Internet Protocols and Digital Inequalities in Global Communications[1]

Open Access Teaching Case Developed for the Tech for Humanity Pathways Minor

Funded by the Andrew Mellon Foundation

Developed by Dr. Padma Chirumamilla and Dr. Fernanda Rosa

Background

On October 4, 2021, Facebook and affiliated services WhatsApp and Instagram disappeared from the Internet. For the span of over five hours, these services were completely unreachable, whether users tried a desktop computer or a mobile phone. Even Facebook's service pages—where one might normally go to find out about a disruption—were unavailable.[2] The cause of the massive outage was later determined to be a configuration error related to the Border Gateway Protocol (BGP).

BGP has been characterized, variously, as a "postal system,"[3] the "dark magic of the Internet",[4] and the "spit and baling wire"[5] holding the internet together. In this case study we will examine the protocol's social implications in global communication. Every time you use the internet to send messages or photos, to open websites or access services, your request is divided into multiple data packets that can traverse many networks before arriving at their final destination.

---

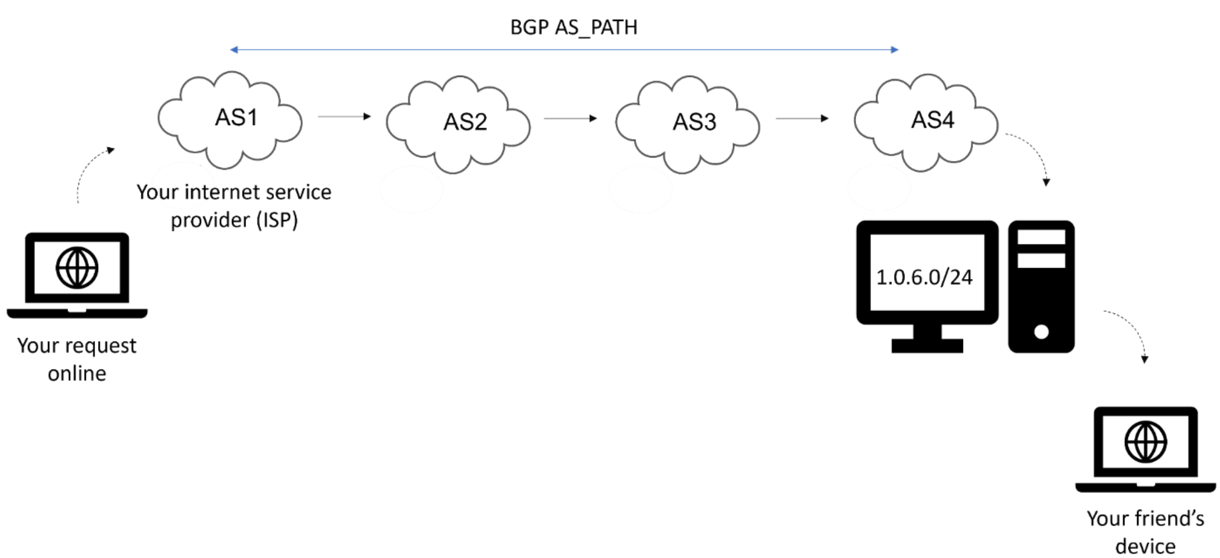
[1] By Fernanda R. Rosa, Assistant Professor, Department of Science, Technology, and Society, Virginia Tech, and Padma Chirumamilla, Ph.D., Consultant.
Case based on the paper Fernanda R. Rosa (2022). "Code ethnography and the materiality of power in internet interconnection infrastructure," *Qualitative Sociology*.
[2] Wodinsky, Shoshana (2021). "Facebook, WhatsApp and Instagram Are All Down [Update: It's Coming Back, Sorry]", *Gizmodo*, October 4. https://gizmodo.com/facebook-instagram-and-whatsapp-are-all-down-1847793555.
[3] Cloudflare, "What is BGP? | BGP routing explained", https://www.cloudflare.com/learning/security/glossary/what-is-bgp/.
[4] Ropek, Lucas (2021). "'The Dark Magic of the Internet' That Brought Facebook To Its Knees," *Gizmodo*, October 5. https://gizmodo.com/the-dark-magic-of-the-internet-that-brought-facebook-to-1847796794.
[5] KrebsOnSecurity. (2021). "The Internet is Held Together with Spit & Baling Wire," November 26. https://krebsonsecurity.com/2021/11/the-internet-is-held-together-with-spit-baling-wire/.

