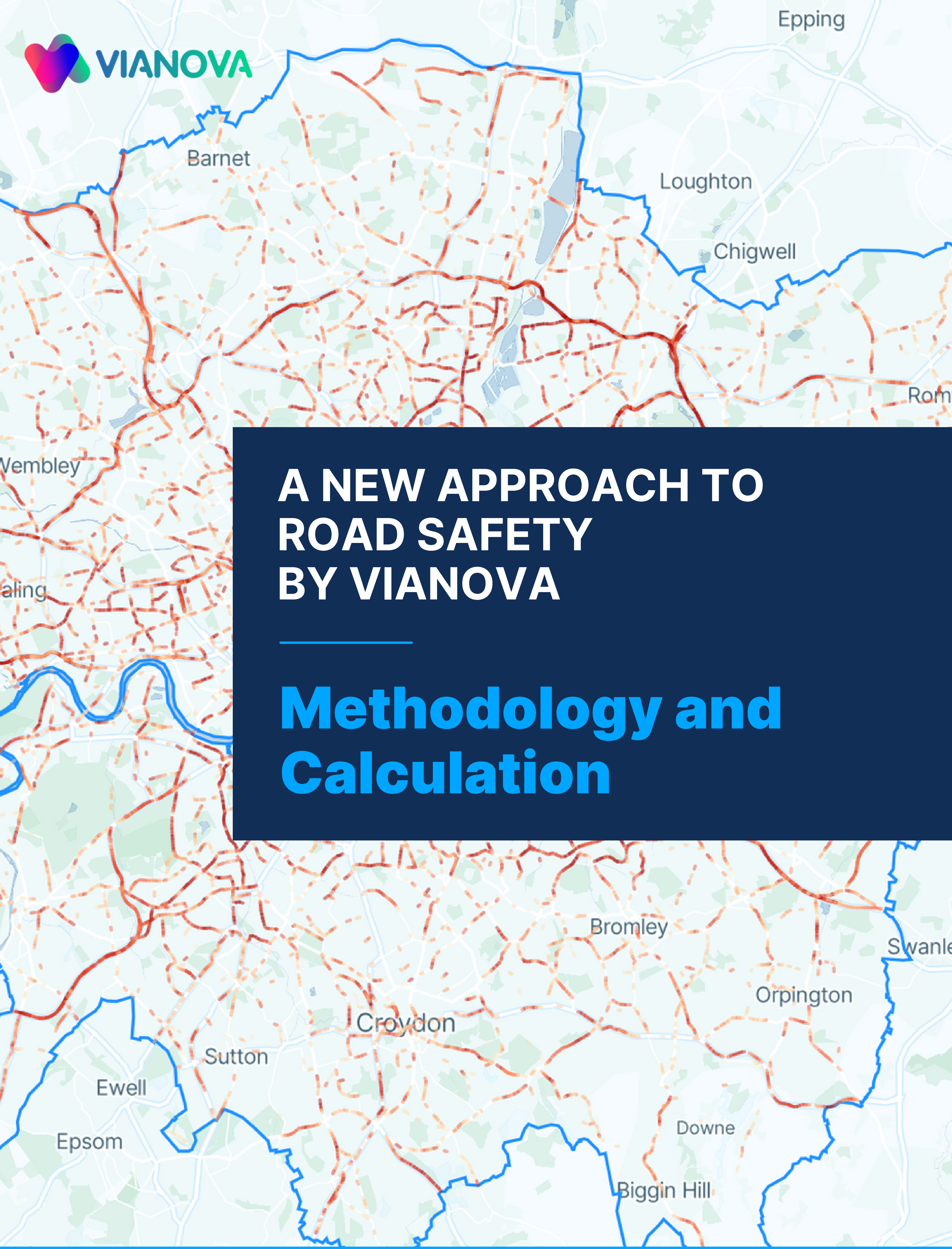
A NEW APPROACH TO ROAD SAFETY BY VIANOVA

Methodology and Calculation



About the author

Meet Alexander Pazuchanics, Vianova's Head of Product & Policy—a seasoned advocate for road safety and urban mobility. With a Master's degree in Public Management from Carnegie Mellon University, Alexander has shaped impactful policies throughout his career.

