 17 mục tiêu phát triển bền vững, bản địa hóa các 115 mục tiêu cụ thể và nêu giải pháp và phân công các bộ, ngành chủ trì thực hiện các mục tiêu và nhiệm vụ cụ thể; Bộ Kế hoạch và Đầu tư là đơn vị đầu mối chịu trách nhiệm giám sát, đánh giá và báo cáo. Đến năm 2019, Bộ KHĐT mới ban hành Thông tư 03/2019/TT-BKHĐT quy định chi tiết nội dung 158 chỉ tiêu thống kê với lộ trình thực hiện 2020 – 2025. Mặc dù có khoảng cách 2 năm giữa Quyết định 622 và Thông tư 03 nhưng khi so sánh với các nước trong khu vực ASEAN thì Việt Nam (xem Hình 3) xếp hạng 13 trên 53 quốc gia khu vực châu Á – Thái Bình Dương với 126 chỉ tiêu có đầy đủ dữ liệu (55% chỉ tiêu SDGs) và 32 chỉ tiêu chưa hoàn chỉnh hoặc thiếu dữ liệu (14% chỉ tiêu SDGs).

Mục 2 và 3 đã trình bày khái niệm 3C của Big Data và ứng dụng của Big data được LHQ triển khai để giám sát, quan trắc và hỗ trợ thực hiện các SDGs ở một số quốc gia đang phát triển. Ở Việt Nam, ứng dụng Big data trong thương mại điện tử và marketing trên các nền tảng Google, Facebook đã được các doanh nghiệp tiếp cận và triển khai thành công ngay từ những năm đầu thập niên 2010. Tuy vậy ứng dụng Big data cho phát triển bền vững, như trong quản lý và phát triển nông nghiệp, bảo vệ tài nguyên biển và rừng, trong y tế và giáo dục, trong quy hoạch không gian đô thị, trong phát triển hạ tầng giao thông và quy hoạch giao thông công cộng vẫn còn nhiều hạn chế và chưa được phổ biến cả trong thực tiễn và nghiên cứu.

Một thực tế, mặc dù Chính phủ rất quan tâm chuyển đổi số và ứng dụng Big data trong quản lý nhà nước nhưng đến nay hạ tầng công nghệ thông tin và cơ sở dữ liệu quốc gia đáp ứng cho hệ sinh thái Big Data chỉ đang trên tiến trình xây dựng và triển khai nhưng chưa đồng bộ, thiếu tính nhất quán và liên thông, chưa tạo được lòng tin của công chúng về bảo mật thông tin.⁸⁶ Tại thời điểm viết báo cáo này, tháng 08 năm 2021, các kế hoạch ứng dụng công nghệ thông tin trong hoạt động và phát triển chính quyền số được Bộ ngành và UBND tỉnh/thành lần lượt ban hành đều nhấn mạnh đến yếu tố bảo mật và đảm bảo an toàn thông tin mạng.



Hình 3: So sánh chỉ tiêu thống kê SDGs của Việt Nam và Khu vực Asia-Pacific⁸⁷

Trong phạm vi giới hạn về chủ đề ứng dụng Big data vào hỗ trợ thực hiện các SDGs, chúng tôi có một số nhận định với hi vọng đóng góp vào sự

⁸⁶ Xem thêm bài viết của Anh Cao (2021), Huyền Thanh (2016), Gia Minh (2021), GSO (2020), Nguyễn và Vũ (2021), và Vũ (2020).

⁸⁷ Biểu đồ thống kê được tạo bởi công cụ trực quan và dữ liệu của UNESCAP SDG Gateway. Truy cập 04/06/2021.

hình thành chính sách dài hạn quốc gia về chuyển đổi số cho một tương lai bền vững. Cụ thể:

Một là: *Xây dựng và duy trì hiệu quả cổng thông tin quốc gia về dữ liệu thống kê chính thức, thông tin tiến trình và kết quả thực hiện các mục tiêu phát triển bền vững*

Cho đến nay Tổng cục Thống kê trực thuộc Bộ Kế hoạch Đầu tư là cơ quan chịu trách nhiệm xây dựng cơ sở dữ liệu cho 158 chỉ tiêu thống kê phát triển bền vững tương ứng với 115 mục tiêu cụ thể của Việt Nam (Thông tư 03/2019/TT-BKHĐT). Thế nhưng trên cổng thông tin chính thức của Tổng cục Thống kê (tại địa chỉ <http://gso.gov.vn>) không có công bố số liệu về các chỉ tiêu thống kê này, cũng như không cung cấp phương tiện truy xuất dữ liệu SDGs dù đây là thông tin rất cần công bố rộng rãi, và báo cáo khoa học này truy xuất dữ liệu của GSO qua cổng thông tin của quốc tế là ESCAP (<http://data.unescap.org>), một đơn vị nghiên cứu của LHQ.

Cổng dữ liệu mở quốc gia tại địa chỉ <https://data.gov.vn/> là một nỗ lực khác của Chính phủ nằm trong chiến lược dữ liệu quốc gia do Bộ Thông tin và Truyền thông vận hành. Tại đây cung cấp nhiều nguồn dữ liệu mở khác nhau nhưng với giao diện và tổ chức thông tin hiện tại thì cổng dữ liệu mở chưa đáp ứng tiêu chuẩn là nguồn dữ liệu thống kê chính thức cũng như chưa có các thông tin về thực hiện các mục tiêu phát triển bền vững.