BGP is the internet protocol that helps the networks communicate with each other, ultimately connecting them.

The networks that your data packets traverse are known as autonomous systems (AS). These systems have unique identifiers known as autonomous system numbers (e.g. AS1, AS2, etc.). Using these identifiers, BGP organizes multiple paths, or AS_PATHs. The AS_PATH is a specific route of networks that your data can take to get to its destination. Figure 1 illustrates this process. In this simplified scheme, imagine that when you send a message to a friend, your data packets follow the AS_PATH AS1 - AS2 - AS3 - AS4, before arriving at their destination: the internet address—or *prefix*—1.0.6.0/24, under which your friend's device can be reached.

Similar to how a postal address is linked to a physical building, prefixes also have a physical presence. Network servers and computers have a physical location—whether in a remote data center or in corporate headquarters—and can potentially be located all around the world. Your mobile phone, your computer, all your internet-connected devices—every one of these is connected to prefixes allocated by your internet service provider (ISP).

Figure 1. An Illustration of a BGP AS_PATH



The operations shown above occur in milliseconds. They are the basis of how data is transmitted through the internet, and they occur thanks to an internet protocol known as BGP.

How does BGP work? Each autonomous system has routers that engineers configure based on BGP. BGP, as an internet protocol, works as a language that network routers use to advertise the network addresses that they are able to reach, as well as the path that they use to get to those addresses. These advertisements constantly feed into and update the BGP routing tables. For instance, if network AS4 receives data packets from network AS1 through networks AS2 and AS3, AS4 now knows that it has to reach AS1, it can follow the path AS3 – AS2 – AS1. Nevertheless, other paths may be available, and AS4 may choose them instead, based on commercial interests.

Importantly, decisions about where to send a data packet are distributed. This means that each autonomous system will decide where to send your data packets next, based on the most updated AS_PATH route available. The internet routing system is very dynamic. Once the routing table is updated with a new path, other networks will know that they can send data packets down that path. At the same time, if an address is *not* advertised for any reason (because of a misconfiguration, for instance) it cannot be reached anymore.

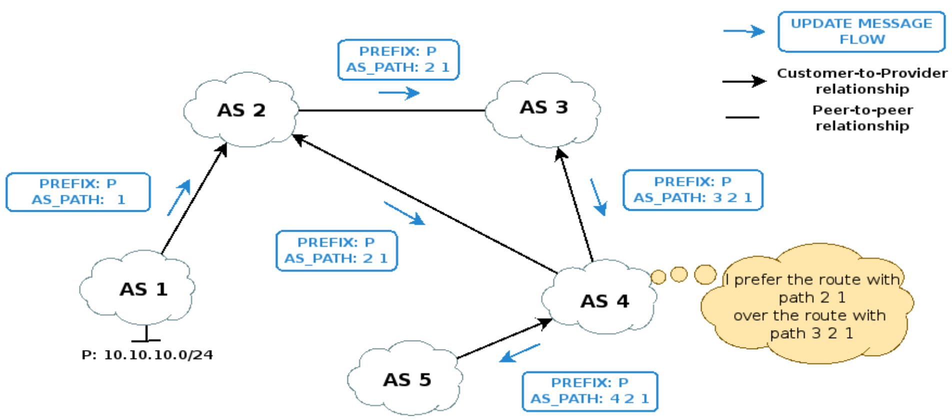
Network engineers commonly program a network's routers to send data packets in the cheapest and quickest way possible. However, what the "cheapest" and "quickest" route is can depend upon commercial agreements, rather than physical closeness.

Operators of autonomous systems can have two types of commercial agreements: *peering*, or a peer-to-peer relationship, and *transit*, a customer-to-provider relationship. When peering with each other, autonomous systems interconnect their networks directly and collaborate by sharing their addresses and the addresses of their customers, usually (but not always) at no cost. In this case, it is said that they exchange traffic with each other. When such traffic exchange is not possible, because, for instance, the prefix an operator wants to reach is not reachable through existing peers, an AS can pay another AS to deliver their traffic to its final destination. This is called a *transit* agreement, and is a customer-provider relationship.

