NIST Technical Note
Publication Identifier

A Taxonomy for Key Performance Indicators in the Public and Infrastructure Sectors

Final Draft

Cheyney M. O'Fallon, PhD
W. Michael Dunaway, PhD

This publication is available free of charge from:
https://doi.org/10.6028/NIST.XXX.XXXX

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Technical Note
Publication Identifier

A Taxonomy for Key Performance Indicators in the Public and Infrastructure Sectors

Final Draft

Cheyney M. O'Fallon, PhD
W. Michael Dunaway, PhD

Smart Connected Systems Division
Communications Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.XXX.XXXX

December 2025

U.S. Department of Commerce
Howard Lutnick, Secretary

National Institute of Standards and Technology
Craig Burkhardt, Acting Under Secretary of Commerce for Standards and Technology and Acting NIST Director

Publication Identifier
December 2025

Publication History
Approved by the NIST Editorial Review Board on 2025-11-14

How to Cite this NIST Technical Series Publication
O'Fallon C, Dunaway M. 2025. A Taxonomy for Key Performance Indicators in the Pubic and Infrastructure Sectors. National Institute of Standards and Technology, Gaithersburg, MD, NIST Series (Series Abbreviation) Publication ID. https://doi.org/10.6028/NIST.XXX.XXXX

Author ORCID iDs
Cheyney O'Fallon: 0000-0002-5931-4173
Michael Dunaway: 0000-0002-1535-1829

Contact Information
michael.dunaway@nist.gov

Abstract

This document presents a taxonomy for key performance indicators (KPIs) tailored to the public and infrastructure sectors. An approach to classify KPIs is outlined and then followed, aiming to enhance the accuracy and utility of performance measurement in these sectors. Through an iterative development process, this taxonomy addresses the structural determinants of transaction costs that impact the accessibility and analytical utility of KPIs. The document serves as a resource for economists, social scientists, and decision-makers, facilitating informed decisions and improvements in public and infrastructure services. This taxonomy not only aids in navigating the complexities of KPI application but also promotes standardization and effective communication across government organizations and communities. The development and implications of this taxonomy are demonstrated using actual public safety and resilience focused data from Portland, OR. The taxonomy has the potential to be adapted and applied to different contexts within the public and infrastructure sectors to improve performance measurement activities.

Keywords

Infrastructure; Key Performance Indicator; Public Safety; Resilience; Taxonomy; Transaction Costs.

Table of Contents

List of Tables

List of Figures

1. Introduction

This document presents a taxonomy artifact and the taxonomy development process through which it was produced. This taxonomy constitutes a means of organizing a form of public performance data known as key performance indicators (KPIs). Specifically, this document is concerned with developing a classification scheme of public and infrastructure sector KPIs for the domains of public safety and community resilience. This represents an initial effort toward developing a broader set of KPIs relevant to other sectors. The taxonomy development process employed follows steps that are well-documented in the scholarly literature on information science and design science research. Development proceeds with the empirical analyst or researcher in mind, whether professional or aspiring. The taxonomy offered in this document is intended as an aid to analysts as they make judgment calls regarding which data resources to use and how in their work. The taxonomy also offers a simple approach for visualizing the level of institutional development and documentation associated with KPIs across government organizations, operating domains, and communities.

The meta-characteristic guiding taxonomy development concerns the ***structural determinants of transaction costs that shape the accessibility, reusability, and analytical utility of KPIs in public and infrastructure sectors***. Intended end users of this taxonomy include applied economists, other social scientists, data scientists, agency decision-makers and leaders, civic organizations and members of the public that want to use public data to inform decisions and improve the services and benefits enjoyed by their stakeholders. This taxonomy should help explain a crucial aspect of the cost structures underlying increasingly complex and consequential performance measurement systems and models driving policy choices.

As a test case, this taxonomy development process was applied to actual data from Portland, OR, and through several iterations, arrived at a concise and fit for purpose taxonomy for public safety and resilience KPIs. After demonstrating that the taxonomy satisfies sets of objective and subjective ending conditions, discussion turns to how the taxonomy may be tested and improved through its application to data from other settings and domains. Methods for visualization are also discussed.