Báo cáo khuyến nghị của LHQ năm 2018 (UNECE, 2018) nhấn mạnh đến sự cần thiết và nhu cầu ngày càng tăng đối với dữ liệu thống kê quốc gia và địa phương chất lượng tốt, đáng tin cậy hỗ trợ cho các quyết định chính sách và phân tích chuyên sâu. Do đó, chúng tôi khuyến nghị:

(i) Xây dựng một cổng thông tin quốc gia cung cấp giao diện truy xuất dữ liệu và trực quan hóa dữ liệu thống kê quốc gia và các chỉ tiêu thống kê SDGs. GSO phải là đơn vị chịu trách nhiệm quản lý và duy trì hiệu quả cổng thông tin này, và xem đây là một trong những nhiệm vụ chiến lược của GSO. Mô hình tham khảo cho các cổng dữ liệu này có thể kể đến Eurostat, FRED, Viện thống kê quốc gia của Tây Ban Nha, Cục thống kê quốc gia của Trung Quốc, Đài Loan, Nhật Bản v.v.⁸⁸ Điểm cần lưu ý là nguồn dữ liệu này cần được truy xuất miễn phí để đạt hiệu quả khai thác tốt nhất.

⁸⁸ Địa chỉ truy cập các cổng dữ liệu lần lượt là: <https://ec.europa.eu/eurostat>; <https://www.ine.es/en/>; <http://www.stats.gov.cn>; <https://eng.stat.gov.tw/>; <https://www.stat.go.jp/english>

(ii) Dữ liệu thống kê ở các địa phương cần được tích hợp vào cổng thông tin quốc gia, truy xuất miễn phí và cập nhật định kỳ các chỉ số / chỉ tiêu kinh tế - xã hội và phát triển bền vững phục vụ cho công tác phân tích, nghiên cứu và hoạch định chính sách địa phương.

Có thể nói, hai đề xuất trên thoạt nhìn có vẻ trùng lặp với hệ thống thông tin thống kê hiện tại nhưng chúng tôi nhận định rằng để đạt được một kết quả có chất lượng và hiệu quả tương tự như các quốc gia vừa nêu trên cần quyết tâm chính trị và một kế hoạch đầu tư ổn định, bài bản từ trung ương đến các tỉnh / thành phố trực thuộc trung ương trong 5 – 10 năm tới trong đó: ngân sách cần chuẩn bị đầy đủ, xây dựng nền tảng công nghệ đồng bộ và có khả năng mở rộng ngay từ đầu, chú trọng phát triển nguồn nhân lực thống kê và phân tích dữ liệu chất lượng cao, tăng cường mạng lưới thu thập dữ liệu và tận dụng tối đa các nguồn dữ liệu mở, internet và mạng xã hội, v.v để tổng hợp và xây dựng thông tin thống kê có giá trị sử dụng (UNECE, 2018; van de Ven, 2018).

Hai là: *Xây dựng và công bố một số nguồn dữ liệu lớn chất lượng cao trong các lĩnh vực trọng điểm*

Một trong những thách thức khi triển khai ứng dụng Big data cho phát triển bền vững là nguồn dữ liệu, chất lượng, và độ tin cậy của nó (MacFeely, 2020). Đối với các mục tiêu SDG1 (xóa nghèo), SDG2 (không còn nạn đói), SDG3 (sức khỏe và chất lượng sống tốt), và SDG13 (hành động đối với biến đổi khí hậu), cần thiết xây dựng và công bố đại chúng cơ sở dữ liệu lịch sử quan trắc khí hậu như lượng mưa, nhiệt độ và ẩm độ, chất lượng không khí theo ngày / tháng / quý / năm của từng tỉnh / thành. Từng bước tiến tới xây dựng cơ sở dữ liệu chất lượng không khí (Air Quality Index, AQI) theo thời gian thực ở các khu vực trung tâm hoạt động sản xuất – kinh doanh ở các tỉnh / thành phố trực thuộc trung ương thông qua mạng lưới các cảm biến quan trắc các chỉ tiêu cơ bản chất lượng không khí: NO_x, CO, CO₂, SO₂, bụi mịn PM_{2.5} và PM₁₀, v.v. Lợi ích kinh tế - xã hội từ mạng lưới cảm biến quan trắc chất lượng không khí dựa trên công nghệ kết nối vạn vật vượt xa chi phí xây dựng và vận hành (Idrees và Zheng, 2020; Kumar và cộng sự, 2015).

Bên cạnh dữ liệu quan trắc khí hậu là dữ liệu viễn thám vệ tinh Việt Nam, hiện nay chưa được phép truy cập đại chúng và miễn phí, trong khi đó cơ sở dữ liệu Earth Observation của NASA hay Google Earth/Maps đều có thể truy cập dễ dàng và miễn phí. Do đó, để khuyến khích các nghiên cứu

ứng dụng trong lĩnh vực nông nghiệp, lâm nghiệp, thích ứng với biến đổi khí hậu và đói – nghèo rất cần xây dựng và cập nhật nguồn dữ liệu mở viễn thám địa lý (dĩ nhiên phải hạn chế một số thông tin liên quan đến quốc phòng – an ninh).

