

# An ecosystem approach to the design of sensing systems for bicycles

Ricardo Cabral<sup>[0000-0002-5471-1778]</sup>, Eduardo Peixoto<sup>[0000-0001-5998-7960]</sup>,  
Carlos Carvalho<sup>[0000-0001-8052-4178]</sup> and Rui José<sup>[0000-0003-3547-2131]</sup>

Algoritmi Research Centre, University of Minho, Portugal  
rui@dsi.uminho.pt

**Abstract.** Bicycles equipped with sensors, processing capacity and communications can be a promising source of data about the personal and the collective reality of urban cycling. While this concept has been attracting considerable interest, the key assumption is the design of a closed system where a uniform set of sensing bicycles, with a concrete set of sensors, is used to support a specific service. The core challenge, however, is how to generalise sensing approaches so that they can be collectively supported by many heterogeneous bicycles, owned by a multitude of entities, and integrated into a common ecosystem of urban data. In this work, we provide a comprehensive analysis of the design space for on-bike sensing. We consider a diverse set of sensing alternatives, the potential value propositions associated with their data, and the collective perspective of how to optimise sensing by exploring the complementarities between heterogeneous bicycles. This broader perspective should inform the design of more effective sensing strategies that can maximise the overall value generated by bicycles in smart cycling ecosystems and enable new cycling services.

**Keywords:** smart mobility, cycling data, sensing bicycle

Smart cycling is a very broad concept encompassing the many paths through which cycling is being incorporated into the connected and smart transport networks of the future. It can be described as the shared, real-time and collaborative application of data, communication technologies, products, and services through both private and public actors, to help best move people individually, and collectively, across the urban environment [1][2]. The application of Information Technologies in this domain is already a dominant factor for the successful adoption of shared bicycles [3], where they provide some of the collective features that characterise those services, such as the ability to find nearby bicycles or seamless pay per use. We can envision the evolution of this paradigm, with increasingly more bicycles being instrumented in ways that can add value to the entire ecosystem.

In this study, we analyse the role that bicycles equipped with sensors, processing capabilities and communications might have as a valuable source of data about the personal and the collective reality of urban cycling. Sensing systems for bicycles have been explored in many different ways [4–7], but the key assumption is always the design of a closed system where a uniform set of bicycles, with a concrete set of sensors, is used

to support a specific service. Regardless of the technical merits of any particular systems, the larger challenge is how to generalise sensing approaches so that they can be collectively supported by many heterogeneous bicycles, owned by a multitude of entities, and operated under a common ecosystem of cycling data. This is a problem that goes beyond the existence of a single platform or the usage of standard protocols for data representation. It also involves the data collection procedures, the types of sensors used, data processing algorithms, the operation of sensors or even the details of how the sensors are deployed on the bicycles. Closed systems lack the generalization that is needed to allow different designs to be explicitly expressed in a way that describes their similarities, their differences and their contributions to a common cycling ecosystem.

### 1.1 Research objectives

A smart cycling approach should allow sensing systems from multiple bicycles to compete for the best way to address specific data collection goals and to complement each other to serve broader data needs. Rather than assuming uniformity, the model should explore heterogeneity as the best way to allow these sensing systems to serve the very diverse set of data needs associated with Smart Cycling.

In this work, we aim to explore this path by providing a comprehensive analysis of the design space of bicycle sensing systems. Our research objectives can be summarized by these fundamental research questions:

- What is the range of sensors that should be integrated into the sensing systems of bicycles to provide a comprehensive view of urban cycling activity and its context?
- What are the key trade-offs involved in the design of sensing bicycles and how can they be optimised to reach the most effective results in the context of broader smart cycling ecosystems?

To explore the possible answers to these questions, we have started by developing our own prototype of a bicycle sensing system. Using this prototype as a research context, we have experimented different designs by creating multiple variants of sensing systems. This provided a key learning context to understand the sensitivities associated with this sensing context and assess the real value that can be obtained from various types of sensors. This experimental work was complemented with an analysis of the sensing design space for bicycles. This analysis was organised around 3 layers, more specifically the universe of viable sensors, the value propositions offered by those services through the services they enable and the collective perspective of how to combine heterogeneous sensing systems to optimise their value to the whole ecosystem. This broader perspective should provide new insights for the development of sensing systems for bicycles and the understanding of their role in the larger context of urban cycling services.

## 2 Related Work

Previous work has already explored many variants of bicycles equipped with various types of sensors. The smart e-bike monitoring system (SEMS) [8] is a platform for the real-time acquisition of usage data from electrically assisted bikes. The system collects e-bike data including location, rider control data and power data. The SEMS data feeds an online interface for data analysis, for riders to view their own data and sharing on social media. BikeNet [9] is a mobile sensing system for mapping the cyclist experience. BikeNet uses a number of sensors embedded into the bicycle so that it collects data about the cyclist's ride. BikeNet uses two approaches for data synchronization: an opportunistic approach where it uploads data when the bicycle opportunistically finds a wireless access point; and real time synchronization using the cellular data channel of the cyclist's mobile phone. SensorBike [7] is mainly focused on capturing data that influences security and comfort while cycling. It includes power sensors, vital sensors, accelerometers and vibration sensor, environmental sensors, distance sensors and cameras. The goal is to understand the cyclist's perspective and contribute to the future of cycling planning. The Smart e-bike [11] explores the combination of sensors associated with the physical characteristics and physical condition of the rider with the motor control system. The aim is to help the cyclist in situations when the values of his physical condition or the parameters of the environment are critical. The system includes pulse sensors, gyroscope/accelerometer, speedometer and GPS.