2. Background

This section reviews literature on the development and implementation of taxonomies deemed pertinent to KPIs for the public and infrastructure sectors. Additionally, this section offers a brief discussion of key concepts, standards, and frameworks relevant to the development of a KPI taxonomy. The intended audience includes researchers and decision-makers interested in the measurement of current performance and its evolution as new technologies and systems are adopted to provide beneficial services.

The efficacy of policy formation improves with high-quality measurement science that ensures the accurate accounting of benefits delivered to stakeholder groups. However, while measurement strategies are incorporated into the official processes of decision makers around the world, neither metrics nor the measures used to produce them, enjoy the high level of standardization required to facilitate benchmarking across physical and temporal settings. As discussed in [1], colloquial use of the terms *measure* and *metric* can demonstrate a lack of consensus on the fundamental definitions of, and distinction between, these terms. The success or failure of public sector technology investments is generally not evaluated using scientific, unbiased evidence of costs and system improvements. In the absence of such indicators of success or failure, community leaders must make sense with the limited and potentially misleading tools of opinion and public sentiment.

This taxonomy development process is based on the concepts and context presented in [2], the *Smart cities and communities: A key performance indicators framework*, developed by the National Institute of Standards and Technology (NIST). The taxonomy development and standards literature reviewed and selected for consideration is evaluated for its alignment with that document [2]. The approach to taxonomy development is informed by the motivating question of which KPIs are sufficiently mature for adoption that they can and should be standardized for the betterment of people's lives. Furthermore, this article develops a taxonomy that can organize KPIs along dimensions that factor into a qualitative and quantitative answer to this motivating question. Those KPIs that exhibit bloated transaction cost structures are less prepared for adoption as the costs of adjusting them to new settings and organizations are likely high, dissipating value propositions if not blocking the realization of net benefits completely. The best candidates for adoption have transaction cost structures that are relatively efficient and balance the gains from insight against the costs of measurement.

2.1. Motivation

The fundamental purpose of developing the taxonomy of holistic key performance indicators is to improve the quality and consistency of performance measurement and estimation processes informing decision makers in public and infrastructure sectors. These processes may evolve out of informal discussions employing a single KPI. From these humble beginnings, an organization might eventually innovate decision support tools founded on digital twins that are synchronized frequently with live data from public and private systems. Coral Gables, FL is an example of a city currently operating a digital twin accessible to the public.[1] The value of these performance measurement processes will reflect the quality of their data inputs and the procedures for drawing inference from their combination. Discussion briefly touches on some general sources of complexity that threaten to reduce the value of models and analyses composed of myriad data resources including KPIs. Ultimately, a taxonomy, well-constructed, can be a tool for managing this complexity.

A raft of new and improved technologies and systems are marketed daily to community leaders who are caught between the dual obligations of solving problems affecting their constituents and being good stewards of public resources. Marketers have a long tradition, informed by their own business interests, of making their offerings difficult to value and compare objectively before deployment [3, 4]. Once system deployment begins, the details and differences of implementation ensure that objective evaluation and management of technology adoption and operation remain a persistent challenge.

The complexity and interaction of the systems driving policy complicates the process of understanding and communicating ground truth to leadership and the public alike. Many publicly reported statistics become conceptual anchors in popular discussion as well as the competing rhetoric of interest groups. However, metrics of widely differing levels of quality are harnessed together in service of sense-making and persuasive exercises all the time. Municipalities can expect the integration of new technologies in their social and infrastructure networks to be a persistent part of planning and operations supporting cities' role as "open-ended social reactors" [5]. Indeed, the combinatorial innovation [6] observed in the domains of system analytics and economic modeling, constitute examples of the integration of numerous data pipelines into distinct documents and dashboards. The effective composition of performance measurement artifacts, such as KPIs, into higher level products and services requires systematic approaches to understanding the qualities of data inputs and their implications.

The statistics that anchor public debate likely deserve deeper study by those who employ them as a matter of civic due diligence. That the level of scrutiny remains suboptimal does not mean that the problems of supporting an informed decision-making process are intractable. An immediate step towards improving the quality of performance measurement is to develop a taxonomy that offers a potential ordering of KPIs along a set of dimensions which capture the costs of conducting good empirical work. Such a taxonomy can help expand the set of people who can interact with and wield public data for better outcomes in their respective communities. Seasoned analysts can develop sound professional judgement around which data sources to use or avoid. However, developing such judgement requires experience from time on task. Not all stakeholders with interests in wrangling and wielding data can be expected to have ample amounts of relevant experience. In this absence, the taxonomy can help emerging professionals to organize their thinking around the attributes necessary to justify a KPI's use.