Cơ sở dữ liệu người dùng điện thoại di động là nguồn dữ liệu được khai thác rất hiệu quả ở nhiều nước trên thế giới. Tuy nhiên những lo ngại về tiết lộ danh tính khiến cho công chúng có nhiều hoài nghi vào mục đích sử dụng dữ liệu cũng như quan ngại bị xâm phạm quyền riêng tư. Năm 2016, ý tưởng về sử dụng dữ liệu điện thoại giám sát giao thông được đưa ra thảo luận nhưng đến nay chưa thể triển khai do vấp phải sự không đồng tình của công chúng về vấn đề bảo mật thông tin người dùng.⁸⁹ Để khai thác nguồn dữ liệu, dựa trên kinh nghiệm của các quốc gia được đề cập ở Mục 3, cần quy định chặt chẽ về bảo mật thông tin cá nhân và xử phạt nghiêm khắc các hành vi lạm dụng thông tin cá nhân cho các mục đích xấu, lừa đảo; ban hành khung pháp lý khai thác dữ liệu người dùng điện thoại ẩn danh cho các ứng dụng Big data hỗ trợ giảm thiểu tác động thảm họa tự nhiên, cứu hộ - cứu nạn, và quy hoạch và giám sát giao thông công cộng; tạo dựng niềm tin với công chúng bằng cách công khai những kết quả ứng dụng trên cổng thông tin điện tử của dự án hay ứng dụng Big data; phổ biến kiến thức về các ứng dụng Big data được triển khai đến công chúng và duy trì hiệu quả kênh đối thoại với công chúng về vấn đề bảo mật danh tính người dùng và an toàn thông tin.

Ba là: *Đầu tư đồng bộ, liên thông hạ tầng công nghệ Big data và khuyến khích sự tham gia đầu tư của khối doanh nghiệp tư nhân*

Báo cáo của nhóm công tác LHQ về ứng dụng Big data cho SDGs đã dự báo cần đầu tư ít nhất 1 tỷ USD hàng năm cho 77 quốc gia đang phát triển nâng cao năng lực thống kê, lưu trữ và khai thác hiệu quả các chỉ tiêu SDGs và 45 – 50 tỷ USD cho cơ sở dữ liệu phục vụ thực hiện SDGs. Dự thảo Chiến lược phát triển Thống kê Việt Nam giai đoạn 2021 – 2030 và tầm nhìn đến 2045 của GSO nhấn mạnh đến “xây dựng cơ sở hạ tầng tạo nền tảng cho chuyển đổi số công tác thống kê; ứng dụng trí tuệ nhân tạo, khoa học dữ liệu, dữ liệu lớn,...vào công tác thống kê,...” (Lan Phương, 2021).

Theo chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng Trí tuệ nhân tạo, trong thập niên 2020s Việt Nam sẽ đầu tư mạnh mẽ vào hạ

⁸⁹ Xem thêm bài viết của Trí Dũng (2016) và Thanh Thúy (2019).

tầng công nghệ thông tin đáp ứng nhu cầu mới về thu thập, lưu trữ và khai thác Big data. Do đó, chúng tôi cho rằng để chiến lược phát triển nêu trên thành công cần đầu tư hạ tầng công nghệ Big data dựa trên mô hình 3C (dữ liệu, năng lực phân tích và sự tham gia của cộng đồng) đồng thời phải đảm bảo tính đồng bộ, có tính kế thừa, có thể linh hoạt mở rộng cùng với sự liên thông dữ liệu giữa các Bộ ngành và địa phương.

Chi phí đầu tư hạ tầng đã lớn nhưng chi phí duy trì hệ thống vận hành hiệu quả hàng năm cũng không hề nhỏ trong khi tốc độ đổi mới của lĩnh vực này rất cao. Ứng dụng Big data hay trí tuệ nhân tạo rất cần các hệ thống xử lý song song có năng lực tính toán lớn, nhưng với nguồn lực có hạn của nhà nước thì đầu tư công vào các hệ thống siêu máy tính có thể không đạt được hiệu suất hoạt động như mong đợi do chi phí điện năng cao và vòng đời hoạt động các thiết bị chỉ từ 3 – 5 năm. Ở một phương diện khác, các doanh nghiệp CNTT tư nhân có đủ tiềm lực thực hiện đầu tư và vận hành hiệu quả các trung tâm tính toán hiệu năng cao do họ có thể cung ứng nhiều loại sản phẩm dịch vụ khác nhau, không nhất thiết phải thực hiện các nghĩa vụ bắt buộc như các đơn vị nghiên cứu khu vực công. Do đó, Chính phủ nên khuyến khích đầu tư tư nhân vào các trung tâm dữ liệu, trung tâm tính toán hiệu năng cao phục vụ nghiên cứu khoa học hoặc các đơn vị nghiên cứu chuyên sâu ứng dụng Big data cho phát triển trong lĩnh vực nông nghiệp, y tế, công nghệ sinh học và tự động hóa bằng các hình thức khuyến khích như miễn/giảm tiền thuê đất khu công nghiệp công nghệ cao, hỗ trợ chi phí khởi nghiệp, ưu tiên tiêu thụ sản phẩm công nghệ Big data, giảm thuế thu nhập cá nhân đối với chuyên gia có trình độ cao / cao cấp và tài trợ các dự án nghiên cứu của doanh nghiệp.

Bốn là: Tăng cường đầu tư, thúc đẩy nghiên cứu khoa học và ứng dụng Big data cho phát triển bền vững

Yếu tố thứ ba bên cạnh cơ sở hạ tầng thông tin và nguồn Big data chất lượng cao, đáng tin cậy chính là yếu tố “con người”, là nguồn nhân lực chất lượng cao tham gia vào công tác vận hành, phân tích, nghiên cứu và phổ biến tri thức tạo ra từ hệ sinh thái Big Data. Đây chính là yếu tố tạo nên sự khác biệt lớn giữa mô hình 3V và 3C, vì thế mô hình sau còn được gọi là mô hình *Big data cho phát triển*.