In addition to special bike configurations, there are also add-ons that include a major sensing functionality. The See.Sense system [10] is a set of augmented cycling products with the capability to collect and share trip information. The See.Sense smart bike lights include a GPS, accelerometers and GSM communication service. The system can collect data from cycling trips and alert cyclists when they are entering a more dangerous zone. The device is able to register 800 data points per second, enabling it to monitor road conditions and even alert the owner when the bicycle is being moved or stolen.

Air quality has been particularly popular as an application domain for sensing on bicycles. The motivation for using bicycles in this context is the possibility to collect data from many locations, many of which might not be reachable by car, and also the fact that the bicycle is not an air pollutant itself and therefore does not interfere with measurements.

Aeroflex [4] is a sensing bicycle for mobile air quality measurements. The system measures Ultra Fine Particles, Particle Matter, Black Carbon and CO. It also collects GPS location, sound, images, vertical acceleration, temperature and relative humidity. The Sniffer bike [6] is equipped with a particulate matter sensor developed in collaboration with the Utrecht Province to be attached to bicycles. It measures air quality every ten seconds and shares data with the Civity data platform every minute. The Copenhagen Wheel [5] is an electric bicycle system developed to transform any bicycle into a smart bicycle by replacing the rear wheel with the Copenhagen Wheel. The wheel is able to collect the location data of the bicycle and other data to improve cycling and cyclists' experience. It also collects data about air (CO<sub>2</sub>, NO) and noise (db) pollution, congestions, road conditions, relative humidity as well as temperature.

Bike sharing companies also use different forms of sensing on their bicycles [12] They typically collect data related with position and timing from routes made by their users. The primary purpose of this data collection is for pricing processes, but it also enables the construction of data related to the users' profile so that the system can be optimized for them. Accordingly, they work on this data to improve efficiency and increase user satisfaction.

These systems demonstrate the variety of approaches that can be explored when combining sensors with bicycles. What differentiates our work is mainly the broader perspective on bicycle sensing. Rather than aiming to propose yet another sensing bicycle, our work aims to uncover the key technical trade-offs and design sensitivities associated with sensing bicycles. The ultimate goal is to allow a heterogeneous and open-ended set of sensing bicycles to coexist, while improving their capability to operate as complementary approaches under a common cycling ecosystem.

Regarding data usage, route characterisation has always been one of the most common goals for on-bike data collection. There is a vast research literature on the main determinants in cyclist's decisions to cycle or to decide where to cycle. These are important indicators for data that should be widely available in cycling ecosystem. Ehrgott et. al. [13] observed that some cyclists prefer to travel longer distances in order to include cycle facilities on their routes. The shortest route is not necessarily the most attractive route to cyclist, making route selection a bi-objective routing problem, where the aim is to generate a set of compromise solutions that is considered efficient. A route is called efficient if, given the same travel time, there will be no route with higher level of suitability and given the same level of suitability, there will be no route with shorter travel time. In another study by Winters et. al. [14] a group of cyclists answered 73 survey items, grouped into 15 factors that might influence their likelihood of cycling. These factors highlight the importance of the location and design of bicycle networks. In another survey with 65 commuter cyclists, Segadilha et al. [15] asked participants to classify the importance of 18 cycling factor in a scale from 1 to 5. The factors identified as the most important were grouped in these 5 categories: road, traffic, environment, trip and route as a whole. A survey conducted by Felix [16] with cyclists and non-cyclists has sought to understand their motivators, triggers and barriers towards cycling. Results have shown that both groups consider the issues related to the perception of safety, physical effort, lack of a safe cycling network, and bicycle ownership as important barriers to take up cycling. Broach et al. [17] developed a GPS model to collect data from cyclists so that stakeholders could answer questions about types of infrastructures, preferences and the links between bicycle infrastructures and cycling behaviour. Su et. al. [18] identified some online bicycle trip planners and concluded that these online trip planners rarely provide the complete set of route selection criteria required for a bicycle trip planner including fast, safe, simple and attractive routes. These studies provided us important insights on the very diverse set of goals that may be relevant when collecting cycling data and on the identification of key target services for cycling.

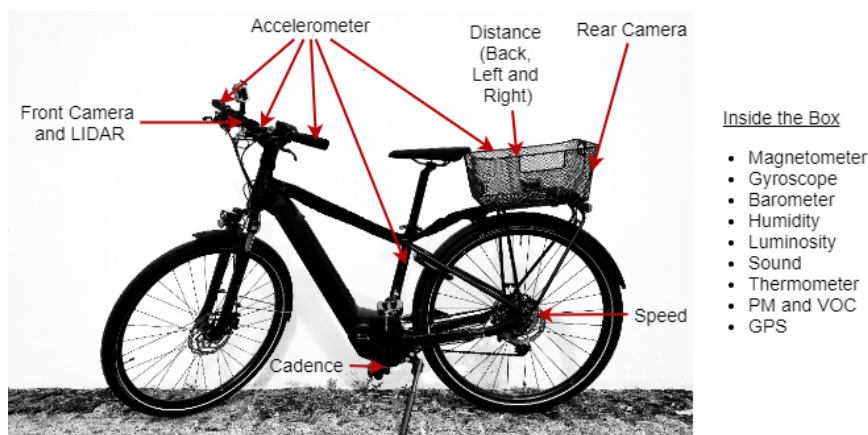
### 3 A prototype of sensing system for bicycles

Our own prototype of a sensing system for bicycles was an important research tool in this work. More than a goal in itself, this bicycle allowed us to experiment many sensing alternatives and uncover multiple challenges associated with sensors on bicycles. The bicycle design is not attached to any particular form of sensing or any particular usage of the data collected. Instead, the system is designed to support a broad range of sensing possibilities and be instantiated in various ways, according to the requirements of specific research goals. Therefore, there is not a unique prototype specification, but rather a set of diverse prototype instances created throughout the project to study specific perspectives of sensing bicycles.