He is the Former Mobility Solutions Manager for the City of Seattle. Before that, he worked as Assistant Director for Policy, Planning, and Permitting at the City of Pittsburgh's Department of Mobility and Infrastructure.

Alexander leads all road safety projects at Vianova, overseeing initiatives from Transport for London to the City of Zurich. Widely recognized for his thought leadership, he authored "Pittsburgh's Smart City Vision: SmartPGH", outlining a comprehensive vision for smart and safe urban environments.

As a key figure at the intersection of policy, technology, and urban planning, Alexander continues to champion road safety, leaving a lasting impact on communities and cities worldwide.

Contents

Introduction

The development of IoT (Internet of Things) technologies has introduced significantly easier access to data generated by connected vehicles.

Vehicles such as cars, trucks, and vans are now equipped with a range of sensors that are able to provide data about positions, speeds, and headings, or the use of features such as windshield wipers, turn signals, and brakes. This type of data is available either directly from the vehicle manufacturer or from after-market fleet management services.

Vianova works with data OEMs and data aggregators representing approximately 40 million vehicles generating data in the United States, as well as approximately 30 million vehicles generating data in Western Europe, representing a broad swath of vehicle makes and models. You can learn more about Vianova's rationale for using connected vehicle data by reviewing our extensive **White Paper** on the subject, which you can find it on our website.

Connected vehicle data is not limited to information about cars and trucks. Increasingly, the use of **cellphone telematic data** or on-vehicle sensors has introduced new data about the behaviors of pedestrians, cyclists, and micro-mobility users.

This data is often less comprehensive but nevertheless provides insights into the density of vulnerable road users and some additional behavioral information.

There are several advantages to the use of connected data, particularly in comparison to empirical data collection (manual observations or the use of cameras or sensors).

Empirical data collection cannot cost-effectively be done over a wide area, whereas connected vehicle data covers the entire geography in which vehicles travel. Additionally, empirical data collection typically does not allow the ability to go “backward in time”- it is not possible to collect data before you know that you want to collect it.

In contrast, connected vehicle data is being collected regardless of whether its use is known, providing the ability to go backward and evaluate time periods in the past. And because connected vehicle data is always collected, it represents observations of **both** abnormal and normal behavior, helping to better understand the ratio between behaviors.

It is important to implement strategies to protect individual user privacy while unlocking the value that connected data can provide. Even when users provide consent to the use of their data for a broad range of purposes, governments can feel uncomfortable making use of data that could ultimately be re-identified to a specific user.

Two techniques, **aggregation** and **anonymization** can be used in parallel to generate insights while reducing privacy risk, especially in the context of GDPR.

Aggregation is the idea of bucketing data in a sufficiently large geographic area, with a sufficiently large time range, while ensuring the buckets have a minimum number of observations, making it difficult to identify a single traveler based on observation. For example, for most analyses, Vianova aggregates data to the street segment level, a geography of approximately 100 meters in length, with each street segment typically containing hundreds or even thousands of unique observations.

Anonymization is the idea of removing the possibility of reidentification of individual unique users. For Vianova, observations are stripped of unique identifiers, or a vehicle ID is used only in a limited circumstance to trace a vehicle’s movement through a small geographic area such as a corridor or intersection. These techniques, when used together, can ensure that the benefit of connected data is available with the appropriate safeguards for the user.

Connected vehicle data is not without its limitations.

Connected vehicles represent a small (but growing) share of the overall population of vehicles. As the vehicle fleet turns over and new vehicles come online, a higher share of the total fleet will become connected. Additionally, errors within an individual vehicle observation may exist, most obviously in the specific position of that vehicle. However, the significant quantity of data collected ensures that, in the aggregate, bias in the data is effectively eliminated.

The Components of Risk Aware

Vianova's **Risk Aware data product** is based on the concept of multi-factor risk assessment. Our approach is to combine multiple different indicators of potential risk, standardize them over uniform geographies, appropriately normalize them, and weight them in a mix appropriate to the user's preference.

There is no universally acceptable list of risk factors to include, nor is there a universally acceptable weighting which captures the priorities of all users.The Risk Aware tool is designed to be flexible and incorporate both the indicators and the weighting which matter to the end user.