Generally, BGP is configured to prioritize AS_PATHs that will have fewer autonomous systems to traverse, as these are quicker. AS_PATHs that incorporate peering tend also to be prioritized, as they are cheaper.

Look at the figure below. Why would AS 4 prefer the AS_PATH indicated in the yellow balloon?

Figure 2. An Illustration of Two BGP AS_PATHs[6]



Now that you know how networks come together to form the internet as a "network of networks," we can ask: how do internet protocols like BGP shape global communication for people in different countries?

If the internet is ultimately a physical entity, and all internet addresses point to devices physically located somewhere, how does the BGP AS_PATH—and its route prioritizations and preferences—affect the flow of internet traffic across countries in the global South and the global North? In the following case study, we will look at how BGP has worked in practice amid global digital inequalities in two different countries: Brazil, in South America, and Germany, in Europe.

Case Study

Let's imagine that you are in the United States and you decide to update a music streaming app on your phone. Your phone has to connect to a network router in order to exchange traffic. It first sends a request for an update, and then receives the update from the developer's servers. As explained earlier, the network routers between your phone and the developer's servers determine the route your request will take to reach its destination.

---

[6] Reproduced from: Improta, Alessandro and Sani, Luca, "How BGP Routing Really Works," Catchpoint Blog. https://www.catchpoint.com/blog/bgp-routing.

To update the app on your phone, data is sent through routers from one network to the other until it reaches the specific network with the app developer's servers in it. Internet users in Brazil who update their iPhone, for example, might have their data routed through Germany—a country over 5,600 miles away—in order for their internet service provider to obtain the cheapest connection to Apple's servers. Internet users in Germany, in contrast, are unlikely to have their data routed through Brazil to update their iPhone. Their ISP would only need to access a local server to reach Apple, in this case.

How is this possible? The answer involves the physical location of the content provider's servers, ISPs' commercial agreements, and BGP AS_PATH.

To optimize their costs, internet service providers will always seek to do *peering* among themselves locally. To facilitate peering, an ISP can use an internet exchange point, or IXP. IXPs are physical locations hosted in data centers where networks can easily peer, interconnecting directly to each other to exchange traffic. IXPs help to shorten the paths between autonomous systems connected to them. For example, while in Figure 1, AS1 and AS4 are connected through the AS_PATH AS1 – AS2 – AS3 – AS4, in Figure 3, they are directly interconnected through the AS_PATH AS1 – AS4.

Brazil has the largest public IXP ecosystem in the world.[7] IX.br is a project under a not-for-profit organization that has built and administered more than 30 IXPs in the country. It is charged by a multi-stakeholder body to encourage greater local exchange and control over data originating in Brazil. IX.br serves a large number of autonomous systems, most of which connect users and companies that are local to Brazil.

The largest IXP in the global South is administered by IX.br. It is in the city of Sao Paulo, Brazil's most populous city. Several big tech companies—including Amazon, Facebook and Google—are participants in IX.br's Sao Paulo IXP. Many ISPs in Brazil, therefore, pass along their data traffic though the Sao Paulo IXP so they can directly access these companies' servers and services that may not be available in their local states and IXPs. Sao Paulo, therefore, serves as a concentration of data flow within Brazil. This perhaps reflects Sao Paulo's history as

[7] Brito, S. H. B., Santos, M. A. S., Fontes, R. dos R., Perez, D. A. L., Silva, H. D. L. da, & Rothenberg, C. R. E. (2016). An Analysis of the Largest National Ecosystem of Public Internet eXchange Points: The Case of Brazil. *Journal of Communication and Information Systems*, 31(1). https://doi.org/10.14209/jcis.2016.23.

Brazil's most important commercial center, but also the commercial decisions of big tech companies about where to place their servers.

