The Joro Carbonizer: A Methodology for Calculating Personal Carbon Footprints

Sanchali Pal
sanchali@joro.tech

Sophie Janaskie
sophie@joro.tech

Sonja Johnson-Yu
sonja@joro.tech

Acknowledgements

Special thanks to members of our Expert Advisory Board, including Reed Miller (Post-Doctoral Associate at Yale Center for Environmental Law Policy), Edgar Hertwich (International Chair in Industrial Ecology at NTNU, Professor at the Department of Energy and Process Engineering), and Na'im Merchant (Climate Change and Development Consultant), for their review and comments. Thank you to Joro team members Gabrielle Gustilo, Nick Reavill, and TJ Considine, as well as external reviewer Paul Wolfram (Postdoctoral Fellow at the Joint Global Change Research Institute, University of Maryland), for their review and comments.

Get in Touch

We are always looking for ways to improve our methodology. Have suggestions or feedback for us? Let us know! Please reach out to us at info@joro.tech.

1. Preface

This white paper describes Joro's methodology for estimating the carbon intensity of financial purchases. The purpose of this paper is to provide transparency into Joro's carbon footprint estimation approach, help advance an international standard for personal carbon footprinting, and inspire ideas and feedback on potential future improvements. This white paper is not intended as a complete catalogue of the market, technological, legal or other risks that Joro may face, or an offer of, or solicitation for, investment in Joro.

1.1. Joro: Our Theory of Change

Joro's initial product is a core set of algorithms (the "Carbonizer") that translate consumer financial transaction data into carbon footprint data. Our long term goal is to empower people to accelerate the transition to a decarbonized economy by 2050.

1.1.1 Generate carbon scores

The first step to decarbonizing consumption is to obtain the data required to inform decisions about where to prioritize carbon reduction and removal. Joro's Carbonizer algorithms automatically convert financial transaction data to carbon footprint data at high enough resolution to provide people real-time feedback on their spending choices. This is the first component of the Joro platform.

1.1.2 Drive rapid emissions reduction

With real-time data on where consumption-based emissions come from, Joro can help people shift to lower-carbon alternatives as quickly as possible. The data and models we develop to inform these recommendations can also be used to inform companies and governments of high-priority opportunities to decarbonize their products and services, facilitating an economy-wide transition to a decarbonized world.

1.1.3 Create a market for net zero

By providing high resolution feedback to people on where the emissions they influence come from, Joro can help them immediately compensate for emissions they cannot yet reduce by supporting promising projects that protect natural carbon sinks and support Carbon Dioxide Removal (CDR) approaches. Joro evaluates, curates, and monitors a basket of such offsets, aligned with achieving a decarbonized society by 2050, and makes them available to anyone with a mobile phone and credit or debit card.

1.1.4 Achieve a just and sustainable economy

Joro recognizes that achieving a decarbonized economy must go hand in hand with measures to promote accessibility, inclusion, and equity. In suggesting actions to drive emissions reduction and removal, we consider and prioritize how those measures will affect local communities. All humans should have the opportunity to live and prosper; biodiversity should flourish; we should be able to use resources at a rate at which the earth can regenerate them. At Joro, that's what "sustainable" means to us.

2. Background

2.1. Why track carbon footprints?

According to the International Panel on Climate Change (IPCC) 2021 report, "achieving global net zero CO_2 emissions is a requirement for stabilizing CO_2-induced global surface temperature increase." [4]

We must slash emissions in half by 2030 and achieve Net Zero emissions by mid-century to avoid the most disastrous impacts of climate change. Achieving these stated climate goals will require a tremendous effort across multiple fronts - from sweeping policy change to bold corporate action.

Understanding our progress, however, requires understanding how much greenhouse gas (GHG) currently exists in, and is being emitted into, the atmosphere, and from which sources. This is where carbon footprinting comes in. A carbon footprint is a measure of the total amount of GHG emissions that are generated by an entity, whether an individual, corporation, city, or other group, over a specified period of time. These emissions may include methane, carbon dioxide, nitrous oxide, and other industrial GHGs. Measurements of these various gases are often converted to and expressed in carbon dioxide equivalents (CO_2e).

Where do consumers fit into this effort? Individuals do not bear the sole responsibility of addressing climate change, but collective action can be a powerful force for change. Every item we buy has a carbon footprint. Consumer choices have been shown to influence over 65% of global carbon emissions [2]. Research has also revealed that an average person can reduce their emissions by approximately 25% through achievable, daily actions [5]. Direct action to reduce emissions can also have indirect consequences, such as contributing to shifting patterns of demand towards more sustainable products and services. These effects are more difficult to accurately measure. Nonetheless, measuring and managing one's personal carbon footprint can help connect individual actions to a broader force pushing for societal change.

2.2. Overview of existing carbon footprint approaches

Over the past couple of decades, carbon footprinting has emerged as an integral tool for measuring and managing the GHG emissions of our activities and achieving global emissions targets. There have emerged two primary approaches for estimating a carbon footprint: (1) a process-based approach, and (2) an input-output model approach.