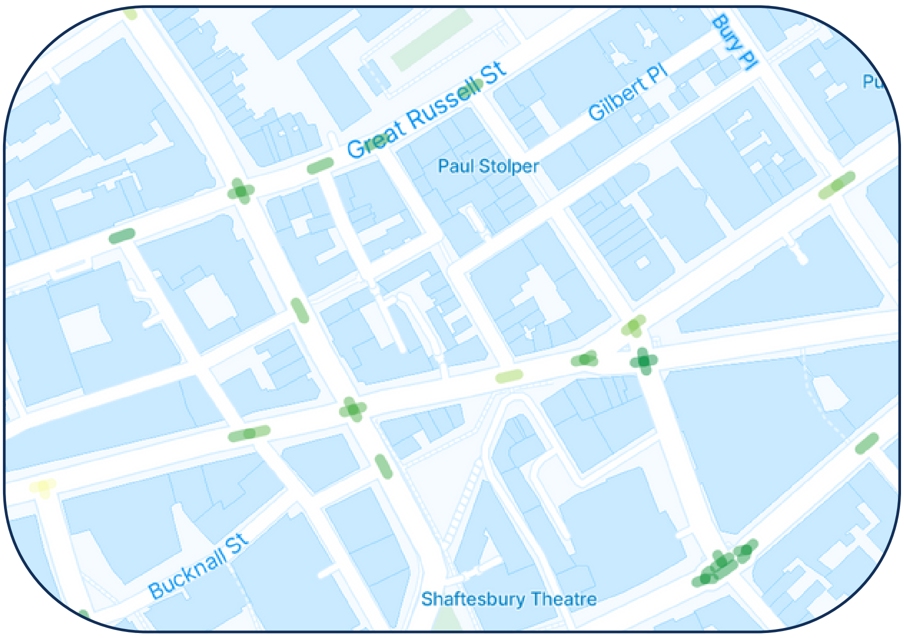
At every stage of the development, users are invited and encouraged to challenge assumptions, make modifications, and propose improvements to make the rankings usable in their specific contexts.

For any city where a Risk Aware ranking is created, Vianova creates a road network, by default built off of Open Street Map. Open Street Map is a globally recognized open-source data source for road networks, including information about the names, characteristics, and rules of roads across the world.

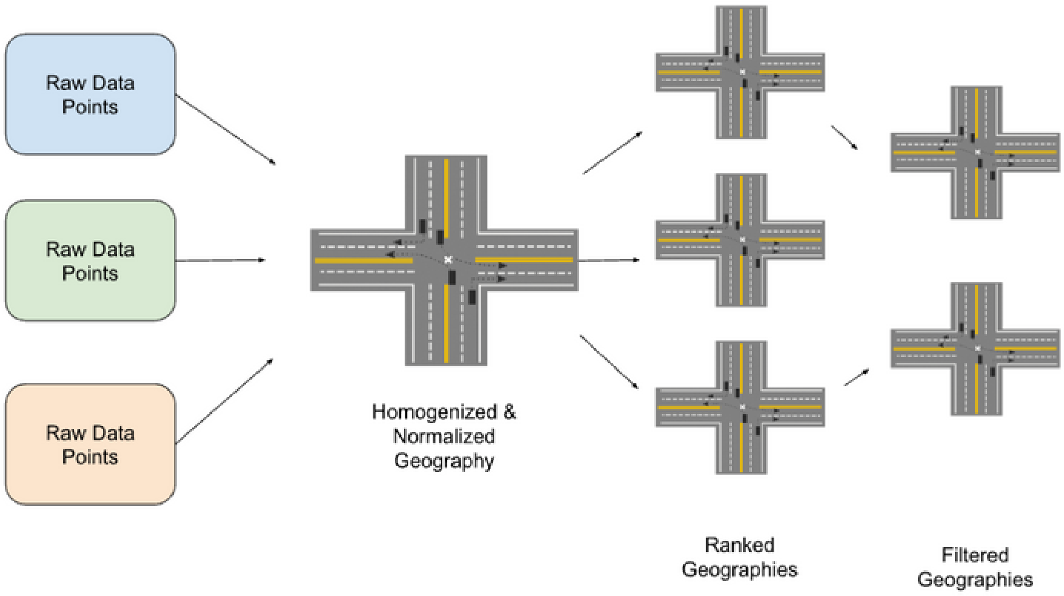
In order to make clear comparisons between roads, roads must be homogenized as much as possible, so that comparable lengths of road can be assessed against one another. The segment size of a road is based on the average segment size in the jurisdiction, but the optimal size of road segment is approximately 100 meters in length. Road segments can be combined together to create a corridor.