Trước hết, để Big Data, một lần nữa chúng tôi nhấn mạnh đến hệ sinh thái Big data, thực sự phục vụ cho mục tiêu phát triển bền vững, Việt Nam cần thúc đẩy nghiên cứu khoa học có công bố quốc tế sử dụng nguồn Big

data trong nước phù hợp với chiến lược tăng trưởng bao trùm và bền vững. Dữ liệu từ Scopus cho thấy chỉ 590 / 123.059 (~ 4.8%) bài báo khoa học có (một phần) nguồn gốc từ Việt Nam viết về các chủ đề liên quan đến kinh tế tuần hoàn, khoa học môi trường, năng lượng và quản trị bền vững, trong đó 6 / 471 (~1,3%) bài báo có đề cập và ứng dụng phân tích Big data. Tuy nhiên, trong sáu bài báo khoa học ghi nhận Việt Nam là quốc gia liên kết nghiên cứu chỉ có 1 tác giả là người Việt Nam.⁹⁰ Vì thế cần có sự khuyến khích và ưu tiên ngân sách nghiên cứu cho các tổ chức giáo dục đại học / Viện nghiên cứu có hoạt động nghiên cứu trực tiếp đến công nghệ và ứng dụng Big data phục vụ mục tiêu phát triển bền vững 2030.

Thực tế cho đến thời điểm 06/2021, Việt Nam chưa có một trung tâm dữ liệu và tính toán hiệu năng cao đạt chuẩn quốc tế. Quyết định số 127/QĐ TTg xác định chiến lược và lộ trình đưa Việt Nam vào nhóm 4 nước dẫn đầu trong khu vực ASEAN và nhóm 50 nước dẫn đầu thế giới về nghiên cứu, phát triển và ứng dụng Big data, Trí tuệ nhân tạo. Theo đó, đến năm 2025 sẽ đầu tư 1 trung tâm dữ liệu lớn và tính toán hiệu năng cao, phần đầu đến 2030 sẽ hình thành hoàn chỉnh 3 trung tâm như vậy cùng với 10 cơ sở nghiên cứu trọng điểm về Big data và Trí tuệ nhân tạo và 50 bộ dữ liệu mở có tính liên thông và kết nối đa lĩnh vực. Một mô hình tốt có thể học hỏi để xây dựng trung tâm dữ liệu và ứng dụng Big data cho mục tiêu phát triển bền vững quốc gia là Trung tâm nghiên cứu quốc tế về Big Data cho Mục tiêu phát triển bền vững của Học viện Khoa học Trung Hoa (CAS' International Research Center of Big Data for SDGs).

Do đó, để tương đồng với chiến lược quốc gia, trong 5 năm tới cần thiết hình thành và duy trì các trung tâm dữ liệu và tính toán hiệu năng cao có quy mô phù hợp đặt tại các phòng thí nghiệm trọng điểm ở cơ sở giáo dục đại học phục vụ cho công tác nghiên cứu và triển khai thử nghiệm ứng dụng Big data và Trí tuệ nhân tạo ở bốn khu vực: Hà Nội, TpHCM, Đà Nẵng và Cần Thơ; đồng thời thúc đẩy các Sở KH-CN ở các địa phương đặt hàng và giao các nhiệm vụ khoa học ứng dụng Big data vào công tác quản lý đô thị, giao thông công cộng, dự báo sản lượng nông nghiệp, quan trắc khí hậu và cảnh báo thảm họa tự nhiên.

⁹⁰ Tác giả truy vấn cơ sở dữ liệu Scopus với câu lệnh: (AUTHKEY ("sustainable") OR AUTHKEY ("sustainability") OR AUTHKEY ("sustainable development") OR AUTHKEY ("sustainability management") OR AUTHKEY ("circular economy")) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "re")) sau đó lọc dữ liệu quốc gia liên kết của Việt Nam (LIMIT-TO (AFFILCOUNTRY , "Viet Nam")). Kết quả lần lượt là 123.059 và 590 bài báo khoa học.

Tăng cường đầu tư trực tiếp vào các dự án khởi nghiệp ứng dụng Big data trong các lĩnh vực nông – lâm – ngư nghiệp, năng lượng, y tế và giáo dục tương thích với SDG 1, 2, 3, 6, 7, 13, 14 và 15. Chính phủ có thể học hỏi kinh nghiệm của Singapore⁹¹ lập quỹ đầu tư mạo hiểm quốc gia thực hiện nhiệm vụ đầu tư chiến lược vào các start-up phát triển các ứng dụng Big data phù hợp với mục tiêu trên. Hơn hết, Việt Nam cần thu hút sự tham gia tích cực của cộng đồng (chữ C thứ ba của mô hình 3C gồm: giới nghiên cứu khoa học, các tổ chức đa phương, khối kinh tế tư nhân) vào các sáng kiến ứng dụng Big data cho các SDGs dựa theo kinh nghiệm của các quốc gia trong khu vực và thế giới như được trình bày ở Bảng 3 của Mục 3.