2.2.1 Bottom-Up Approach (Process-Based)

Process-based models for carbon footprinting are bottom-up approaches to estimate carbon impacts. They typically take the form of Life Cycle Assessments (LCAs) and take into account all of the steps, activities, and inputs required to produce a product or activity. If spanning the entire life cycle of a product, they may also take into account the use phase and end-of-life of the product. LCA is a well-established methodology subject to ISO standards.

This type of approach is particularly useful when trying to determine the footprint of a specific individual product, especially if that product is made using a unique process. For example, if a manufacturer has produced a garment from sustainably grown cotton or by using a more sustainable dying process, then these specific processes and their relative benefits and drawbacks can be reflected in an LCA analysis.

However, LCA analyses face a number of logistical challenges. First, an accurate LCA requires a great deal of detailed data. Collecting this data is costly, both in terms of time and labor. Accessing this data may also be a challenge, as some suppliers prefer to keep the specifics of their production activities proprietary. Second, there is a certain degree of subjectivity involved in LCA. This type of analysis requires the analyst to draw a "system" boundary around the process that they are researching. This system boundary may differ across studies, leading to a lack of comparability across different LCAs of the same process.

For these two reasons, LCA can lack scalability and consistency, making it most appropriate for applications where (1) reliable data is readily available, and (2) a specific process or product is the subject of study.

2.2.2 Top-Down Approach (Input-Output)

Input-output (IO) models are top down approaches that rely on two types of data: (1) macroeconomic data in the form of input-output tables, which record the flow of money between economic sectors, and (2) national environmental datasets to generate top-down GHG impact estimates of activities in various sectors of the economy.

These GHG emissions factors ($tCO_2e/\$$) are estimated from environmentally-extended input-output (EEIO) models, which assign national GHG emissions to industry sectors that directly emit them. The model traces the economic flows between sectors and the associated flows of embodied GHG emissions. GHG emissions are allocated to final products as they accumulate along the supply chain. Therefore, this method connects GHG estimates at the industry level to the purchaser (e.g. the payment of the individual consumer), thus estimating the carbon impacts of a product upstream from the point of purchase, including raw material extraction, supply-chain transport, and manufacturing.

The input-output approach, while lacking specificity on a product-level relative to LCA, is comprehensive across the economy of a given geography. Its internal consistency and simplified data requirements make it highly scalable.

However, a number of challenges also face the input-output approach. First, as the GHG estimates provided by this method are sector averages, nuances within that sector can be lost. For example, while the IO method is effective for estimating the average carbon footprint of money spent at a clothing store, it will not be able to differentiate between different types of clothing (e.g. t-shirts vs. trousers) nor specific processes used to produce clothing (e.g. using conventional cotton vs. organic cotton). For sectors that are highly heterogeneous, this problem is further exacerbated.

Second, input-output models are often constructed at the national, and occasionally subnational, level. Expanding a model across geographic regions into multi-region input-output models requires data harmonization and introduces additional requirements. A challenge with using an input-output model at the national level is that import products are assumed to have the same footprint as products made domestically, which may be an over- or underestimate depending on the source country's relative environmental performance.

Third, a key feature of the input-output approach is the assumption that there exists a linear relationship between monetary and environmental flows.[1] In other words, it assumes that, for every dollar spent, GHG emissions will increase in a linear fashion. This tends to be sufficient at the national scale, and when goods/services within a category have similar prices. However, at the individual customer level, when different customers pay very different prices, it can lead to inconsistencies. As an example, consider the case of a shopper A who purchases an item at full price. Shopper B, meanwhile, uses a half-off coupon for the same product. Ideally, both shoppers would be assigned the same carbon footprint for that product, but the model would assign Shopper B half the carbon footprint of shopper A.

2.2.3 Existing Methodologies & Calculators

A number of standardized carbon footprinting protocols have developed over the past decade. The GHG Protocol, for example, is a nonprofit entity that "provides standards, guidance, tools and training for business and government to measure and manage climate-warming emissions." [8] They have developed perhaps the most widely used GHG accounting standards globally.

While GHG Protocol provides comprehensive GHG accounting standards and frameworks for companies and cities, there does not currently exist a universally accepted method for personal or household carbon footprinting. In the absence of such a standard, a number of nonprofits, academics, and companies have sought to develop robust methodologies for measuring and managing personal carbon footprints. Some of these entities - such as the World Wildlife Fund, UC Berkeley Cool Climate, and Klima - are based on user inputs and questionnaires addressing their food, energy, transport, and purchasing behaviors. Others - such as Ducky and Svalna - have developed methodologies that blend top-down approaches with bottom-up approaches. Finally, other entities such as Aerial are investigating ways to leverage non-financial data sources, such as transportation data, to construct carbon footprints.