Figure 3. An IXP with Four Participants

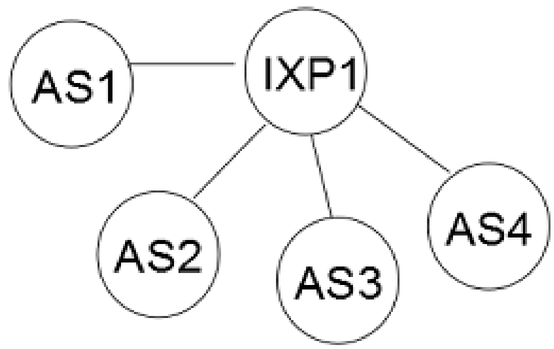
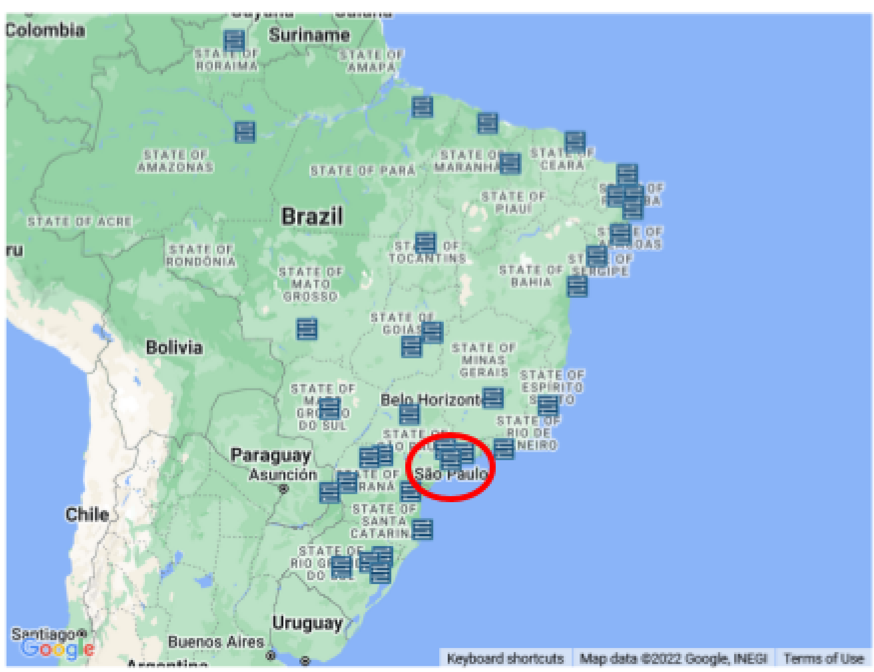


Figure 4. IX.br Map[8]



[8] Source: IX.br. https://ix.br/localidades/atuais

However, what if the content that a Brazilian user wanted—in this case, an update for their Apple iPhone—is not available inside Brazil? Although IXPs were built to facilitate peering, and content providers can bring their servers there, it is up to the companies themselves to connect to an IXP. 93% of Apple's public points of interconnection, where internet service providers can do peering with the company's servers, are based in countries located in the global North. They do not have a presence in any Brazilian IXP.[9]

Commercial preferences are built into the routes that your (or a Brazilian user's) data takes to get to its destination. While these are decisions executed at the level of code and routers, they have important social and practical consequences. If an internet service provider cannot shorten their AS_PATH with a content provider in a local IXP because the company is not present there, they can pay for a *transit* agreement through an intermediary network, or they can look for an IXP located elsewhere where that company is a participant. This is where the largest IXP in the global North —DE-CIX in Frankfurt, Germany—comes into our story.