These various instances may differ in many ways, such as their specific set of sensors, the way those sensors are physically deployed or the specifics of the data collection process. The structured exploration of these many possibilities provides a broader view of the whole range of on-bike sensing possibilities and creates actionable knowledge on how to define specific sensing strategies for specific sensing goals. It also creates a context where it can be possible to compare multiple competing alternatives for similar forms of sensing. To improve the generalization of the results, we also avoid any major assumptions about the ways in which a system like this could be embedded into a common bicycle, e.g. embedded directly on the bike or attached as a removable add-on. We will however, try to explicitly identify any effects associated with sensor positions or with other design decisions that may also impact on the viability of particular deployment approaches.

#### 3.1 Bicycle instrumentation.

Our prototype bicycle was based on an electrical bike equipped with a Bosch motor. **Fig. 1** shows an instance of our sensing bicycle with a common combination of sensors and their deployment positions.



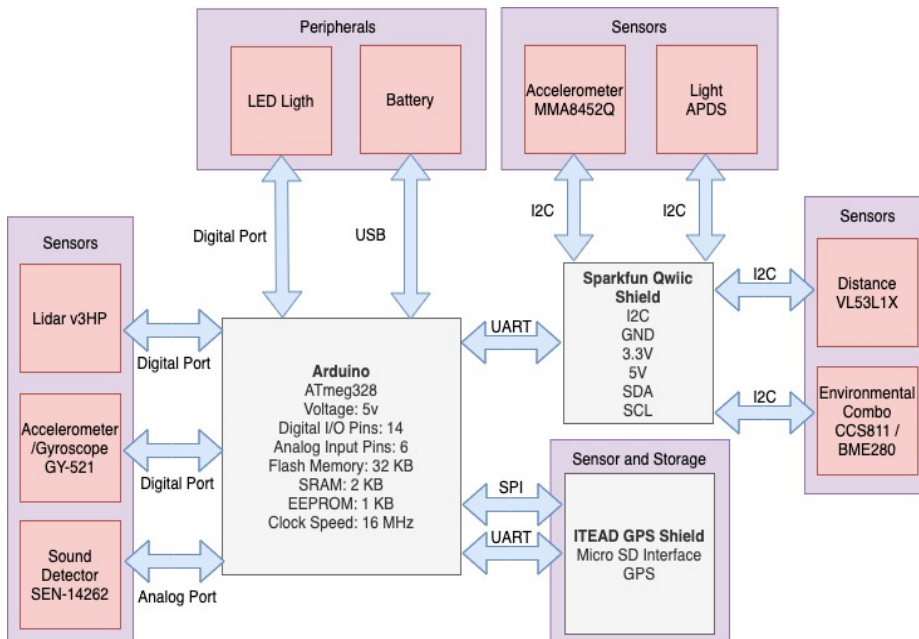
**Fig. 1.** Sensing Bicycle instrumentation

The instrumentation of the bicycle essentially involved the selection of specific positions for particular sensors and for the central control unit, as well as the installation of cables to support the connections between them. To facilitate deployment, a rear basket was added. This is where the central control unit was placed, together with most sensors and the power battery. To support the cable connection between the central unit in the rear basket and the various sensors that needed to be placed at the front of the bicycle, two small plastic tubes were attached to the bicycle frame and wrapped with black fabric.

### 3.2 System Architecture

The system follows a centralized architecture with a single control unit connected to multiple sensors via cable connections. In this integrated design, all core functionality is provided as part of a single system, developed and deployed by a single entity.

The system comprises an Arduino UNO R3, a shield with Qwiic extension ports, a micro-SD storage shield, a LED to signal the system status, a 12V battery and a number of sensors controlled from the Arduino board. **Fig. 2** represents a particular instantiation of the system with four distance sensors, two accelerometers, one environmental sensor, one light sensor and one sound sensor.



**Fig. 2.** Sensing system architecture

Since all the sensors are connected with cables, their maximum number is limited by the number of available connections. However, this has not been a major problem at this stage, because for each prototype instance there is normally a specific goal, for which only a particular sub-set of the potential sensors is needed.

### 3.3 Sensor sampling algorithm.

The control software running on the Arduino board executes all the necessary tasks for continuously collecting sensor data and writing it to an output file in the SD Card. The sensor sampling algorithm begins with the setup phase, which involves the activation and initialization of the sensors. A LED light on the bicycle handlebar provides basic awareness about the current system status and will be on if all sensors defined as critical are operational. Once the setup is successfully completed, the system enters a data collection loop to obtain data from the various sensors. Given the specific properties of the Arduino Board, each cycle is a sequential process which needs to go through all the system's sensors to obtain their data.