Năm là: *Bảo mật thông tin và trách nhiệm giải trình*

Trong bối cảnh xu hướng thiên tả, tin giả, thông tin sai lệch, lỗ hổng bảo mật và rò rỉ thông tin diễn ra tràn lan đến mức khó kiểm soát trên không gian mạng ở Việt Nam và thế giới khiến niềm tin của công chúng vào việc khai thác dữ liệu người dùng phục vụ cho các dự án phát triển bền vững của Chính phủ ngày càng suy giảm. Sự hoài nghi này ngày càng trở nên tiêu cực khiến người dùng có xu hướng không sẵn sàng chia sẻ dữ liệu sử dụng điện thoại di động, lịch sử lướt web hay dữ liệu y tế cá nhân phục vụ các nghiên cứu khoa học.

Bên cạnh các tiêu chuẩn đạo đức y sinh được rộng rãi công nhận của Beauchamp và Childress (2001) và đạo đức khai thác Big data được nhiều chuyên gia khác như Richards và King (2014), Richterich (2018) và Nair (2020) phân tích mà trong báo cáo này chúng tôi không đi sâu vào chi tiết, bảo mật thông tin là điều Chính phủ có thể làm được và kiểm soát được thể hiện cam kết với cộng đồng rằng tất cả dữ liệu người dùng, thông tin cá nhân đều được khai thác đúng mục đích phục vụ nghiên cứu khoa học mang lợi ích và phúc lợi chính đáng cho công chúng.

Trách nhiệm giải trình cũng hết sức quan trọng vì công chúng cần được thông báo, giải thích, bảo vệ và cập nhật thông tin về các dự án ứng dụng Big data có khai thác dữ liệu cá nhân ẩn danh. Các nghiên cứu của Dubois và cộng sự (2018), Popham và cộng sự (2018) hoặc Hargittai (2018), Shahin và Zheng (2018) và Smith và cộng sự (2016) đã chỉ ra những mối quan tâm lớn của cộng đồng về các loại ứng dụng này bao gồm nhưng không giới hạn: tính công khai của đạo đức dữ liệu, các nguy cơ tiềm

⁹¹ Xem tại <https://www.nrf.gov.sg/funding-grants/early-stage-venture-fund>. Truy cập ngày 25/07/2021.

tàng xâm phạm quyền riêng tư, khả năng gạt bỏ ý kiến của những (nhóm) cộng đồng cư dân nhỏ và khác biệt dẫn đến thiên vị hay làm sai lệch mục tiêu chính sách, mục đích sử dụng thật sự dữ liệu của chương trình/dự án, sự cần thiết của các quy tắc cơ bản cho thực hành đạo đức dữ liệu, v.v.

Vì thế, bằng cách minh bạch thông tin các dự án, công khai chính sách bảo mật và giải trình với các bên liên quan về quản trị đạo đức khai thác Big data, xây dựng hành lang pháp lý chặt chẽ bảo vệ thông tin cá nhân, Chính phủ sẽ thuyết phục được công chúng tin tưởng, ủng hộ và tích cực tham gia vào các dự án Big data phục vụ cho mục tiêu phát triển bền vững.

5. KẾT LUẬN

Vai trò hỗ trợ của Big data đối với các SDGs của Chương trình Nghị sự 2030 đã được LHQ khẳng định qua các nghiên cứu, thử nghiệm ứng dụng thực tế ở nhiều quốc gia khác nhau. Việt Nam là một trong những quốc gia đang phát triển có cam kết mạnh mẽ thực hiện các SDGs vì thế có nhiều lợi ích khi ứng dụng Big data cho việc theo dõi và giám sát thực hiện 115 mục tiêu cụ thể. Lợi ích này đã được nhận diện trong Dự thảo Chiến lược phát triển Thống kê Việt Nam giai đoạn 2021 – 2030 và tầm nhìn đến 2045 của GSO.

LHQ định hình khái niệm Big data cho phát triển hay 3C (Crumbs, Capacities và Communities) Big data hàm ý về hệ sinh thái Big Data gồm 3 yếu tố chứa đựng vào nhau: dữ liệu, năng lực phân tích và cộng đồng tích cực tham gia. Vì thế, chúng tôi khảo cứu một số tình huống điển hình triển khai ứng dụng Big data cho phát triển bền vững ở các quốc gia đang phát triển (Colombia, Nepal, Sri Lanka, Indonesia và Phillipines) và quốc gia, khu vực kinh tế có trình độ phát triển cao hơn Việt Nam như China, Singapore và Liên minh Châu Âu, từ đó đưa ra những gợi ý chính sách phát triển và khai thác hiệu quả nguồn Big data trong nước cho hỗ trợ thực hiện các SDGs 2030.

Tổng hợp các tình huống cho thấy, những dự án Big data được triển khai ở các quốc gia đang phát triển tập trung vào các vấn đề giảm thiểu và khắc phục hậu quả thiên tai, phát triển nông lâm và ngư nghiệp bền vững, thích ứng và dự báo tác động biến đổi khí hậu bằng cách phân tích dữ liệu người dùng di động, quan trắc thời tiết, viễn thám Trái đất, v.v. Trong khi đó, ở quốc gia phát triển, các ứng dụng Big data đang hỗ trợ giải quyết nhiều bài toán về nâng cao chất lượng chăm sóc sức khỏe, cải thiện hiệu quả logistics, khai thác năng lượng bền vững, quản lý đô thị thông minh và

tối đa hoá tiềm năng tăng trưởng của nền kinh tế số. Nguồn dữ liệu cho các dự án này phức tạp hơn, lưu lượng và tần suất cao hơn, cần nhiều cảm biến kết nối với hạ tầng công nghệ thông tin hiện đại.