This area of the field is still evolving, and this paper aims to provide a framework for personal carbon footprinting, which incorporates a blend of top-down and bottom-up techniques, developed by Joro's team.

2.3. Why use financial data to estimate personal carbon footprints?

Studies have shown that personal carbon calculators have historically suffered from a lack of completeness, an inability to keep users engaged, and an inability to inform real-time decision-making. We find that incorporating data on consumers' spending activity, namely in the form of their credit card[2] transactions can help address some of these challenges.

2.3.1 Benefits

There are several key benefits to using financial data to estimate personal carbon footprints:

First, financial spending data is an automated data feed. A significant barrier to the adoption and repeated use of personal carbon management tools is the need for users to manually input data related to their behaviors, actions, and consumption. Using credit card transaction data allows this tracking to be updated automatically, significantly reducing the amount of manual user input that is required in order to measure and manage a footprint. Furthermore, using an automated data feed reduces the likelihood of human error, making it more reliable. Financial transaction data in particular is robust and reliably provided and recorded by multiple, secure, third party sources, rendering it a valuable and rich source of information.

Financial data provides a holistic view across energy domains. Carbon estimation based on financial transactions demonstrates carbon footprint trade-offs across spending categories. For example, one is able to compare the relative estimated impact of a public transportation ticket purchase compared to a meal at a restaurant. This type of cross-domain comparison aids in the building of a "carbon intuition," or an understanding of the relative impacts

[1] Note that the linear relationship is present in LCA as well, but typically connects mass or number of a product purchased to the impact. See this paper for more detail.

[2] For the purpose of this paper, we will use "credit card" as a blanket term for a payment card product, which includes credit, debit, and prepaid cards.

of spending choices, and informs decision-making.

It is immediate and action-oriented. Credit card purchases are reflective of daily user actions, such as paying for a taxi or buying groceries or home goods. Capturing this type of data paints a granular picture of peoples' activity, allowing Joro to inform immediate potential behavior changes to lower emissions. Financial transaction data offers near-immediate pathways to action. Given the short time frame to achieve international emissions reduction goals, this is a significant benefit.

Spending data is consistent with existing footprinting standards. Over the past two decades, EEIO carbon footprinting methodologies have evolved significantly. Originally they were developed to apply to countries. Since then, they have been adapted for use by corporations and other sub-national actors. Extending this methodology even further to consumers allows us to be consistent with these previously developed accounting approaches.

2.3.2 Limitations

There are several known limitations of this approach that we are aware of and must take into consideration as we advance a new standard in personal carbon footprint accounting. Many of these limitations have already informed our work in developing the Carbonizer, and some present opportunities for future improvement:

Financial transaction data is not all-encompassing. While credit card transaction data does capture a meaningful amount of consumer behavior, especially in the United States and among our target demographic, 25-44 year olds, who use credit, debit, and electronic payments for over 70% of all transactions [3]. However, there are a number of activities and purchases that do not find their way onto a card or bank statement. For example, paying rent by check or making cash purchases in a store are not included in this particular data stream. Thus, some manual input will still be required. We are also aware that this data feed does not include use of free or public services, such as public education. However, as Joro's intention is to help inform sustainable consumption based on resources that people directly influence through their personal spending, we consider these types of services out of scope for the Carbonizer. Instead, to address these areas, Joro aims to help people understand other types of actions they can take to influence emissions that do not directly result from their spending, e.g. via civic engagement, political action, etc.

It lacks granularity at the product level. Credit card transaction data includes several key pieces of information - including vendor and transaction amount. However, most credit card data does not include information on the individual items that are purchased. This lack of granularity presents a challenge when trying to develop highly accurate carbon impact estimates of purchases. Moving forward, Joro intends to help inform consumers on the relative carbon impact of different purchases in a way that will inform decision-making, while working around the lack of product data, for instance by distinguishing sustainable practices by vendor or by product type.

Assigning ownership can be challenging. The footprint of every transaction executed on a credit card is not necessarily attributable to the person who purchased it. For example, an individual could decide to cover the entire bill at a dinner with multiple friends. Here, attributing the full carbon impact of that transaction to the cardholder could be perceived as inaccurate. In another example, it may be the case that a credit card is shared by spouses, and the stream of data it produces represents purchases for an entire multi-person household. As a team, we are taking steps to allow people to better represent and attribute the carbon footprint of various purchases accurately, to the extent that it can inform emissions reduction.

Transactions are not perfectly categorized. Banks and other third-party entities often automatically assign categories to credit card transactions. This categorization is helpful in determining the appropriate carbon multiplier to assign to a given credit card transaction. However, transactions are not always perfectly categorized, or categories assigned may not be the most informative for determining the emissions associated with that transaction, introducing a potential source of error into the process of constructing a carbon footprint. To address this, we are taking steps to improve categorization of transactions as relevant to carbon footprint accounting.

3. Joro's Carbonizer

3.1. Overview

At the highest level, Joro's Carbonizer uses three key sources of data to provide a holistic view of the carbon footprint of a person's consumption.