Where roads converge, Vianova creates a special road segment, an “**intersection**”. The intersection combines small lengths of road from each road feeding into the road network. The amount of the road network taken into the intersection is based on the “safe stopping distance” of the road based on its speed limit.

In other words, an intersection approached by a 20 km per hour road will have legs “shorter” than an intersection approached by a 90 km per hour road, because the distance necessary to stop on the 90 km per hour road is greater.



While this approach characterizes the default Risk Aware product, road layers can ultimately be customized depending on the characteristics of the jurisdiction and the preference of the user. Like the sources of data, each risk aware layer is tailor-made to support the prioritization process of the city.



Vianova typically begins by creating a Risk Aware layer using a mix of historic collision information and connected vehicle data. These components are standard by default, in a Vianova Risk Aware data layer, and are by default assembled with the following weighting:

Collision Data

- Historical collisions: 20%

Floating Car Data

- Heavy braking events: 20%
- Frequency of overspeeding events: 20%
- Intensity of overspeeding events: 20%

Connected VRU Data

- Density of VRUs- 20%

We will now review each of these factors in detail:

1. Collision Data

As discussed previously, a total reliance on collision data can leave critical blind spots in road risk identification.

However, this fact should not result in the total discounting of the data source. The reality is that collisions are one data set which should be incorporated into the overall picture of risk.

As collision data is not collected or reported in a standardized manner, Vianova works with the local jurisdiction to identify the best source of collision data. The quantity and quality of data on collisions varies significantly across jurisdictions.

Most data has a lag in collection and reporting of several months or one year. By default, Vianova uses the five most recent years of data of historic collisions available. Where information about severity is available, data included is limited to collisions involving a fatality or bodily injury. Collisions are associated with the road segment or the intersection they occur on.

Where the data exists, the Risk Aware process retains “raw” information about the collision which could be useful for further filtering and refining afterwards, such as the time of day, the meteorological conditions, the types of vehicles involved, and other information. These fields may be useful in further refinement at the project evaluation stage (see below).

2. Floating Car Data

Risk Aware incorporates a range of floating car data providers in order to generate the **optimal mix of points.**

We try to target approximately 5% of the total vehicles on the road, in order to produce a representative sample of observations.

In the Risk Aware approach, the number of observations is less critical than their distribution within the city- the goal is to have broad geographic coverage in order to compare road segments to one another- allowing for easier prioritization. In such a large data set, there is “smoothing” of the noise in the data, allowing trends to become more visible. Additionally, the initial Risk Aware layer contains a minimum of three months and as much as 18 months of data in order to smooth for seasonal variability.

Some floating “car” data is not from sedans or SUVs at all, but rather from trucks, vans, buses, or other types of vehicles.

This data provides a unique window into an important asset class and one with a disproportionate effect on the safety of vulnerable road users, as the weight and size of vehicles creates greater risk for pedestrians and cyclists. A mix which over-samples these vehicle types can produce additional insights into the relative risk of various road segments.

The specific floating car data partners Vianova works with in each jurisdiction will vary depending on the presence of the vehicle manufacturer in the market.

Because there is no standardized format for collection of floating car data, this means that the specific set of fields available for each observation may vary. Additionally, the particular data anonymization technique may also vary (ie, whether multiple observations are attributable to the same unique vehicle ID).

However, all floating car data providers will include some form of information about:

- The timestamp of the observation
- The speed of the vehicle
- The vehicle's heading
- The character or class of the vehicle

As long as this information is present, Vianova is able to generate the three components present in the default risk score- heavy braking events, frequency of overspeeding, and intensity of overspeeding.

Depending on the preference of the user and the availability of the data, other behaviors can be similarly mapped and ranked, including cornering actions, rapid accelerations, or other vehicle properties.