Đúc rút từ những kinh nghiệm thành công, chúng tôi hình thành năm gợi ý chính sách đóng góp vào chiến lược quốc gia về phát triển bền vững và phát triển ứng dụng Big data phục vụ mục tiêu phát triển bền vững của Chương trình Nghị sự 2030. Cụ thể, (i) Việt Nam cần đầu tư bài bản xây dựng và duy trì hiệu quả công nghệ thông tin thống kê quốc gia, cung cấp kịp thời dữ liệu về tiến trình thực hiện các SDGs; (ii) tập hợp và công bố miễn phí trước tiên các nguồn dữ liệu chất lượng cao có sẵn về quan trắc thời tiết và vệ tinh viễn thám; (iii) đầu tư đồng bộ, liên thông hạ tầng công nghệ Big data và khuyến khích sự tham gia đầu tư của khối doanh nghiệp tư nhân; (iv) thúc đẩy nghiên cứu khoa học và ứng dụng Big data cho phát triển bền vững; và (v) củng cố hành lang pháp lý về bảo vệ quyền riêng tư dữ liệu, cải thiện trách nhiệm giải trình để khuyến khích sự tham gia tích cực của cộng đồng vào các dự án Big data trong lĩnh vực y tế, giáo dục và quản lý đô thị, v.v.

Trong giới hạn của bài báo khoa học, chúng tôi chưa đề cập hay phân tích chuyên sâu kết quả ứng dụng Big data đang trong quản lý nhà nước của các Chính phủ trên toàn cầu, hay bỏ sót những kinh nghiệm triển khai ứng dụng Big data thành công xuất sắc của LHQ tại nhiều quốc gia trên thế giới. Tuy nhiên, chúng tôi nhận thấy nghiên cứu có thể được mở rộng bằng cách xem xét kỹ lưỡng hơn tính khả thi về mặt công nghệ và ứng dụng của các dự án Big data cho phát triển bền vững của LHQ, châu Âu và các tổ chức đa phương phi chính phủ vào điều kiện kinh tế, xã hội và trình độ công nghệ của Việt Nam.

TÀI LIỆU THAM KHẢO

1. Alliance, D. P. (2017). Leveraging Big Data for Sustainable Development: Toolkit. Unpublished document.
2. Baum, Anja. (2020). Vietnam's Development Success Story and the Unfinished SDG Agenda. IMF Working Paper WP/20/31.
3. Beauchamp, T. L., & Childress, J. F. (2001). Principles of biomedical ethics. Oxford University Press, USA.
4. Bragazzi, N. L., Dai, H., Damiani, G., Behzadifar, M., Martini, M., & Wu, J. (2020). How big data and artificial intelligence can help better manage the COVID-19 pandemic. *International journal of environmental research and public health*, 17(9), 3176.
5. Campbell, J. (2019). We lack data for 68% of SDG Indicators. <https://un-spbf.org/guest-insights/jillian-campbell/we-lack-data-for-68-of-sdg-indicators-closing-data-gaps-essential-to-achieving-sdgs/> (truy cập 31/05/2021).
6. Curry, E., Metzger, A., Zillner, S., Pazzaglia, J. C., & García Robles, A. (2021). The Elements of Big Data Value: Foundations of the Research and Innovation Ecosystem.
7. Dubois, E., Gruzd, A., Jacobson, J. (2018). Journalists' use of social media to infer public opinion: The citizens' perspective. *Social Science Computer Review*.
8. ElMassah, S., & Mohieldin, M. (2020). Digital transformation and localizing the Sustainable Development Goals (SDGs). *Ecological Economics*, 169, 106490.
9. Favaretto, M., De Clercq, E., Schneble, C. O., & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PloS one*, 15(2), e0228987.
10. King, G. (2013). Big data is not about the data!. Golden Seeds Innovation Summit.
11. Giest, S. (2017). Big data analytics for mitigating carbon emissions in smart cities: opportunities and challenges. *European Planning Studies*, 25(6), 941-957

12. Guo, H. (2017). Big Earth data: A new frontier in Earth and information sciences. *Big Earth Data*, 1(1-2), 4-20.
13. Haleem, A., Javaid, M., Khan, I. H., & Vaishya, R. (2020). Significant applications of big data in COVID-19 pandemic. *Indian journal of orthopaedics*, 54, 526-528.
14. Hargittai, E. (2018). Potential biases in big data: Omitted voices on social media. *Social Science Computer Review*.
15. Hassani, H., Huang, X., MacFeely, S., & Entezarian, M. R. (2021). Big Data and the United Nations Sustainable Development Goals (UN SDGs) at a Glance. *Big Data and Cognitive Computing*, 5(3), 28.
16. Huadong, G. (2021). Big Earth data in support of the sustainable development goals. *EDP Sciences*.
17. Idrees, Z., & Zheng, L. (2020). Low cost air pollution monitoring systems: A review of protocols and enabling technologies. *Journal of Industrial Information Integration*, 17, 100123.
18. Jones, A., Mair, S., Ward, J., Druckman, A., Lyon, F., Christie, I., & Hafner, S. (2016). Indicators for sustainable prosperity? Challenges and potentials for indicator use in political processes.
19. Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., ... & Britter, R. (2015). The rise of low-cost sensing for managing air pollution in cities. *Environment international*, 75, 199-205.
20. Letouzé, E. (2015). *Big Data & Development. An Overview. Data-POP Alliance Primers Series*.
21. Letouzé, E., & Jütting, J. (2015). *Official statistics, big data and human development. Data-Pop Alliance White Paper Series. Data-Pop Alliance, New York*.
22. Liengpunsakul, S. (2020). Artificial Intelligence and Sustainable Development in China. *The Chinese Economy*, 1-14.
23. Maarroof, A. (2015). *Big data and the 2030 agenda for sustainable development. Report for UN-ESCAP*.