1. First, the algorithms connect to a user's credit card via a **financial API**, which provides an automated stream of data on user purchases. The current Carbonizer integrates with Plaid.

2. Second, users take a **Carbon Survey** within the mobile app to provide key inputs to the algorithm that cannot be captured by credit card transactions but will meaningfully inform footprint estimates.

3. Third, **external datasets** from academia, government, and other trusted sources are incorporated to increase the accuracy of the Carbonizer's estimates.

The foundation of the Carbonizer is built using a top-down, input-output approach. The algorithm ingests

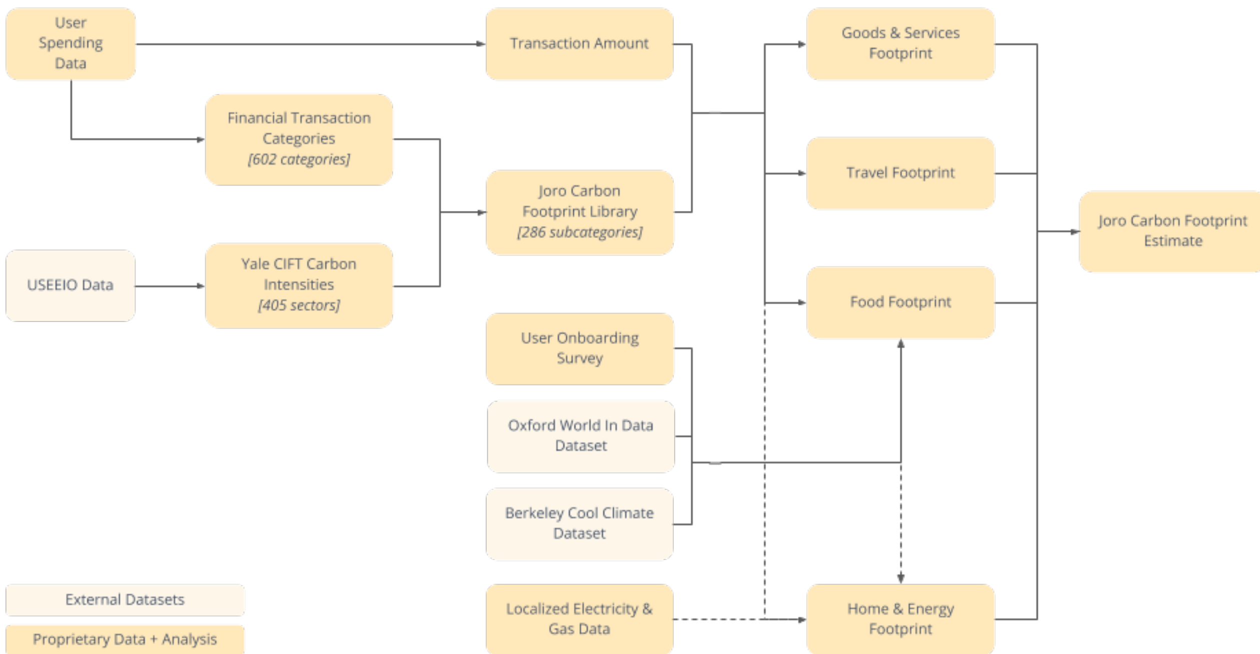

Figure 1: Overview of Carbonizer inputs

users' credit card transactions, sorts them into categories, and uses those categories to determine a kg CO_2e carbon multiplier per dollar for each purchase. These carbon multipliers are calculated based on rigorous academic and publicly available datasets on the carbon intensity of US industrial sectors. The dollar value of each purchase is then multiplied by its assigned carbon multiplier to derive a carbon footprint estimate, per the guiding equation below.

Carbon Footprint of Purchase *(kg CO_2e)*
= ***Transaction Amount*** *($)* × ***Transaction Category Carbon Weight*** *(kg CO_2e/$)*

As discussed in the previous sections, input-output approaches suffer from drawbacks that impede the accuracy and actionability of the carbon estimates they produce. To address some of these challenges, Joro's Carbonizer incorporates additional localized or process-based data retrieved from user input or other trusted sources.

The following sections provide more detail on how these additional calculations are incorporated into each of Joro's **four supercategories**:

1. Goods & Services (Shopping)
2. Food & Drink
3. Transport
4. Home Energy

A high-level summary of the inputs of the algorithm can be seen in Figure 1, and a more in-depth explanation can be found in the Data Sources section.

3.2. Calculation Methodology

3.2.1 Calculating carbon multipliers

Joro receives information on a user's purchases via a financial API provided by Plaid, a financial data platform. In addition to the transaction date, amount, and memo, this data also includes a "category id", which identifies the transaction as belonging to one of several hundred financial categories. A foundational piece of Joro's Carbonizer is generating a carbon multiplier (in kgCO_2e/$) for each of these financial categories.