Even though there is a single data collection cycle, we need to accommodate very different requirements regarding the sampling rates of various sensors. Accelerometers will typically have the highest rate. For example, for road anomaly detection in a car driving at 50 Km/h, Silva et al. suggest a 50 Hz sampling rate [19]. On the contrary, for GPS data a single sample per second can already be seen as intensive tracking. To address these different sampling rates, we allocate to each sensor a variable that defines the respective frequency.

These sampling frequencies are determined by the envisioned usage of the data, which affects the potential relevance of higher sampling rates, and by the specifications of the sensors, which may limit the maximum rate at which data can actually be produced. The maximum sampling rate possible is also bounded by the performance of the board and the execution time of each cycle.

### 3.4 Data collection.

Data collection is the ultimate goal of any sensing system and is therefore a critical part of this research on sensing bicycles. Our data collection process is structured around three complementary data collection processes, as represented in **Fig. 3**. They are distinct processes because they serve different goals, but they also have many interdependencies and shared steps.

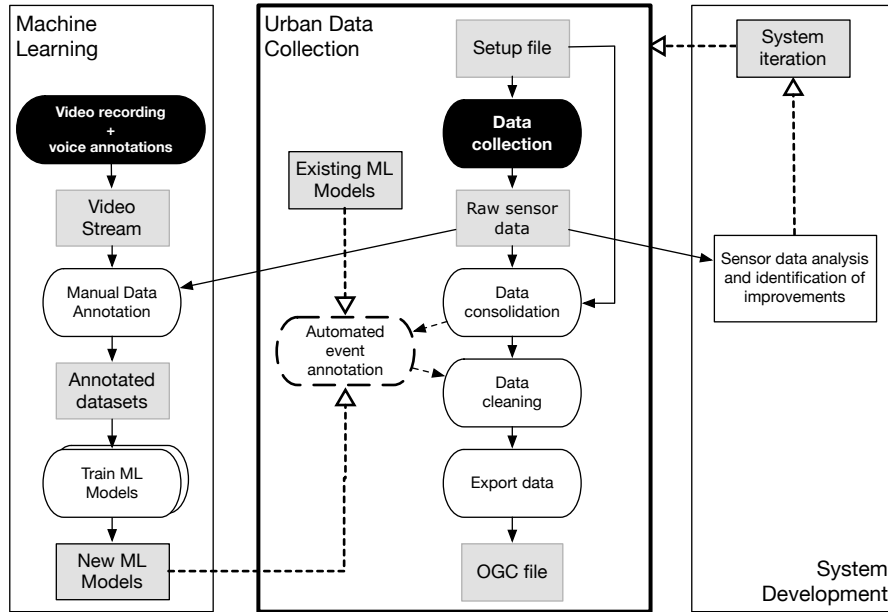


Fig. 3. Complementary Data Collection Processes

Urban data collection is what we envision as a generic, large scale data generation process, where a large number of different types of sensing bicycles are in operation to regularly produce the data needed to support various types of urban cycling services. This is the most generic process supported by the sensing bicycle and essentially corresponds to the normal execution of the software on the control board unit. The raw sensor data is a CSV file containing the data collected from sensors during a trip. The data is structured in a Tidy Data format [20], with each sensor output being registered as a new observation, consisting of line with a time stamp, the data context and the respective value. This format provides the flexibility needed to operate the system under multiple sensing bicycle configurations, each with its own particular set of sensors. It also provides a simple solution to address the very different sensing rates of the various sensors in our system, as each sensor is free to output its data to the file without any dependencies on the others. At the end of the process, data is exported to OGC SensorThings format [21]. This json format provides the necessary self-description of the data, allowing it to be processed without the need for further information, other than what is described in the file itself.

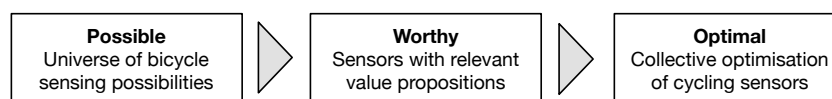
The sensing bicycle can also be used to produce training data for machine learning models. In this case, data collection should be done under a more structured protocol to control the specific variables involved and facilitate the annotation of relevant events. Videos can be used to support this annotation process. They show the concrete situations being experienced during the ride and they will normally include voice descriptions of the events made by the rider. Each study will conduct whatever specific data



manipulation processes it may need to produce specific Machine Learning models. Those ML models may at some point integrate the normal data processing flow to support the automated identification of such events during the ride and shared them as an additional data stream produced by the bicycle.

## 4 What to sense in a sensing bicycle

A central research question in our work is how to determine the set of sensors that should be available on bicycles to support their role as sensing entities. Our experimental work on sensing systems for bicycles may provide us with many insights about the implications of certain design decisions or about the use of concrete sensors, but it does not take us closer to the answer to this question. This is a very open-ended question, which would normally be easier to answer if made in the context of a concrete systems with specific sensing goals. Trying to provide a generic answer can be much harder, as it requires the generic exploration of sensing possibilities, their viability in the cycling context and also the potential value of the respective data for concrete application domains in the cycling ecosystem. The overall process can be seen as involving three successive layers of analysis, which progressively reduce the set of sensors to be considered, as represented in **Fig. 4**.



**Fig. 4.** Layers of analysis in the selection of sensors

The following sub-sections will explore each of these perspectives in more detail.