Harsh braking events identify a rapid deceleration of a vehicle. Several studies have identified the relationship between harsh braking events and collisions, indicating that the factor should be considered an indicator of potential risk.

In some cases, these events could be a “near miss”, while in others, they may not meet the technical definition but remain an indicator of risk.

Vianova calculates a “harsh braking” event in different ways depending on the data available from the specific provider. For some data providers, harsh braking is determined through a proprietary measure of force on the brake pedal, with brake pressures above a certain threshold being determined by a harsh brake event.

For other providers, brake events are identified as “harsh” based on the change in speed between two observations of the same vehicle (excluding events where the original speed was low, or where the rate of deceleration is gradual and over a long period of time).

Speeding is another behavior strongly correlated both to the presence of collisions, and particularly to their lethality. Even small changes in vehicle speed can dramatically affect the prognosis of a pedestrian or cyclist hit by a vehicle.

Vianova uses the concept of “overspeeding” to assess speeds above the posted speed limit.

An “overspeed” event is calculated based on the known speed limit of the road segment (by default provided in the Open Street Map road network).

If an observation exceeds the posted speed limit, it is treated as an “overspeed” event. Overspeed events are measured to produce two distinct scores.

“Overspeed frequency” indicates the share of observations on a road which exceed the speed limit. This measure is to give a sense of the scale of overspeeding- in other words, what percentage of travel on the road is above the posted speed limit.

“Overspeed intensity” compares the average amount by which overspeeding events exceed the speed limit.

Overspeeding intensity is measured to better understand the variance in overspeeding events. A street with many vehicles overspeeding by only 1 or 2 kilometers per hour may be treated differently than a street with a small number of vehicles overspeeding, but by significantly greater speeds.

3. Connected Vulnerable Road User Data

The protection of cyclists, pedestrians, and other travelers unprotected in an automobile (referred to as “vulnerable road users” by the European Union) is of paramount importance, particularly as cities and nations try to shift travelers out of automobiles for shorter trips in urban areas.

Collisions involving VRUs create a disparity in safety- a driver in an automobile with mandated safety features is significantly less likely to die or be seriously injured than a pedestrian or cyclist without similar protection.

The challenge historically has been collecting a robust set of data about pedestrian and cyclist behavior, as the modes are less connected than modern automobiles. However, three recent trends have enabled better understanding of the behavior of these travelers.

Shared micro-mobilty represents a rich data source of vulnerable road user behavior and presence.

Since the development of shared bikes and scooters occurred in the digital era, these vehicles are connected by default, providing rich information about vehicle position, trip starts and ends, and most importantly, routes used.

Though there are certain biases present in the data (shared micro-mobility users skew younger, wealthier, and more male than the vulnerable road user population as a whole), the data remains a valuable, easily accessible source of information on the street segments most frequented by shared mobility users.

The application of cellphone telematic data also provides a chance to develop new insights about vulnerable road user behavior.

Originally collected in order to support footfall studies and target marketing for retail operations, telematic data, when appropriately aggregated, can provide insights into the presence of slow-moving cell phones, giving cities a sense of the relative density of vulnerable road user traffic.

In certain circumstances, specialized connected vehicle hardware is available in order to collect data about traveler behavior. For example, hardware attached to bicycles can provide data on the swerving and heavy braking patterns of cyclists. If these sensors are deployed, cities can gain additional insights into vulnerable road user behavior.

Vianova uses a combination of these data sources, depending on their availability. The objective is to visualize the relative presence of vulnerable road users on each road segment. The density of observations over a period of six months is incorporated to account for seasonality and to smooth out special events.

Combining and Weighting Data

Vianova's Risk Aware score combines several data sources together to create a profile of risk for each road segment. Other data can be incorporated into this score, pending data availability.

For example, some cities prefer to incorporate the relative volume of vehicles travelling on a road in the assessment of risk. Vehicle volume can be a factor of risk, given that the larger number of total observations is correlated to the number of total near-miss events.

For others, data about road conditions or points of interest may be a factor in assessing risk. For example, the proximity of a road segment to a school, or its use as a freight corridor may be a factor in its risk profile.

Other cities may incorporate socioeconomic or demographic information in order to assess risk as part of a transportation equity focused approach to prioritization, such as streets in low-income neighborhoods or in areas with disproportionate numbers of young or elderly residents.