To generate these carbon multipliers, Joro maps each of these financial categories to one or more industry sectors from the "Capital Inclusive Footprint Tool - United States" (CIFT-US) [1].

The CIFT-US is based on the US Environmentally Extended Input Output (USEEIO) model, which captures monetary flows through 405 US industry sectors and also estimates the total carbon emissions of each sector based on national greenhouse gas accounting [10]. This allows for the calculation of the carbon footprint intensity (kg CO_2e

Mapping Scenario	Financial Category Carbon Multiplier Calculation
1 EEIO sector to 1 Financial Category	Equal to the EEIO sector weight
1 EEIO sector to Many Financial Categories	Equal to the EEIO sector weight
Many EEIO sectors to 1 Financial Category	Calculated as the weighted average of the many mapped EEIO sector weights, weighted by the dollars of consumption in those sectors
No match	Calculation varies, including: 1. Setting carbon weight to zero 2. Flagging for inclusion in the Carbon Survey 3. Seeking alternative sources for the weight

Table 1: Mapping from EEIO sectors to Financial Categories

per dollar) of the goods and services produced by these industry sectors. The CIFT-US builds upon the USEEIO in two important ways: (1) by adjusting them to reflect the price of the good paid by the final purchaser, rather than the (lower) price paid by the upstream producer, and (2) by incorporating the impacts of the capital assets used along the product supply chain (such as equipment, structures, and R&D) typically accounted for separately from EEIO models.[3]

Further information on CIFT-US, the USEEIO model, and the financial API are available in Appendix A.

Mapping EEIO sectors to financial transaction categories: Because the CIFT-US industry sector categories are defined differently than the financial transaction categories from the financial API, they do not map perfectly to one another. Thus, Joro has constructed a custom-built category hierarchy that maps the financial data categories to the appropriate sectors from the CIFT-US. For a given financial transaction category, there exist four possible mapping scenarios, summarized in Table 1. Using this mapping, a carbon multiplier in kg CO_2e per dollar can be generated for each financial transaction category.

Exceptions: Through this mapping exercise, we identified a limited number of low-carbon intensity Plaid category types for which transactions do not consistently adhere to a linear relationship between dollars spent and increased carbon impact. Due to the fact that this is a fundamental assumption of the input-output method and the low footprint of these types of transactions, we assigned a carbon multiplier of zero to these categories. These categories have been flagged for future work to develop more robust ways of accounting for these impacts.

3.2.2 Calculating the base carbon estimates for a transaction

When Joro receives a transaction from a financial API, the Carbonizer first categorizes the transaction and then retrieves the carbon multiplier associated with that category. The transaction purchase amount is multiplied by this carbon multiplier to calculate the base carbon footprint for the purchase, measured in kg CO_2e.

Goods & Services: Whereas certain transaction types that belong to the three other supercategories integrate user and/or external data to adjust the base carbon footprint, at this time Goods & Services transactions simply use the base carbon footprint.

3.2.3 Improving upon base carbon estimates

Food & Drink: Studies have demonstrated that dietary choices can be a significant driver of personal carbon footprints [6][9]. When a user makes a Food & Drink purchase, however, the data received from Plaid is not sufficiently granular to identify what specific food items were purchased. Thus to increase the actionability of our estimates, the Joro Carbonizer combines user inputted data and data on the carbon intensities of various dietary types (e.g. "vegetarian" or "pescetarian").[4]

Modifying Food & Drink carbon footprint estimates using the Carbon Survey: For users who have connected their digital spending accounts, the Carbonizer uses the Carbon Survey to improve the granularity of Food & Drink carbon footprint estimates.

First, the Carbonizer recognizes a Food & Drink purchase based on its financial category id and looks up the

[3]Most EEIO models are built to model overall emissions in a specific year at the national level, and so they look at the footprint of creating capital assets in the year they are created. In CIFT-US, the goal is to account for the impact. For example, if a factory was built 5 years ago and 10,000 kg CO_2 was generated during its construction, a fraction of that construction impact would be assigned to the products it made this year.

[4]See Appendix A for more information on data sources.

corresponding base carbon multiplier.

Second, the Carbonizer considers supplementary information on a user's dietary habits from their Carbon Survey to construct a food multiplier. Food multipliers are calculated based on the user's protein intake, based on data on the carbon intensity of various protein types. For instance, a user who purchases red meat will have a higher food multiplier applied to their transaction-based carbon footprint estimates than a vegan user.

While an average American purchasing from a grocery store might consume a combination of meat, dairy, and produce, a vegan might primarily consume produce. Without the integration of the Carbon Survey, a $50 grocery purchase would have the same carbon intensity, whether purchased by an average American or a vegan. The Carbon Survey allows the Carbonizer to differentiate these cases, offering accuracy to a user's estimates in a way that can inform lower-carbon decision making.

Transport: Studies estimate that 30-40% of the average US citizens' carbon footprint can come from transportation emissions [6][9]. For most categories within the Transport supercategory, Joro uses the standard base carbon footprint calculation.