### 4.1 Universe of cycling sensing possibilities

The starting point in this analysis is the universe of bicycle sensing possibilities, which we define as the range of physical phenomena that are measurable, meaningful in the cycling domain and viable. Sensor viability is determined by their cost and by the practical implications of their deployment on the bicycles, such as dependability, size, volume, mass, and longevity [10]. For the purpose of this study, we will assume a loose interpretation of viability, and consider that the universe of sensing possibilities corresponds to all the sensors previously suggested for this purpose in other studies and market products.

Given the wide range of sensor possibilities and their various applications domains, we structured this analysis around a set of sensing profiles, which represent particular sensor types that share the same type of sensing phenomena and a similar application domain. This profile structure explores the fact that the various sensor possibilities are not independent between each other. Similar sensors will typically be used for similar purposes and most of the time will just be alternative solutions to the same problem.

By treating them as a whole, we significantly reduce the complexity of the analysis, while still maintaining the capability to make meaningful connections between sensors and their value propositions. We will now present the proposed list of 8 sensing profiles:

**Position profile.** The position profile assumes the existence of position sensors that can determine the position of the bicycle. Position data play a major role in the sensing process, not just as a core data itself, but also as a way to georeference data generated by other sensors. This profile may thus be often used in combination with other profiles. Position sensors rank high on viability as they can be low-cost, discrete and pose no major processing requirements, especially when no real-time data is involved. Whenever there is the need to determine the position of the bicycle on a regular basis, a GPS receiver is the common solution.

**Motion profile.** Motion measurements can provide key data to understand the smoothness of the ride. When properly analysed by machine learning models this data can produce high-level knowledge about the motion patterns of a ride, including the driving style or the identification of riding events, such as braking, turning, road bumps or irregular tracks. This type of data can serve many relevant purposes, including those related with safety. The motion profile is mainly composed by IMU sensors, such as 3-axis accelerometers, 3-axis gyros and 3-axis magnetometers. These sensors can be low-cost. However, they produce large volumes of data that needs to be processed locally or otherwise transferred to a server. Either way, this may introduce additional requirements regarding processing capability, storage or network connectivity.

**Environmental profile.** Environmental sensors can measure a wide range of environmental characteristics, such as gas concentrations, particles in the air, light intensity, humidity level, atmospheric pressure or temperature. In the cycling domain, this data can be useful to inform each cyclist about the level of exposure to hazards elements experienced during daily rides and to complement information about route quality, e.g. the presence of certain particles can be a predictor for heavy motor traffic. This data can also be very useful beyond the cycling context, as bicycles are frequently recognized as ideal vehicles for mobile environmental data collection [22–24]. These sensors do not place many new additional requirements, but some of these sensors can be particularly expensive and require specific installation settings or calibration procedures that may not be compatible with large scale crowdsourcing data collection.

**Surrounding profile.** The surrounding profile includes distance sensors to provide a perspective of route quality and safety. Distance between the bicycle and nearby objects defines the free surrounding space, which can be an important indicator for the safety risks associated short distances to other vehicles or to potential obstacles. Common distance sensors, possibly pointing into different directions, may provide simple data about nearest objects. A more sophisticated perception of the surrounding space can be

created with data obtained from Lidar and Radar devices, as they offer the additional capability to make 3D representations and identify object sizes. However, LIDAR and Radar sensors, can be much more demanding in regard to their deployment on the bicycle and their data processing requirements.

**Rider.** The rider is at the centre of the cycling experience and therefore data from the rider can also provide key insights about that experience. The Heart beat rate is commonly used in sports contexts, but it may also be used as proxy for the level of stress experienced during the ride. A cadence sensor can measure the rotations per minute performed by the rider on the pedal and therefore the effort the cyclist is making.

**Video.** With proper processing capability, video can be a powerful form of sensing, but the automated collection of data from computer vision processes is not common on bicycles, mainly due to the strong processing requirements. For the purpose of this study, we will embrace video mainly as a source of ground truth data through the creation of autonomous video streams for later annotation of relevant events. In this context, we can expect the video profile to be used mainly in the context of professional data collection activities.

**Sound profile.** While potentially a form of environmental sensing, sound can be used to support more advanced interpretations of the cycling context. The level of sound obtained from a simple sound level sensor may be used as a proxy for traffic levels. Sound data obtained with a microphone may enable machine learning methods that explore the particular sound frequencies associated with riding events, e.g. different types of surface will typically produce distinct combinations of sound frequencies and being overtaken by a car may also produce a unique sound signature.

**Proximity.** Proximity sensors enable bicycles to detect the presence of nearby entities without any physical contact. This is not concerned with the physical proximity to surrounding objects as is the case with distance sensors. This is about logical proximity to recognisable entities, such as other bicycles, cars or bike counters, that are able to identify themselves and possibly engage in more sophisticated communications. Bluetooth, BLE or RFID are commonly used for this purpose.

These 8 profiles do not correspond to specific bicycle instances. They are meant to be combined in many different ways. In particular, the position profile can be expected to be often included when one of the other profiles is also included. By describing sensing systems from the perspective of the supported sensing profiles, it becomes possible to have a common framework to discuss many different sensing bicycle designs.