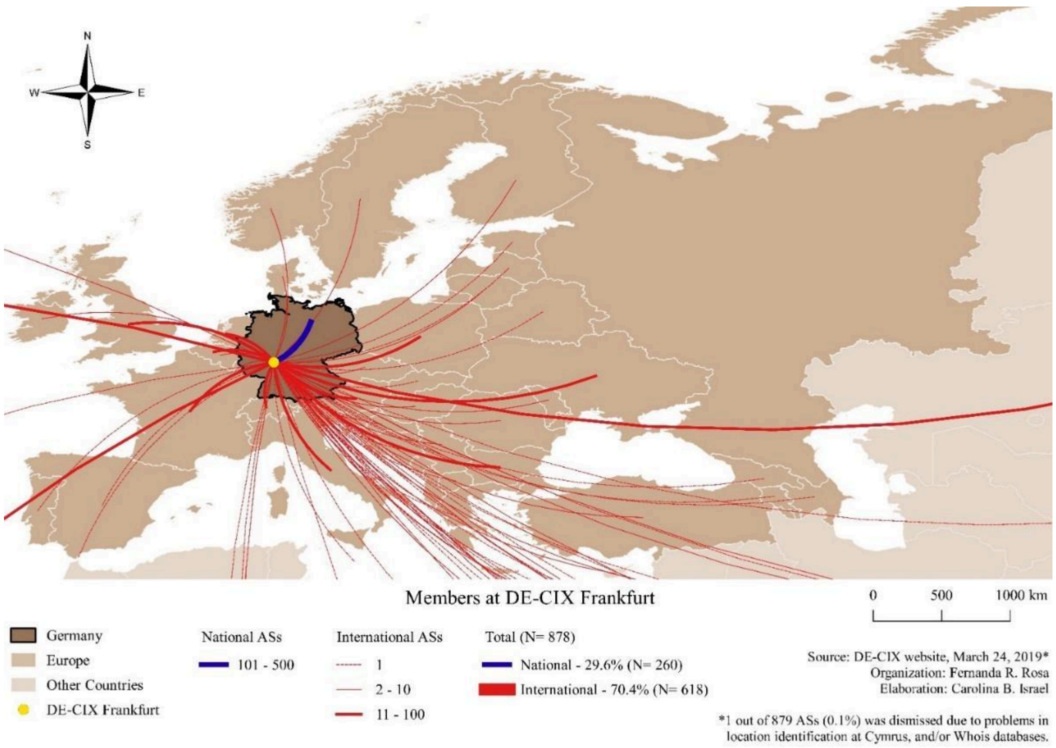
DE-CIX is a for-profit corporation. 70.4% of its participants come from *outside* Germany, as can be seen in Figure 5. This means that internet service providers from other countries send data through here, as illustrated by our example of a Brazilian ISP. In contrast, only 4% of the participants in IX.br Sao Paulo come from abroad. DE-CIX Frankfurt has become a central internet node on the internet, where the world's largest internet companies—like Apple, Amazon, Facebook/Meta, or Google—choose to be participants and are available to do peering. It may therefore be a reasonable market decision for an ISP in Brazil to be a DE-CIX customer, even if that means passing its users' data through Germany. By becoming a customer of DE-CIX, the Brazilian ISP can peer to Apple and other content providers all at once. This might be preferable, in the ISP's view, to passing data through transit agreements for reasons already explained related to costs and the interest in shortening the BGP AS_PATH.

Why does this matter? From an individual perspective, users will not know or notice where in the world their online requests are being fulfilled. But for a variety of reasons, users from certain countries are more likely to have their data cross national borders. This means that users' data will be subjected to laws in countries other than their own. From a global communications

[9] Rosa, Fernanda R. and Hauge, Janice A. GAFA's information infrastructure distribution: Interconnection dynamics in the global North versus global South. *Policy & Internet,* 14(2) 424-449. https://policyreview.info/articles/analysis/internet-interconnection-infrastructure-lessons-global-south

standpoint, this means that more and more of the world's data is concentrated in and passes through networks and IXPs in the global North.

Figure 5. Level of International Appeal of DE-CIX



More ISPs send their traffic *from* the Global South *to* the Global North than the other way around. Examining the flows of traffic between IXPs reveals an infrastructural interdependency between IXPs in the Global North connecting many major autonomous systems together—as in DE-CIX Frankfurt—and internet service providers in the Global South who need access to those particular autonomous systems. The fact that most of the largest content providers are primarily owned and operated by corporations based in the Global North is one reason for this trend.

The fact that IXPs and content providers in the global North attract customers from thousands of miles away illustrates their power in the global communication market. At the same time, the unequal investments that internet service providers in the global South have to make in order to have access to popular content, and BGP AS_PATH's preferences to route data from the global South to the global North, also illustrates the inherently uneven concentration of this market power. This concentration of internet routes, content, and giant internet nodes interconnection facility challenges the notion of the internet as a simple "network of networks", driven only by the free flow of information. BGP reveals that not all networks are equal, and not all data circulate in the same way. Digital inequalities between the global North and the global South stand out at the level of internet infrastructure itself.

Focus Questions

1) Using your own words, explain what is the Border Gateway Protocol (BGP)?

2) What is an autonomous system (AS)?

3) What is a prefix?

4) What are the main drivers that ASes use to select the best route for data on the internet?

5) Why is an internet exchange point (IXP) important for the internet service providers' businesses?

6) What are the main differences between the IXPs in Frankfurt and in Sao Paulo? What are the implications of that?

7) Why is there a greater chance that data from internet users from the global South traverse the global North than the other way around? What does the prioritization of BGP AS_PATH have to do with that?