24. MacFeely, S. (2020). Measuring the sustainable development goal indicators: An unprecedented statistical challenge. *Journal of official statistics*, 36(2), 361-378.
25. Malhotra, C., Anand, R., & Singh, S. (2018). Applying big data analytics in governance to achieve sustainable development goals (SDGs) in India. In *Data Science Landscape* (pp. 273-291). Springer, Singapore.
26. Nair, S. R. (2020). A review on ethical concerns in big data management. *International Journal of Big Data Management*, 1(1), 8-25.
27. Ng, R. (2018). Cloud computing in Singapore: Key drivers and recommendations for a smart nation. *Politics and Governance*, 6(4), 39-47.
28. Popham, J., Lavoie, J., Coomber, N. (2018). Constructing a public narrative of regulations for big data and analytics: Results from a community-driven discussion. *Social Science Computer Review*.
29. Pulse (2013). *Big Data for Development: A primer*.
30. Quyết định số 622/QĐ-TTg ngày 10/5/2017 của Thủ tướng Chính phủ về việc Ban hành Kế hoạch hành động quốc gia thực hiện Chương trình Nghị sự 2030 vì sự phát triển bền vững.
31. Quyết định số 127/QĐ-TTg ngày 26/1/2021 của Thủ tướng Chính phủ về việc Ban hành Chiến lược quốc gia về nghiên cứu, phát triển và ứng dụng Trí tuệ nhân tạo đến năm 2030.
32. Richterich, A. (2018). The big data agenda: Data ethics and critical data studies (p. 154). University of Westminster Press.
33. Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest L. Rev.*, 49, 393.
34. Schmidt-Traub, G., & Shah, A. (2015). Investment needs to achieve the sustainable development goals. Paris and New York: Sustainable Development Solutions Network.
35. Shahin, S., Zheng, P. (2018). Big data and the illusion of choice: Comparing the evolution of India's aadhaar and China's social credit system as technosocial discourses. *Social Science Computer Review*.

36. Smith, M., Patil, D., Muñoz, C. (2016). Big risks, big opportunities: The intersection of big data and civil rights. White House, Washington, DC.
37. Sun, K., Chen, J., & Viboud, C. (2020). Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health*, 2(4), e201-e208.
38. Thông tư 03/2019/TT-BKHĐT của Bộ Kế hoạch và Đầu tư quy định Bộ chỉ tiêu thống kê phát triển bền vững của Việt Nam.
39. Ting, D. S. W., Carin, L., Dzau, V., & Wong, T. Y. (2020). Digital technology and COVID-19. *Nature medicine*, 26(4), 459-461.
40. UN Desa (2016). Transforming our world: The 2030 agenda for sustainable development.
41. UNECE (2018). Value of Official Statistics: Recommendations on Promoting, Measuring and Communicating the Value of Official Statistics. Available at <https://unece.org/DAM/stats/publications/2018/ECECESSTAT20182.pdf>
42. van de Ven, P. (2018). Economic Statistics: How to Become Lean and Mean?. *Journal of official statistics*, 34(2), 309.
43. Van Halderen G., Bernal I., Sejersen T., Jansen R., Ploug N., & Truszczynski M. (2021) Big Data for the SDGs: Country examples in compiling SDG indicators using non-traditional data sources. UNESCAP Working Paper Series SD/WP/12/January 2021.
44. Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature communications*, 11(1), 1-10.
45. Wu, B., Tian, F., Zhang, M., Zeng, H., & Zeng, Y. (2020). Cloud services with big data provide a solution for monitoring and tracking sustainable development goals. *Geography and Sustainability*, 1(1), 25-32.