## 4.2 Sensors with relevant value propositions

The universe of cycling sensors can be very large, but any realistic cycling sensing setup will have to be shaped by the worthiness of the data produced. Worthiness is essentially determined by the demand side and the concrete value propositions that can be associated with the data produced by those sensors. To scope this analysis, we will focus on cycling routes as the core data entity in our model. Our problem can thus be defined as identifying the types of data that are the most relevant to characterise cycling mobility on a given route network. We mainly consider two major types of data: trips that define movement, i.e. where bicycles are passing and at what speed; and route annotations that characterise routes, i.e. what type of road surface is there or what types of riding events are generated. These two dimensions can be combined with the individual and the collective perspective. For example, combining trips with the collective perspective corresponds to route traffic measurements, i.e. volume metrics. From an individual perspective, trip information can be used to track individual progress and achievements. The combination of these dimensions and their key results are depicted in Fig. 5.

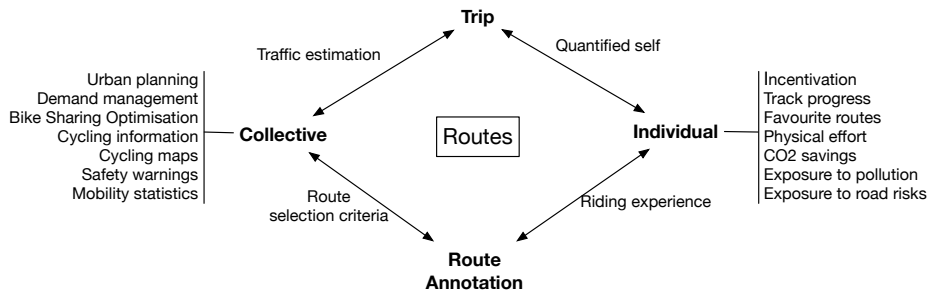


Fig. 5. Simplified model of cycling data needs

We can explore worthiness by exploring the possible associations between the many types of sensors and these different types of data needs. Trip information is largely provided by GPS devices, which generate a time series of positions along with the respective speed and orientation information. Usage of this type of data is very common with Traffic Management services for motorized vehicles where many of these same metrics are intensively explored to produce a general perspective of urban traffic, supporting the generation of live traffic maps and accurate and dynamic estimations of travel times. While there are many commonalities to be explored, cycling data differs from those models in that real-time data is not as relevant as it is with cars. A major goal with cars is to manage ongoing traffic and avoid congestion. With bicycles, volume is also very important, but mainly to understand demand trends and adjust accordingly. We can thus consider trip information as strongly relevant, even if it does not need to be shared in real-time.

On the contrary, route annotation data is very unique and much more relevant for bicycles than it is for cars. With cars, the navigation focus is very strongly on the fastest

route. With bicycles, route selection requires a much broader set of data about routes, particularly in regard to safety, comfort or road gradient. Data collected by bicycles should thus be able to support the generation of indicators that represent complex and multi-criteria route selection processes.

The research literature includes a vast body of research on route selection criteria [13–18], which we have used to match sensing possibilities to worthy data applications. In particular, we have considered data services along the various layers of the Bicycle Pyramid model proposed by the BITS project [25]. The pyramid model is derived from Maslow’s pyramid of needs and is also organized according to the principle that the needs at the bottom of the pyramid need to be met before the next level becomes relevant. The five proposed layers, from bottom to top, include Safety & Reliability, Speed, Convenience, Comfort and Experience. This is an extensive and prioritized view of multiple elements that can shape cycling activities and ultimately determine its adoption, and therefore provides a good framework for matching data with the value it can generate. Once again, we will use the abstractions offered by sensing profiles to facilitate the association between any specific sensing bicycle designs and concrete urban cycling services. The result is summarised in **Table 1**, which describes the mapping between different sensing profiles and particular sets of services.

**Table 1.** Mapping sensing profiles to specific services

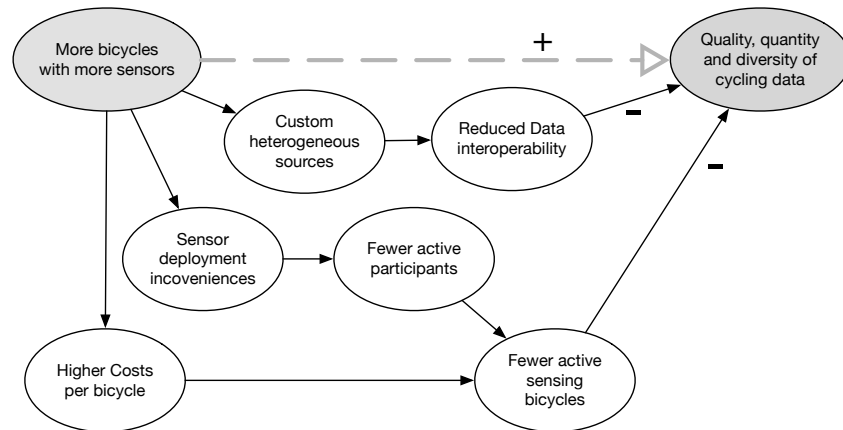
Profile	Service
Position	Cycling Maps, Network planning, Volume estimations, Quantified self, Cycling network hierarchy, A->B patterns, Travel times
Motion	Ridding Conditions, Road surface classification, Automated cyclist profile, Bump detection, Riding flow metrics
Environment	Local pollution levels, Traffic level estimations
Surrounding	Safety index based on free surrounding Space
Rider	Physical condition, Riding stress, Exertion
Video	Route annotation. Data interpretation.
Sound	Road surface classification, Surrounding traffic, Sound level
Proximity	Bicycle count, symbolic positioning

By providing a perspective of the mapping between sensing system and key data-centric services this table should help developers to consider which profiles to include to obtain specific services, or, given a set of profiles, to understand what can realistically be achieved with the available data.

