Data-Centric Solutions to the Plastic Waste Problem

How data can be part of the solution beyond describing the status quo

By Julia Nikulski, Wuppertal Institute (julia.nikulski@wupperinst.org)

Overview of data-centric solutions to the plastic waste problem. Source: Julia Nikulski with icons from ultimatearm, Freepik, photo3idea_studio, mynamepong, and Smashicons from Flaticon.

The debate around plastic waste on a macro-level is characterized by a lack of data availability: production amounts, trade volumes, waste generation, consumption patterns, and end-of-life treatments are insufficiently documented (see Breaking the Plastic Wave, 2020). While individual companies dealing with plastics production, consumption, or treatment capture data related to their operations, this data is usually not shared or aggregated with that of other companies to provide a comprehensive picture of the plastic flows. If aggregated data is publicly available, then usually only in annual frequency without detailed distinction among different plastic types. Some of this data is collected and reported with time delays of multiple years (see Eurostat).

This shows the gap that exists within the applications of plastics data. On one side of the spectrum, company-level data is used for operational optimizations in individual companies. On the other side, globally aggregated data is leveraged to describe the amount of current and past plastic waste generation and different scenarios for future developments (e.g. Geyer et al., 2017). What is missing in between is data that can be used to achieve desired future outcomes. However, the right data applications can help change behaviors and create a positive impact.

The broad field of data science uses data to extract knowledge – not just about the present, but particularly about the future. Using machine learning, previously unknown patterns and trends can be identified. In the following paragraphs, I will describe three opportunities for how data

can be part of the solution to the plastic waste problem and what needs to happen to implement these ideas.

Predicting plastics availability for recycling

One major barrier to using recycled plastics in products or packaging is the uncertainty around the available quantity of such plastics (see Milios et al., 2018). The amount of plastics available for recycling depends on temporal, spatial, and other factors. Machine learning can be leveraged to predict how much plastics would be available next week, next month, or during any other period. However, the generated information is only valuable if plastic types are differentiated. Therefore, different prediction models for different plastic types (PP, PET, HDPE, etc.) are necessary. If producers could guarantee the supply of recycled plastics because they know how much will be available in the market, consumers could be more willing to adjust their packaging and products to incorporate those plastics.

To generate these quantity predictions, data on the amount and types of plastic waste disposed of needs to be collected. Ideal providers of such data are waste collection facilities and sorting facilities. The former capture how much plastic waste is generated in total in a certain area. The latter can provide details on the amount of plastic waste per plastic type.

Spatiotemporal patterns in the plastic waste disposal

Waste prevention measures are only effective if they target the underlying cause of waste generation. Identifying those causes is not trivial. They depend on time and location. Are there any differences in waste disposal patterns between urban and rural areas? Does seasonality affect the amount and types of waste disposed of? Self-reported data on waste prevention or consumption behaviors can be biased due to the social desirability bias. Studies measuring the waste generated by selected households are not necessarily representative of an entire city and forget about the non-household related waste. Moreover, surveys as a data collection method are usually only conducted over a specified period. But spatiotemporal patterns need to be analyzed throughout multiple periods. Therefore, alternative sources of data are necessary for meaningful pattern analysis.

The granularity of spatiotemporal patterns depends on data availability. Considering the current availability of public data, no or only limited pattern analysis would be possible. However, data from waste collection facilities could be used to identify trends on a community level. Additional granularity can be provided through images of plastic waste littering captured by citizens or drones in different neighborhoods of a city. Both options can provide useful insights and inform decisions on waste prevention measures.

Image recognition to identify ocean plastic hot spots

Different approaches to identifying plastic hot spots in the ocean have already been tested in various studies (e.g. Biermann et al., 2020). Satellites and drones can be used to collect images of plastic debris in oceans, rivers, and coastal areas. With the help of image recognition models areas with an increased need for intervention can be identified (see Martin et al., 2018). Clean-ups can be targeted at particularly polluted areas. Measures to prevent plastics from entering rivers and oceans can be introduced in hot spot areas.

Continuous monitoring through satellites and drones can provide data to derive patterns of waste flows into oceans and to predict future hot spots. A necessary condition for this type of analysis is the availability of satellite and drone imagery.

Final thoughts

Useful solutions with a data-centric approach do not necessarily need to be based on artificial intelligence and machine learning. Creating a database with information on alternatives for plastics packaging for different products can already be a useful first step to support plastic prevention. A marketplace that connects producers with consumers of plastics and other material solutions can facilitate access to alternatives and increase the transparency of the packaging material market. Recommendation engines – similar to how Amazon recommendations work – could be created once enough data on users and transactions in this marketplace is collected to suggest packaging solutions matched to the needs of the users.

The advantage of data-centric solutions is simultaneously their disadvantage. Data is valuable, both for the producers and consumers of data. One reason for the lack of data in the plastics debate is the technical or organizational inability to collect and process all relevant information. Equally contributing to this problem, however, is that already available data can be proprietary, and owners may be reluctant to share this information with governmental actors, other for-profit organizations, or the public. Incentives and regulations around the usage of the data are necessary to make this much-needed data available.