[1] For a current example of a municipal digital twin, please see https://coral-gables-smart-city-hub-2-cggis.hub.arcgis.com/.

Imposing a strict, prescriptive set of requirements and best practices for performance management of public and infrastructure sector technology adoption and operation holds the potential to create as much friction as value. Nevertheless, aids to facilitate common understanding and discussion are needed to anchor community decision-making on a solid foundation of empirical analysis and measurement. The taxonomy developed in this manuscript seeks to reduce the barriers and frictions that disrupt and dissipate the opportunities for communities to learn from each other and come to defensible conclusions regarding their own performance when adopting technologies, and operating infrastructure systems. The resulting taxonomy endeavors to help infrastructure owners and operators understand how to better present their analyses and value judgements with respect to the KPIs they produce and consume.

2.2. Concepts and Definitions

This section contains a discussion of several key concepts relating to the development and use of this taxonomy, as well as definitions of relevant terms. What follows is intentionally more illustrative than exhaustive. Detailed surveys of related literature are cited where appropriate so that the reader may find additional resources as needed.

2.2.1. Key Performance Indicators

The taxonomy developed here seeks to organize a representative subset of key performance indicators (KPIs) and illustrate an approach to identifying and using KPIs as a basis for decision-making. Complicating matters is the fact that many organizations report KPIs, but few will agree on a precise definition of the phenomena. One view, expressed in [7], outlines seven characteristics of KPIs, stating that they are: nonfinancial measures; measured frequently; acted on by an organization's executive team; of high impact; conducive to appropriate actions; clear regarding the actions required by staff; and provide linkages between responsibility for action and the teams charged with their execution. Even a cursory investigation will produce many examples of objects labeled KPIs that do not have all these characteristics.

Reasonable people may disagree regarding which if any of these characteristics are strictly necessary for the purpose of performance measurement. However, definitional consensus is possible to find across settings. For a global survey of 3200 senior executives on how their organizations use KPIs [8], the MIT Sloane Management Review defined KPIs as "the quantifiable measures an organization uses to determine how well it meets its declared operational and strategic goals." Recent NIST research [9] uses the relatively concise definition of a key performance indicator as "[a] measure of progress toward intended results".

Any taxonomy of public and infrastructure sector KPIs, if it is to be broadly applicable, must be able to handle differences in how organizations define the concept. The phrase public and infrastructure sectors is intended to encompass municipalities as well as other infrastructure owners and operators. In the absence of a single uniformly preferred definition of a KPI, this taxonomy seeks to accommodate a wide set of the measures offered by organizations under that label. Furthermore, variability in KPI definition, development, and deployment provides motivation for the next turn in discussion.

2.2.2. Transaction Costs

Measurement science with respect to transaction costs has long encountered challenges of a fundamental variety. Härting and Sprengel [10] note that while “it is not primarily possible to carry out a measurement, it is possible to generate a basis for decision-making with regard to efficiency”. This manuscript has the express purpose of producing a process and taxonomy to contribute to that basis for efficient decision-making.

Before diving into transaction costs, it is necessary to briefly explain what constitutes a transaction for the purpose of this manuscript. Williamson explains that a transaction occurs “when a good or service is transferred across a technologically separable interface” [11]. This interface is the point where the output of one distinct phase of production is clearly separated from the next. In theory, if this separation is sufficiently clean, the output of the prior phase of production could be routed to one of many potential next phases. The next phase may involve the next robot along a conveyor belt within a given firm’s production line, or it may be the further machining or treatment of a component by another firm entirely. In the example concerning KPIs, one might think of the posting of KPI data to an organization’s website for download and use by any interested party as a transaction.

Specifically, this taxonomy focuses on the dimensions of cost associated with a KPI’s acquisition and application to empirical questions. The ease of downloading such data generally masks the potentially costly work of the analyst to reduce uncertainty with respect to the conceptual definition, collection methodology, and product quality of the data that will be incorporated into scientific research. The costs of working with KPI data thus reflect the effects of design and operating decisions made upstream along the information pathways that feed into the KPI. Should a researcher publish their findings employing such KPIs as a technical report or research article, that publication could constitute another transaction. For now, it should be clear that the costs of these different types of transactions may vary considerably.