In the Risk Aware result, each road segment is ranked by the risk factors, producing an individual rank (the lowest rank being 1 and the highest rank being the number of road segments in the city).

This rank can be described in several ways, including a categorization (for example between 1-5) or a percentile (between 0-99).

These descriptions allow two roads to be compared against one another, thus enabling more clear prioritization. Road segments can be compared by the aggregate, **“weighted”** risk profile, or by the individual, **“component”** risks.

Streets can be prioritized using a “waterfall” prioritization, where one factor is emphasized. For example, roads within the top 10th percentile for collision history can be identified first, and then the balance of roads can be ranked by the combination of the other criteria. After rankings, road segments can be filtered or excluded based on certain criteria which are not factors in the risk score. For example, roads can be filtered by their speed limit, their typology, or their proximity to particular points of interest. This filtering may be important to respond to specific funding sources or program streams, for example a scheme to support safe transport for school students.

The Risk Aware layer is recalculated on a quarterly basis, as new data is made available. A quarterly update ensures that risk information is contemporary and changes in risk can be measured over time.

Incorporating Risk Aware into a Road Safety Program

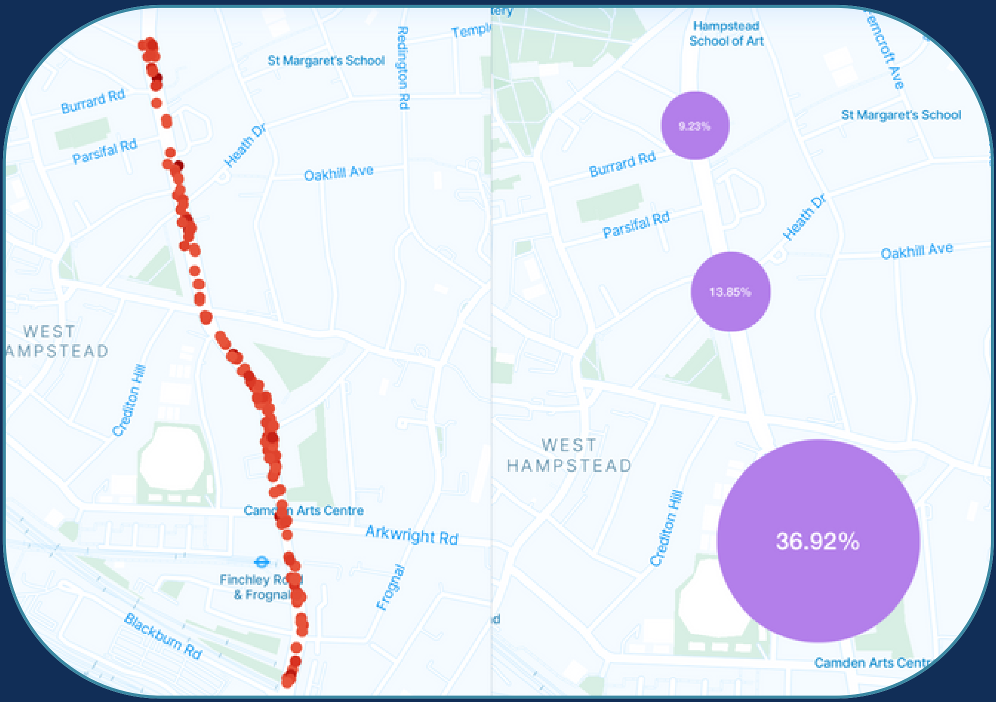
The identification of comparative risk is only the first step to producing a safer transportation network. The ultimate objective of the risk score is to assess risk over time, and to make interventions to make roads safer.

1. Evaluating Project Level Details

Once particular segments or corridors are identified, the various component data sources can be used to produce a more granular analysis of a project area.

In this case, the unique observations of collision history, floating car observations, or vulnerable road user observations are re-queried, enabling more fine-grained analysis of patterns within a smaller geographic scope.

As all of the data is time-coded, observations can be matched to other data sources sharing the same time, for example weather or traffic conditions.



The objective of performing project detail analysis is to pinpoint the appropriate type of road safety intervention to undertake.