### 4.3 Collective optimisation of cycling sensors

The final layer in the analysis is to move beyond the individual bicycle and consider how to optimise an entire cycling sensing ecosystem. Ideally, once we define the target services and identify the sensor profiles that need to be included to support those services, we would like to have all the bicycles equipped with all sensors for all those profiles. In reality, this assumption is challenged by multiple effects, which occur

across the various parts of the system and are common to any sensing bicycle design. These effects represent a design trade-off between the quantity and quality of the data that needs to be collected and the costs and hassles involved in that data collection process, as represented in the conceptual map in **Fig. 6**.



**Fig. 6.** Key design trade-offs

The key design decision regards the sensing profiles to deploy on each bicycle. In our work with the prototype bicycle, we found, from our many integration attempts, that the number and range of sensors deployed on a particular bicycle seems to have a major impact across a number of other key design goals, particularly cost, convenience and heterogeneity. The first two, cost and convenience, will significantly impact the number of bicycles that one can expect to be involved in the sensing system. The impact of cost is self-explanatory. The impact of sensor diversity is mainly associated with the potentially strong increase in system complexity. There will be more complex integration issues, and especially, there will be more and more specific deployment requirements, leading to all sorts of hindrances, such as specific pre-ride operations, additional cables on the bicycle, special configurations, regular maintenance or sensors placed at inconvenient positions in the bicycle. These will impact negatively on the cycling experience and would probably be enough to drive cyclists away from any form of sensing.

While previous work has often assumed sensing bicycles as part of a uniform fleet equipped with a specific set of sensors, this analysis led us to the opposite conclusion. Future smart cycling ecosystems are much more likely to be composed by a very heterogeneous set of bicycles with various combinations of sensing profiles. This will be even more the case, whenever we consider that a broad range of services, and consequently a broad range of sensing profiles, should be available. This is also much more aligned with an ecosystem view which naturally assumes the existence of many stakeholders with very different needs and, consequently, with different priorities on what sensors to include in their bicycles. Ultimately, each individual cyclist will also be making his or her own decisions on what sensing add-ons to use when cycling.

Therefore, rather than pushing for more uniform bicycles, we leverage heterogeneity as an opportunity to explore complementarities between large numbers of heterogeneous bicycles in an urban ecosystem. More specifically, we envision an ecosystem where different proportions of different types of sensing bicycles would be complementing each other to produce a balanced vision of the cycling reality and guarantee the necessary data while avoiding unnecessary redundancies. A larger number of low-cost sensing bicycles, with a basic sensing profile could collect large volumes of data, especially data that requires higher spatial and temporal coverage, e.g. traffic volume. These would be complemented with a few, more sophisticated, bicycles that could combine two or more sensing profiles and collect data that is updated only occasionally, such as route characterisation. Additionally, one or very few probe bicycles could be deployed for professional data collection regarding more critical data. These probe bicycles could include multiple sensing profiles, but its most distinctive property would be the professional nature of the data collection process, which could thus involve more rigorous processes and more credible data. The quantitative relation between specific sensing profiles and the proportion of bicycles where they should be included to produce the data needed by their target services is not addressed in this work, but becomes a relevant research topic for future work.

While bicycle heterogeneity may hold the key for sensing in urban cycling, it also raises interoperability issues that may negatively affect the value of the data generated by the whole system. In a context of heterogeneity, there are many design possibilities regarding the specific sensor model, the way it is deployed, its position on the bicycle or even the bicycle itself, that may affect the data generated. These challenges may go beyond the usual issues of standard data representations. They may also involve the explicit description of the data collection circumstances, which in some cases might be crucial for the correct interpretation of data. If these relevant elements are not considered, the ability to aggregate data from multiple sources and consequently the value that could be obtained from combining heterogeneous bicycles, can be compromised. This suggests the need for a reference design for sensing bicycles that allows many different bicycles to be used together as part of one common ecosystem, while explicitly dealing with any relevant deployment sensitivities.

## **5 Conclusions and Future work**

Sensors and bicycles are a powerful combination that is attracting significant attention from research and industry perspectives. While many technical approaches have already been explored, there is a gap on how to approach this topic from an ecosystem perspective where many stakeholders can be involved and data is being collected by very diverse sensing bicycles all of which with their own technology and sensing concepts. The novelty of this contribution is in the broader and more generic perspective of this study, which aimed to pursue a more explicit and thorough analysis of the various dimensions shaping the key trade-offs for cycling sensing systems.

## 5.1 Future work

This is an ongoing work, where we are plan to develop further research to provide more thorough indications on how to design, describe and combine very diverse sets of sensing bicycles. In particular, we plan to explore the quantification of the size and relative proportion of the samples needed to offer the services associated with each sensing profile. This knowledge is crucial to support the planning of smart fleets of heterogeneous bicycles where the combination of the various profiles is optimized according to the target services. We also plan to evolve the definition of a reference design for sensing bicycles. This should offer a common reference for describing the sensing affordances of a particular bicycle design, from the simplest ones to a full-fledged probe bicycle for professional data collection. This should allow those sensing possibilities to be combined in many different ways, while allowing their results to be analysed consistently. With this model, it should become possible to relate multiple sensing bicycle designs as variants of a common model and it should be easier to integrate data generated from different instances into shared datasets.