A brief discussion on transaction costs, property rights, and public goods is warranted here for clarity. Allen uses the concise definition that “transaction costs are the resources used to establish and maintain property rights” [12]. Simply put, property rights are “the rights of individuals to the use of resources” [13]. Public and infrastructure sector KPIs are public goods, meaning that they are both non-excludable and non-rivalrous. Non-excludable, in an economic sense, means that the prevention of the good’s use by people is either impossible or prohibitively costly. Non-rivalrous means that the use of the good or service by one consumer does not diminish its availability to others. Once a KPI is posted online for popular consumption, it is clearly non-excludable and non-rivalrous. A core challenge faced by a taxonomy that attempts to understand the dimensions and characteristics that determine the qualitative transaction cost structures of KPIs is that the property rights driving these costs are often weak and undefined because of the inability and inappropriateness of claims to exclusive ownership. The uncertainty arising from public good property rights contributes additional costs to using

KPIs to guide investments and policy choices to improve system performance and related value propositions.

Williamson [14] traces the antecedents of transaction cost economics to a collection of eminent economists [15-17]. Surveys of the considerable empirical evidence to support transaction cost economic theory can be found in [18] and [19]. Transaction cost theory also has promising applications with respect to platform markets, artificial intelligence, and efforts to accomplish non-pecuniary goals [20].

No shortage of nuance may be found in attempts to categorize the contributors to transaction costs, but the discussion here will begin with a simpler conceptualization. Matters of uncertainty, specificity, and frequency are generally understood to influence transaction costs [10]. Higher levels of uncertainty relative to human or external factors tends to increase transaction costs. The same is true of high levels of asset specificity. An example of high asset specificity might be a proprietary sensor network and software suite used by an organization to understand infrastructure utilization. Key to the term specificity in this context is the presence of limited alternative options for the product of an irreversible investment in monitoring systems. Efforts to improve opportunities for KPI reuse, see [2], offer a clear avenue for reducing asset specificity and per-unit transaction costs. Higher frequency transactions present opportunities for cost-reducing learning effects and the establishment of efficient contracting mechanisms over time. As discussed below, metadata on the frequency of publication of KPI data readily becomes a dimension of the taxonomy. Evaluating metadata for what it can tell us about the transaction costs of working with KPIs is central to the taxonomy development efforts discussed here.

Transaction costs arise in both the activities under measurement and the performance measurement processes. This duality can foster the recursive problem of cost escalation due to the conduct of measurement activities undertaken for the express purpose of reducing operational transaction costs.

The origins of transaction costs associated with KPIs are myriad. Organizations publishing KPIs must first engage in costly research and stakeholder engagement, as well as testing and validation efforts. The production of documentation entails costs for the original KPI developer, but poor documentation ensures higher transaction costs of adoption by peer organizations operating in other settings. Incomplete provision of implementation details is a common source of additional transaction costs associated with translating the solutions of peers to new adopters. A lack of widely adopted standards for KPIs also increases transaction costs.

2.2.3. Replication, Reuse, and Salience

Ultimately, this taxonomy is intended to help with the identification, reduction, and avoidance of KPI transaction costs so that organizations may more readily replicate empirical findings, reuse metrics across multiple applications, and increase the salience of performance metrics to the decision-making process.

Replication in this context refers to the ability to reproduce analytical findings about organizational performance across different timeframes, organizational units, or institutional contexts. Successful replication requires clear documentation of methods and consistent data collection practices to verify results and build cumulative knowledge.

The term reuse is defined here in the same manner as it is in the holistic KPI (H-KPI) Framework.

> A 'use' is defined as the use of a specified data type and source to enable a defined infrastructure service or community benefit. A 'reuse' is the use of that same data for additional infrastructure services or community benefits. [2]

Salience, adapted from network theory [21], measures the extent to which a link participates in the shortest paths between pairs of nodes in a network. In performance measurement systems, highly salient metrics are those that frequently provide critical pathways between analytical questions and meaningful answers across multiple organizational contexts. Just as salient network links facilitate efficient paths between many node pairs, salient performance measures serve as key connection points enabling both replication and reuse.