46. Xu, Z., Chau, S. N., Chen, X., Zhang, J., Li, Y., Dietz, T., ... & Liu, J. (2020). Assessing progress towards sustainable development over space and time. *Nature*, 577(7788), 74-78.
47. Zillner, S., Gomez, J. A., Robles, A. G., Hahn, T., Le Bars, L., Petkovic, M., & Curry, E. (2021). Data Economy 2.0: From Big Data Value to AI Value and a European Data Space. In *The Elements of Big Data Value* (pp. 379-399). Springer, Cham.
48. Tồng, T. P. (2019). Tạp chí Cộng sản. Truy cập https://www.tapchicongsan.org.vn/media-story/-/asset_publisher/V8hhp4dK31Gf/content/quoc-hoi-viet-nam-thuc-ay-viec-thuc-hien-cac-muc-tieu-phat-trien-ben-vung
49. Nguyen, M. Q., Vu, V. C (2021). Tạp chí Quản lý Nhà nước. Truy cập <https://www.quanlynhanuoc.vn/2021/03/18/phat-trien-ha-tang-big-data-du-lieu-lon-o-viet-nam-hien-nay/>
50. Vũ, C. M. Đ (2020). Cục Tin học hoá. Truy cập <https://aita.gov.vn/big-data-va-phat-trien-chinh-phu-so>
51. Anh Cao (2021). Bộ Nội vụ. Truy cập <https://www.moha.gov.vn/kstthc/tintuc/ke-hoach-ung-dung-cong-nghe-thong-tin-phat-trien-chinh-phu-so-va-bao-dam-an-toan-thong-tin-mang-nam-2021-cua-bo-noi-vu-46095.html>
52. Huyền Thanh (2016). Tạp chí Công an nhân dân. Truy cập <http://cand.com.vn/Su-kien-Binh-luan-thoi-su/Chinh-phu-dien-tu-kho-hinh-thanh-neu-thieu-su-lien-thong-387571/>
53. Gia Minh (2021). Tạp chí An toàn thông tin thuộc Ban Cơ yếu Chính phủ. <http://www.antoanongtin.vn/chinh-sach---chien-luoc/ban-co-yeu-chinh-phu-dam-bao-bao-mat-xac-thuc-he-thong-co-so-du-lieu-quoc-gia-ve-dan-cu-va-can-cuoc--106840>
54. GSO (2020). Cổng dữ liệu Quốc gia thuộc Bộ Thông tin và Truyền thông. Truy cập https://data.gov.vn/web/guest/news/-/asset_publisher/FRkblAs8yr3H/content/udcnttdieutrathongke
55. Trí Dũng (2016). Tạp chí Thời báo tài chính Việt Nam. <http://thoibaotaichinhvietnam.vn/pages/thoi-su/2015-03-18/nam-2016-se-su-dung-du-lieu-dien-thoai-di-dong-giam-sat-giao-thong-18916.aspx>

56. Thanh Thuý (2019). Báo Giao thông thuộc Bộ Giao thông vận tải. <https://www.baogiaothong.vn/ung-dung-co-so-du-lieu-dien-thoai-di-dong-trong-quy-hoach-gtvt-d407931.html>
57. Lan Phuong (2021). Báo điện tử của Chính phủ nước Cộng hoà Xã hội Chủ nghĩa Việt Nam. <http://baochinhphu.vn/Kinh-te/Chien-luoc-phat-trien-Thong-ke-Viet-Nam-giai-doan-2021-2030/432037.vgp>

PHỤ LỤC: 17 mục tiêu phát triển bền vững (Sustainable Development Goals)

- Mục tiêu 1 (SDG1). Chấm dứt mọi hình thức nghèo ở mọi nơi.
- Mục tiêu 2 (SDG2). Xóa đói, bảo đảm an ninh lương thực, cải thiện dinh dưỡng và thúc đẩy phát triển nông nghiệp bền vững.
- Mục tiêu 3 (SDG3). Bảo đảm cuộc sống khỏe mạnh và tăng cường phúc lợi cho mọi người ở mọi lứa tuổi.
- Mục tiêu 4 (SDG4). Đảm bảo nền giáo dục có chất lượng, công bằng, toàn diện và thúc đẩy các cơ hội học tập suốt đời cho tất cả mọi người.
- Mục tiêu 5 (SDG5). Đạt được bình đẳng giới; tăng quyền và tạo cơ hội cho phụ nữ và trẻ em gái.
- Mục tiêu 6 (SDG6). Đảm bảo đầy đủ và quản lý bền vững tài nguyên nước và hệ thống vệ sinh cho tất cả mọi người.
- Mục tiêu 7 (SDG7). Đảm bảo khả năng tiếp cận nguồn năng lượng bền vững, đáng tin cậy và có khả năng chi trả cho tất cả mọi người.
- Mục tiêu 8 (SDG8). Đảm bảo tăng trưởng kinh tế bền vững, toàn diện, liên tục; tạo việc làm đầy đủ, năng suất và việc làm tốt cho tất cả mọi người.
- Mục tiêu 9 (SDG9). Xây dựng cơ sở hạ tầng có khả năng chống chịu cao, thúc đẩy công nghiệp hóa bao trùm và bền vững, tăng cường đổi mới.
- Mục tiêu 10 (SDG10). Giảm bất bình đẳng trong xã hội.
- Mục tiêu 11 (SDG11). Phát triển đô thị, nông thôn bền vững, có khả năng chống chịu; đảm bảo môi trường sống và làm việc an toàn; phân bổ hợp lý dân cư và lao động theo vùng.
- Mục tiêu 12 (SDG12). Đảm bảo sản xuất và tiêu dùng bền vững.

- Mục tiêu 13 (SDG13). Ứng phó kịp thời, hiệu quả với biến đổi khí hậu và thiên tai.
- Mục tiêu 14 (SDG14). Bảo tồn và sử dụng bền vững đại dương, biển và nguồn lợi biển để phát triển bền vững.
- Mục tiêu 15 (SDG15). Bảo vệ và phát triển rừng bền vững, bảo tồn đa dạng sinh học, phát triển dịch vụ hệ sinh thái, chống sa mạc hóa, ngăn chặn suy thoái và phục hồi tài nguyên đất.
- Mục tiêu 16 (SDG16). Thúc đẩy xã hội hòa bình, dân chủ, công bằng, bình đẳng, văn minh vì sự phát triển bền vững, tạo khả năng tiếp cận công lý cho tất cả mọi người; xây dựng các thể chế hiệu quả, có trách nhiệm giải trình và có sự tham gia ở các cấp.
- Mục tiêu 17 (SDG17). Tăng cường phương thức thực hiện và thúc đẩy đối tác toàn cầu vì sự phát triển bền vững.