## Acknowledgements

This work is supported by: European Structural and Investment Funds in the FEDER component, through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) [Project n° 039334; Funding Reference: POCI-01-0247-FEDER-039334]

## References

1. Stratta, P.: Towards a Smarter Cycling. On the brink of a Smart (R)evolution, <https://ecf.com/what-we-do/cycling-new-technologies/towards-smarter-cycling>.
2. Nikolaeva, A., te Brömmelstroet, M., Raven, R., Ranson, J.: Smart cycling futures: Charting a new terrain and moving towards a research agenda. *J. Transp. Geogr.* 79, 102486 (2019).
3. Lee, J., Leem, Y.T., Lee, S.H.: Categorizing U-Bike Service and Assessing Its Adoptability Under It-. 12th World Conf. Transp. Res. 1–10 (2010).
4. Elen, B., Peters, J., van Poppel, M., Bleux, N., Theunis, J., Reggente, M., Standaert, A.: The Aeroflex: A bicycle for mobile air quality measurements. *Sensors (Switzerland)*. 13, 221–240 (2013).
5. Outram, C., Ratti, C., Bitterman, A.: The Copenhagen Wheel: An innovative electric bicycle system that harnesses the power of real-time information and crowd sourcing. Mayor’s Summit at the COP15 United Nations Climate Change Conference (2009).
6. SODAQ: SODAQ- Sniffer Bike, <https://sodaq.com/projects/sniffer-bike/>.
7. Wirtschaft, W. und S. des V. an der H.K.– T. und: Feintuning für die Radverkehrsplanung, <https://www.hs-karlsruhe.de/presse/feintuning-fuer-die-radverkehrsplanung/>.
8. Kiefer, C., Behrendt, F.: Smart e-bike monitoring system: Real-time open source and open hardware GPS assistance and sensor data for electrically-assisted bicycles. *IET Intell. Transp. Syst.* 10, 79–88 (2016).



9. Eisenman, S.B., Miluzzo, E., Lane, N.D., Peterson, R.A., Ahn, G.S., Campbell, A.T.: BikeNet: A mobile sensing system for cyclist experience mapping. *ACM Trans. Sens. Networks*. 6, (2009).
10. See.Sense: See.Sense, <https://seesense.cc/>.
11. Makarova, I. V., Boyko, A.D., Shubenkova, K.A.: Cycling intellectualization in Smart Cities. *EAI Endorsed Trans. Smart Cities*. 2, 153498 (2017).
12. Shen, S., Wei, Z.Q., Sun, L.J., Su, Y.Q., Wang, R.C., Jiang, H.M.: The shared bicycle and its network—internet of shared bicycle (IoSB): A review and survey. *Sensors (Switzerland)*. 18, 1–24 (2018).
13. Ehrgott, M., Wang, J.Y.T., Raith, A., Van Houtte, C.: A bi-objective cyclist route choice model. *Transp. Res. Part A Policy Pract.* 46, 652–663 (2012).
14. Winters, M., Davidson, G., Kao, D., Teschke, K.: Motivators and deterrents of bicycling: Comparing influences on decisions to ride. *Transportation (Amst)*. (2011).
15. Segadilha, A.B.P., Sanches, S. da P.: Identification of Factors that Influence Cyclists Route Choice. *Procedia - Soc. Behav. Sci.* 160, 372–380 (2014).
16. Félix, R.: Barriers and motivators to bicycle in low cycling maturity cities: Lisbon case study. 259 (2019).
17. Broach, J., Dill, J., Gliebe, J.: Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transp. Res. Part A Policy Pract.* 46, 1730–1740 (2012).
18. Su, J.G., Winters, M., Nunes, M., Brauer, M.: Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada. *Transp. Res. Part A Policy Pract.* 44, 495–505 (2010).
19. Silva, N., Soares, J., Shah, V., Santos, M.Y., Rodrigues, H.: Anomaly Detection in Roads with a Data Mining Approach. *Procedia Comput. Sci.* 121, 415–422 (2017).
20. Wickham, H.: Tidy Data. *J. Stat. Softw.* 59, (2014).
21. Liang, S., Huang, C.Y., Khalafbeigi, T.: OGC SensorThings API Part 1: Sensing. *Open Geospatial Consortium. Implement. Stand.* 1–105 (2016).
22. Elen, B., Peters, J., van Poppel, M., Bleux, N., Theunis, J., Reggente, M., Standaert, A.: The Aeroflex: A bicycle for mobile air quality measurements. *Sensors (Switzerland)*. 13, 221–240 (2013).
23. Liu, X., Xiang, C., Li, B., Jiang, A.: Collaborative bicycle sensing for air pollution on roadway. *Proc. - 2015 IEEE 12th Int. Conf. Ubiquitous Intell. Comput. 2015 IEEE 12th Int. Conf. Adv. Trust. Comput. 2015 IEEE 15th Int. Conf. Scalable Comput. Commun.* 20, 316–319 (2016).
24. Corno, F., Montanaro, T., Migliore, C., Castrogiovanni, P.: SmartBike: An IoT crowd sensing platform for monitoring city air pollution. *Int. J. Electr. Comput. Eng.* 7, 3602–3612 (2017).
25. Boot, M.: A digital future for cycling: the role of data and ITS in encouraging more cyclists, <https://www.intelligenttransport.com/transport-articles/78216/digital-future-cycling-data-its-encouraging-cyclists/>.