Automotive fuels such as gasoline and diesel, however, are examples of Transport purchases that have significant geographical variance in price. This variance can be a source of inaccuracy for carbon estimates produced using an input-output method. For example, a gallon of gasoline in Florida may cost two dollars. That same gallon of gasoline in California, however, may cost four dollars. Though the California gallon costs twice that of the Florida gallon, the carbon impact of those gallons remains equivalent. Thus a foundational assumption of the input-output method, namely that carbon impact trends linearly with dollars spent, does not hold. To increase the accuracy of the Joro Carbonizer's transport estimates, we incorporate additional fuel-related datasets to account for localized prices.

Incorporating localization data for gasoline

To address the price variance issue, the Joro Carbonizer adopts a location-sensitive method for calculating the quantity of gasoline a user purchases. First, the Carbonizer identifies a transaction as a gasoline purchase (via the "Gas Stations" Plaid category). Second, the Carbonizer combines datasets on the price of gasoline with zip code information provided by users to identify the average price of gasoline in their area.[5] Finally, Joro combines this information with the user's transaction amount in dollars to calculate the actual quantity in gallons of gasoline purchased.

Incorporating combustion data for gasoline

A significant portion of the carbon impact of a gasoline purchase occurs during the use phase, i.e. during combustion. We incorporate both the upstream carbon impacts (from the CIFT-US dataset), as well as the combustion carbon impact (from the US Environmental Protection Agency) in Joro's carbon estimates.

Home Energy

Energy consumption is another large driver of household greenhouse gas emissions. As such, a wealth of data is available on the consumption, price, and greenhouse gas impacts of electricity and natural gas. In light of both the heightened importance of this sector and the rich datasets available, the Joro Carbonizer incorporates additional energy-related datasets to increase the granularity and accuracy of its Home & Energy carbon estimates.

Incorporating localization data for electricity and natural gas

When it comes to utilities, there are two types of users: those who pay their utility bills on their connected cards, and those who do not.

For users who do pay their utility bill on their connected spending accounts, the Carbonizer modifies the ongoing estimates of their home energy footprint from utility transactions based on their zip code. This estimation uses a localization method similar to the one applied to relevant Transport transactions. The Carbonizer integrates data on the price of electricity and natural gas by state.[6] It then uses user zip code information to identify the average price of electricity ($/kWh) and natural gas ($/thousand cubic feet) in a given users' area. Combining this localized price information with the dollar value of utilities purchases enables the algorithm to more accurately estimate the quantity of energy consumed by a user. Using greenhouse gas intensities for electricity and natural gas provided by the US Energy Information Administration, the algorithm can convert this quantity of energy used to an estimated carbon footprint. Modifying Home footprint estimates using the Carbon Survey For users who do not pay their utility bill on their connected spending accounts, the Carbonizer approximates the user's home energy footprint based on their home size and number of people in their household, as specified by the user.

3.3. Splitting transactions and saving recategorizations

Splitting transactions

Sometimes a user's purchases do not reflect their carbon footprint alone. Users who own shared financial accounts or frequently make purchases on behalf of groups will have the collective carbon emissions reflected in their personal footprint, even if they are only responsible for a fraction of it. To address this issue, Joro allows users to "split" emissions from transactions by designating the number of people sharing the purchase.

Saving recategorizations

[5] See Appendix A for more information on data sources.

[6] See Appendix A for more information on data sources.

If a transaction is incorrectly categorized, Joro users have the opportunity to correct the categorization manually. A user can also specify for Joro to automatically recategorize future transactions from that vendor.

4. Future Work

Joro is constantly looking for ways to improve the rigor, relevance, and actionability of the Carbonizer's carbon footprint estimates. In this spirit, there are a number of ways in which we anticipate continuing to improve our methodology moving forward.

Further improving the granularity and actionability of estimates. We have identified areas in our category mapping that could be made even more specific to better inform action. For example, some purchases are categorized in by financial data as "Convenience Store," which alone, does not provide a clear picture of the types of products purchased. By contextually soliciting more information about more heterogeneous purchases, Joro's Carbonizer can deliver more nuanced insights and recommendations.

Estimate carbon emissions according to the vendor. While the Carbonizer's estimates are helpful for developing a carbon intuition, users want more feedback on how to make sustainable lifestyle decisions. Specifically, users want to know which vendors are more sustainable and be rewarded for buying from them. This requires developing carbon multiplier estimates for specific vendors by assessing vendor-provided information and drawing on third-party evaluations. As a result, a user's carbon footprint will be more influenced by their purchasing behavior. This will aid users in making more environmentally-conscious decisions and may put pressure on companies to adopt more environmentally-conscious behaviors.

Expansion to global markets. The climate crisis is a shared global challenge. Part of Joro's mission is to help people connect with others across the world to engage in meaningful collective climate action. Our methodology is currently based on US datasets. Moving forward, we aim to integrate datasets from different countries and regions to expand the geographic reach of the algorithm.