In these project details, additional empirical data can be incorporated to compliment the data connected by connected vehicles. For example, field observations or sensor based turning counts can be incorporated into a project detail analysis to pinpoint the right type of intervention.

2. Making the Improvement

Once data has been collected and analyzed, traffic engineers can begin to design the optimal solutions for a road safety intervention.

Every project requires on-the-ground investigation by a professional engineer, and no amount of connected vehicle data is going to replace the expertise of a local traffic expert.

But depending on the type of problem identified through the data and observations (unsafe mixing of vehicle types, speed, visibility, etc.) the right types of designs can be implemented.

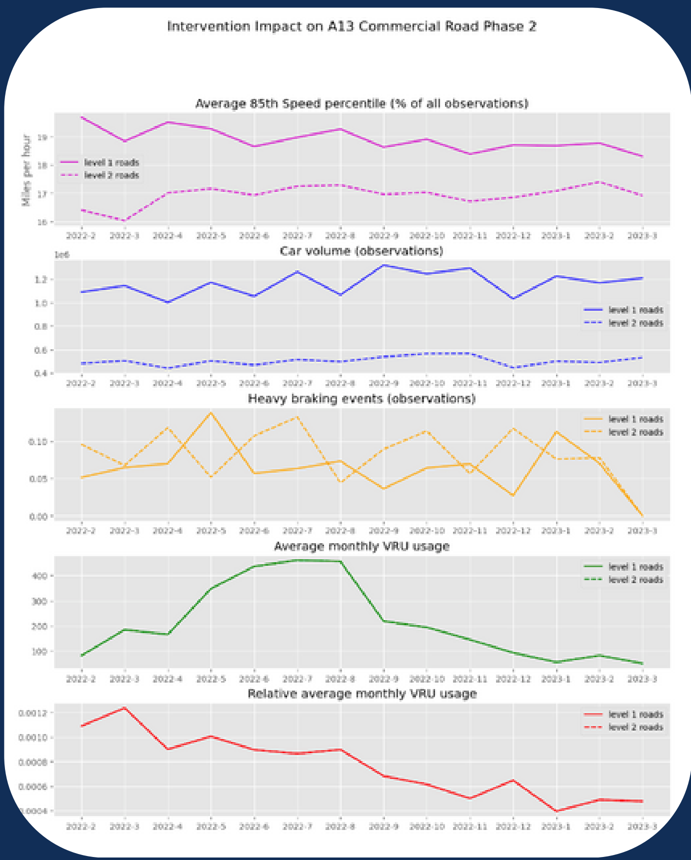
Tools such as the Global Road Safety Facilities Guide for Road Safety Interventions, the National Association of City Transportation Officials Urban Street Design Guide, and the CROW Road Safety Manual are examples of comprehensive guides for the most appropriate intervention strategies to be deployed in urban and semi-urban environments.

Engineers will also likely incorporate other objectives in the project, for example to support better urban design, improve stormwater management, and enable better connections for public transport.

3. Monitoring the Results

A collision-centric data collection technique could take months or even years before sufficient data is gathered to determine the success of an intervention. Because collisions are (thankfully) rare events, a long window of time is needed to collect data on whether they have occurred.

In contrast, connected vehicle data is being collected before, during, and after the projects are implemented, enabling an easy opportunity to monitor changes in vehicle speeds, driver behavior, and vulnerable road user density.



An analysis of this behavior can be conducted in a much shorter time frame, and can also be complimented by additional empirical data, such as traffic cameras or sensors or manual observation.

Connected vehicle data can even be used to monitor the potential “spillover” effects of an intervention, for example by examining adjoining streets, streets in close proximity, or known detours. The goal should be not to shift risk around, but rather to reduce it entirely.

Communicating Progress

An important part of the implementation of a successful road safety program is communication to stakeholders and travellers. Through simple, legible dashboards, information can be made public through reports, updating frequently or even in real-time as new observations are taken.

Additional Uses of Risk Aware Methodology

While the application of Risk Aware scoring for infrastructure managers is evident, there are additional potential applications of a connected vehicle -based approach to risk identification.

As connected vehicles become more ubiquitous and the data produced by them becomes better understood, new opportunities emerge to support better decision-making from other partners.