Thematic Reflection and Discussion Questions

Theme 1: Data Inequalities

According to the United Nations Conference on Trade and Development (UNCTAD), "new products and services generated from data in turn generate even more data, which thereby further accentuates the market power of the digital giants."[10] UNCTAD reports that most digital firms that benefit from data gathering and production are concentrated in the United States and

[10] United Nations Conference on Trade and Development, "Digital Economy Report 2021: Cross-border data flows and development: *For whom the data flow*," p.46. https://unctad.org/system/files/official-document/der2021_en.pdf

China, with only a few comparable platforms in other regions, such as Latin America's Mercado Libre and Africa's Jumia.

As cross-border data flows increase in the global economy, UNCTAD reports that "those that can extract or collect the data—and have the capacity to further process them, mainly global digital corporations from the United States and China—are in a privileged position to appropriate most of the value of the data. By contrast, those who can be considered producers or sources of the data in raw form—i.e., the users of the platforms, with a large number of them in developing countries, who are also contributing to that value—do not receive development gains."[11]

Discussion Questions

1) How does the unbalanced flow of data illustrated in the case study contribute to the inequalities highlighted by the UNCTAD?

2) What would a more equitable environment for data look like? How could we develop and promote such an environment?

3) How can governments be involved in creating this equitable environment for data? How about the public?

Theme 2: Data Concentration

The Center for Applied Internet Data Analysis (CAIDA), located at the University of California's San Diego Supercomputer Center, publishes a ranking of autonomous systems (ASes) from across the world. The ranking is publicly available at https://asrank.caida.org/. One of the rough metrics that the Center uses in determining its ranking is the number of addresses that can be reached from that AS. In short, the larger the reach of the AS, the higher its rank. For the following discussion questions, go to the website and explore the current ranking of the world's ASes.

Discussion Questions

1) Where are the top 20 ASes in the world primarily operated?

[11] United Nations Conference on Trade and Development, *Digital Economy Report 2021*, p.47.

2) How far down the list must you go before you can find an AS operated by a company based in South America? In the Middle East and Africa?

3) What problems might arise from having so many large ASes operated by organizations located in only one or two regions of the world?

Theme 3: Physical internet

While the internet and connectivity is commonly imagined to be seamless and free-flowing, it is actually highly reliant on physical devices and locations, like data centers, cables, servers, routers, etc. In January 2022,[12] Hunga Tonga-Hunga Haʻapai—an underwater volcano off the coast of the island nation of Tonga—erupted, sending massive plumes of ash into the air and creating tsunamis that affected nations across the globe. The eruption destroyed the single underwater cable connecting Tonga to the internet, completely disconnecting the country's people. Internet would not be restored to the main island until over a month later.[13]

Other physical aspects of the Internet—such as the data centers where large corporations host their content—are also increasingly affected by weather and "offline" events. According to the U.S. Department of Energy, data centers consume 10 to 50 times the energy per floor space of a typical office building, and their projected power consumption is expected to increase as the number of data centers across the country increases.[14] Outages due to extreme weather have also been reported at data centers in the United Kingdom.[15]

Discussion Questions

1) What are some "offline" problems that internet infrastructure might encounter in the future? Which direction should we take to address those problems?

---

[12] Stokel-Walker, Chris (2022). "Tonga's volcano blast cut it off from the world. Here's what it will take to get it reconnected," *MIT Technology Review*, (January 18). https://www.technologyreview.com/2022/01/18/1043790/tongas-volcano-internet-reconnected/.

[13] Associated Press (2022). "Tonga's internet is restored 5 weeks after big volcanic eruption," *NPR*. (February 22) https://www.npr.org/2022/02/22/1082483555/tongas-internet-restored-5-weeks-after-big-eruption.

[14] Office of Energy Efficiency and Renewable Energy, "Data Centers and Servers," https://www.energy.gov/eere/buildings/data-centers-and-servers.

[15] Moore, Mike (2022). "Oracle and Google data centers taken down by UK heatwave," (July 20) *Techrader.* Available at: https://www.techradar.com/news/oracle-and-google-data-centres-taken-down-by-uk-heatwave.

2) How do you think that the consistent digital inequalities between the global South and the global North will intersect with these emerging environmental issues? What possible outcomes can we imagine?