Improving transaction categorization. The accuracy of the Carbonizer is limited by Plaid's capacity to correctly categorize transactions. If a transaction is incorrectly categorized, then its carbon estimate will likewise be inaccurate. Joro aims to improve categorization accuracy to improve footprint estimation accuracy.

Revealing the emissions we cannot reduce alone. The current Carbonizer methodology does not capture the footprint of publicly-provided goods and services, such as public education, infrastructure, and government services. These services are part of an "immutable" portion of our carbon footprints, which cannot be changed through personal action alone. It is important to acknowledge that this is a meaningful part of our footprints that we can influence through civic action, activism, and engagement. Looking forward, Joro hopes to make this part of our footprints visible and actionable, too.

Appendices

A. Data Sources

The Carbonizer integrates data from a number of sources to calculate a personalized carbon footprint estimate for each unique user. This data includes automated financial transaction data collected from Plaid, sector-level greenhouse gas emissions data from the USEEIO and CIFT-US, user-inputted data from Joro's Carbon Survey, and other external data sources from academia, government, and other trusted sources.

A.1. Financial Transaction Data

For card-connected Joro users, Joro uses the Plaid API, a secure fintech platform, to access information about users' purchases. Plaid is the same API that Venmo, Robinhood, and most cutting-edge fintech applications use to securely process bank-related information. Joro does not store any sensitive information about accounts, but rather simply reads the names, categories, and amounts of a user's transactions. This is the first input to the Carbonizer.

A.2. Carbon Intensity Data

A.2.1 USEEIO

The Joro Carbonizer algorithm builds upon the United States Environmentally-Extended Input-Output (USEEIO) model, which is a source for greenhouse gas emissions factors of US economic sectors. The USEEIO combines data on economic transactions between industry sectors with environmental data for these sectors to build a life cycle model of US goods and services. The most recent USEEIO model is constructed using data from 2012, as it takes several years for the relevant government agencies to update this data.

While the USEEIO model is a robust source of data, it has two key shortcomings for the purposes of Joro's Carbonizer. First, the USEEIO carbon intensities are presented in kg CO_2e/producer-\$. As the Joro Carbonizer aims to apply the GHG intensities to consumer purchases, it requires carbon intensities in kg CO_2e/purchaser-\$. Second, given the methodology used to create the USEEIO, the environmental cost of long-term capital assets (e.g. machinery, factories, IT, vehicles, roads) used to produce certain goods are not included. This omission results in a systematic underestimation of the true climate impact of EEIO estimates [10][1].

A.2.2 CIFT-US

To address these two issues, the Joro Carbonizer uses carbon intensities sourced from the Capital-Inclusive Footprint Tool for the United States (CIFT-US), developed by researchers at the Yale Center for Industrial Ecology. This tool is built on top of the USEEIO and addresses these two limitations. First, the authors of the CIFT-US construct a matrix that converts "producer's price" into "purchaser's price." Second, the authors develop a capital flow matrix to incorporate capital assets (also called "endogenizing" capital). According to their analysis, the use of capital assets for production 2012 accounted for 13% of the economy-wide carbon footprint, underscoring the need to include them in the input-output approach to greenhouse gas accounting [1]. Thus the CIFT-US is the second input to the Carbonizer.

A.3. User-generated Inputs

The Carbonizer takes into consideration certain user-provided inputs about lifestyle choices to improve the granularity of purchase-based carbon footprint estimates. The mobile app solicits this information as part of the Carbon Survey, a component of the user onboarding flows.

A.3.1 Setting and supplementing baseline carbon emissions estimates

When a person creates a Joro account, the app calculates the user's starting carbon footprint based on the previous 90 days of transactions before account creation, to represent an estimate of the user's emissions before the introduction of Joro. If a user chooses not to connect their digital spending accounts, the app estimates a static footprint based on their answers to the carbon survey. For users who do not complete the carbon survey, the app assumes values for the carbon survey in line with the average American, supplementing any gaps in data with publicly available average values from UC Berkeley's Cool Climate Dataset.

A.4. External Data Sources

A.4.1 Food & Drink Data Sources

In our models, we use Poore & Nemecek's estimates of the carbon intensities of various protein-rich foods [7]. We quantify dietary habits based on four different types of protein types: plant-based proteins (e.g. nuts, tofu, pulses), non-meat animal proteins (e.g. eggs, dairy), white meat proteins (e.g. seafood, fish, chicken, pork), and red meat proteins (e.g. beef, lamb). We assume all dietary types consume similar amounts of grains and other foods that are not protein-rich.

To apply this data, we ask a user to approximate their consumption based on the number of meals of each type of protein they eat in a week. We assume all meals include a foundation of plant-based grains, and any non-specified meals include plant-based proteins.