1. Routing Applications

Understanding the relative risk of a road segment can be an important additional tool for vehicle routing apps.

The total risk of a route, or the relative risk of specific segments can be weighed against other routing optimization goals, such as fastest time, lowest cost, or lowest potential emissions. Roads can be dynamically “re-ranked” based on the journey in order to avoid segments which cross an unacceptably high threshold, particularly at a particularly risky time of day or under adverse meteorological conditions.

This ranking technique can be particularly beneficial to routing for pedestrians and cyclists, who may route based on a different set of criteria than travelers in a vehicle.

The presence of historical risk can also be used to suggest that a traveler has entered an area of increased risk. Routing tools must be careful to avoid providing alerts which distract a driver more than they provide the driver with information.

But with effective user experience design, alerts can prepare a traveller to be extra aware when they are approaching an area known to have historic risk factors.

2. Insurance Applications

Over time, a historically large data set emerges, which can support the insurance industry's objective to reduce risk and promote safety.

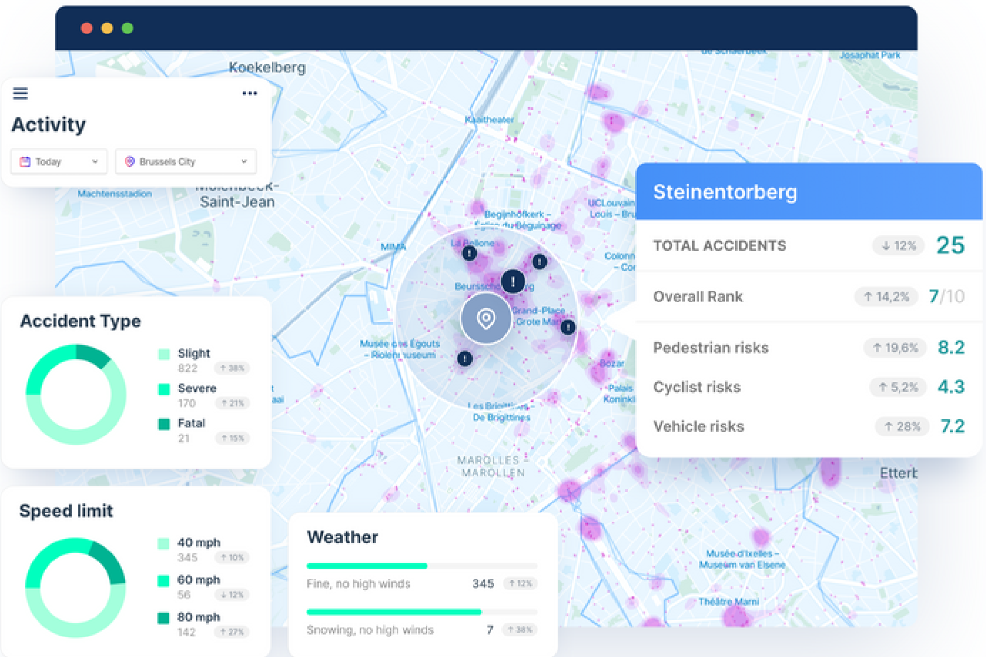
The emergence of "pay by the kilometer/mile" insurance, as well as premium discounts for safe driving behavior, produce an additional set of connected vehicle data which could support a more robust Risk Aware assessment. Infrastructure managers such as city and regional governments can work with local insurance providers to identify an appropriate technique for sharing data with appropriate privacy protections.

An understanding of the baseline risk of the road network can also support more sophisticated pricing of insurance applications. For example, an individual driver's behavior can be compared to the relative risk profile of the road, especially as it relates to speed and braking behavior.

Conclusion

The use of connected vehicle data in risk identification and safety development is an exciting and iterative new development. As a software as a service and data solutions as a service company, Vianova is excited to continue to improve on the use of new data sources to help make prioritization decisions easier and to better document improvements to the road network.

Achieving **Vision Zero** will take a comprehensive effort, incorporating new strategies and techniques to build safer streets for all travelers.



Want to learn more about our methodology?
Interested in making roads safer in your city? Contact us!

inquiry@vianova.io

www.vianova.io

THANK YOU

inquiry@vianova.io

www.vianova.io