Energy Source	Data Input	Source
Gasoline	Fuel type	Joro Carbon Survey
	Fuel price, by region and type	EIA Gasoline and Diesel Fuel Update
	Fuel price, by state	AAA
	GHG intensity, combustion of gasoline	US EPA
	GHG intensity, production of gasoline	CIFT-US, “petroleum refineries” sector

Table 2: Data sources for gasoline emissions calculation

Energy Source	Data Input	Source
Electricity	Residential electricity retail price, by state	EIA, Electric Power Monthly
	CO_2 emissions factor of electricity, by state	EIA State Electricity Profiles
Natural Gas	Price of natural gas, by state	EIA Natural Gas
	CO_2 emissions factor of natural gas	EIA
	Heat content of natural gas	EIA

Table 3: Data sources for electricity and natural gas emissions calculation

A.4.2 Transport Data Sources

Additional fuel-related datasets used are summarized in Table 2.

A.4.3 Energy and Utilities Data Sources

Additional datasets used are summarized in Table 3.

B. Glossary

See next page for glossary.

References

[1] P. Berrill, T. Miller, Y. Kondo, and E. Hertwich. Capital in the american carbon, energy, and material footprint. *Journal of Industrial Ecology*, 24(3):589–600, June 2020. 5, 9

[2] D. Ivanova, K. Stadler, K. Steen-Olsen, R. Wood, G. Vita, A. Tukker, and E. Hertwich. Environmental impact assessment of household consumption. *Journal of Industrial Ecology*, 20, 2015. 2

[3] R. Kumar and S. O'Brien. Findings from the diary of consumer payment choice. 2019. 4

[4] V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, N. C. S. Berger, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, O. Y. T. Waterfield, R. Yu, and B. Zhou. Ipcc, 2021: Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. *Cambridge University Press*, 2021. 2

[5] D. Moran, K. Kanemoto, M. Jiborn, R. Wood, J. Többen, and K. C. Seto. Carbon footprints of 13 000 cities. 13(6):064041, jun 2018. 2

[6] D. Moran, R. Wood, E. Hertwich, K. Mattson, J. F. D. Rodriguez, K. Schanes, and J. Barrett. Quantifying the potential for consumer-oriented policy to reduce european and foreign carbon emissions. *Climate Policy*, 20(sup1):S28–S38, 2020. 6, 7

[7] J. Poore and T. Nemecek. Reducing food’s environmental impacts through producers and consumers. *Science*, 360(6392):987–992, 2018. 9

[8] G. G. Protocol. Greenhouse gas protocol about us. 3

[9] K. Song, S. Qu, M. Taiebat, S. Liang, and M. Xu. Scale, distribution and variations of global greenhouse gas emissions driven by u.s. households. *Environment International*, 133:105137, 2019. 6, 7

[10] Y. Yang, W. Ingwersen, T. Hawkins, M. Srocka, and D. Meyer. Useeio: A new and transparent united states environmentally-extended input-output model. *Journal of Cleaner Production*, 158:308–318, Aug. 2017. Publisher Copyright: © 2017. 5, 9

Glossary

Capital-Inclusive Footprint Tool for the United States (CIFT-US)	A tool developed by researchers at Yale to facilitate the estimation of carbon, energy, and material footprints associated with consumption of goods or services, including the impacts of fixed capital assets
Carbon dioxide equivalent (CO_2e)	A unit of greenhouse gas equivalent to the amount of CO_2 that would have the same climate impact over a period of 100 years, as defined by the IPCC
Carbon footprint	The the amount of greenhouse gases emitted from a person's activities during a given period of time
Environmentally Extended Input-Output data (EEIO)	A family of macroeconomic-environmental models that facilitate the top-down estimation of potential environmental impacts associated with the production and/or consumption of goods and services
Greenhouse Gas Emissions (GHG)	Gases that trap heat in the atmosphere, such as carbon dioxide, methane, nitrous oxide, and fluorinated gases
International Panel on Climate Change (IPCC)	A United Nations body of 195 member governments that regularly assesses the scientific basis of climate change, including its drivers, impacts, risks, and options for mitigation and adaptation. IPCC reports are a key input into annual international climate negotiations.
Life Cycle Assessment (LCA)	A methodology used to estimate the potential environmental impacts of a product or service during a defined portion of its lifetime.
Net Zero	The achievement of balance between GHGs emitted into the and GHGs removed from the atmosphere. Achieving "Net Zero" generally implies measuring one's GHG emissions, reducing these emissions as much as possible, and removing only what could not be reduced.
Paris Agreement	A landmark international treaty on climate change signed by 197 nations and adopted in 2015. It set the goal "to limit global warming to well below 2, preferably 1.5 degrees Celsius, compared to pre-industrial levels" by achieving net zero emissions globally by mid-century.
Producer price	In the context of input-output methods, the amount receivable of a producer of a good or service, excluding "wholesale and retail trade margins and transportation costs," but including "sales and excise taxes collected and remitted by producers."
Purchaser price	In the context of input-output methods, the price paid "by intermediate and final purchasers for the goods and services that they buy. These prices are equal to producers' prices plus domestic transportation costs and trade margins."