2.2.4. Metadata

Metadata, or "data about data," provides structured information that describes the essential attributes and context of data resources [22]. In the information systems literature, metadata serves multiple critical functions including enabling data discovery, ensuring interoperability between systems, supporting digital object management, and facilitating long-term preservation [22]. The National Information Standards Organization (NISO) distinguishes three fundamental types of metadata: descriptive metadata that enables resource identification and discovery, administrative metadata that supports resource management, and structural metadata that documents relationships between components [22]. Markup languages may also be considered a distinct type as they combine content and metadata in a single document.

For developing taxonomies of real-world phenomena, metadata provides a strong point of departure for identifying and analyzing the key characteristics that differentiate objects of interest. Taxonomies may be derived either conceptually through deductive reasoning or empirically through examination of actual objects [23]. When working with published data resources like KPIs the metadata associated with these indicators offers a rich empirical basis for taxonomy development. The metadata captures important qualitative attributes such as definition parameters, measurement methodologies, data quality characteristics, and usage conditions that can inform the identification of taxonomy dimensions [24].

Following the method proposed by [23] and refined by [25], metadata can be systematically analyzed to develop taxonomy dimensions that align with the meta-characteristic of interest (See Sec. 3.4). This approach combines the conceptual strength of having a clear theoretical foundation (the meta-characteristic) with the empirical rigor of deriving dimensions from actual object attributes documented in metadata. The resulting taxonomy benefits from being grounded in observable characteristics while maintaining focus on its intended purpose and users [25]. For taxonomies aimed at understanding transaction costs in empirical research, as in the case of public and infrastructure sector KPIs, the metadata provides insight into the practical determinants of these costs through its documentation of data accessibility, quality, and complexity characteristics.

A brief digression into the value of metadata to system design follows with an instructive example. Development of the Metacat [26] system faced three key challenges in managing ecological data that parallel this taxonomy development effort: data heterogeneity, data dispersion, and local control requirements. The Metacat team developed a flexible metadata framework that could accommodate diverse data structures while maintaining consistent searchability and interoperability across distributed research sites. This approach resonates with KPI taxonomy development, where heterogeneous indicators are dispersed across multiple agencies and jurisdictions, each maintaining local control over their measurement frameworks. Just as Metacat used structured metadata to bridge these divides in ecological research, taxonomy development can leverage KPI metadata as a standardizing layer that helps identify common dimensions across otherwise disparate performance measurement systems. The success of Metacat in enabling cross-site analysis while preserving local autonomy demonstrates the value of using metadata as an intermediary structure when developing taxonomies of distributed real-world objects.

2.2.5. Taxonomy

Many definitions of the term *taxonomy* exist and can inform the development process. Nickerson et al. notes in [23] that taxonomy is a form of classification and is therefore what Wand et al. describes in [27] as "a fundamental mechanism for organizing knowledge". Alternatively, "taxonomy is a list of words and phrases, most often organized into a hierarchy. It is used to classify, categorize or tag just about anything" [28]. More formally, Lippell defines a taxonomy in [28] as "[s]trictly, a *Controlled vocabulary* arranged in a *Hierarchy* without other fields and relationships." ANSI/NISO Z39.19 is a leading standard for monolingual controlled vocabularies [29]. A discussion of the development and refinement of the ANSI/NISO Z39.19 standard can be found in [30].

Another prominent taxonomy standard, ISO 25964 [31], was designed to facilitate information retrieval and indexing in settings that extend beyond the boundaries a single organization, especially within the context of the internet as it evolved towards greater prevalence of machine-readable content. This overarching design consideration is relevant to these efforts in that the taxonomy is intended to be of use to empirical researchers and the decision-makers they support. Improving information retrieval and comprehension by external researchers is crucial to the better integration of KPIs into the analyses informing decision-making processes and public discourse.

2.2.6. Ontology

While the development of an ontology is not within the scope of this document, future work may expand on this taxonomy in that direction. Gruber defines ontologies in [32] as specifications of conceptualizations, and outlines five design criteria for ontologies intended to facilitate knowledge sharing: clarity, coherence, extendibility, minimal encoding bias, and minimal ontological commitment. The Simple Knowledge Organization System (SKOS), described by a World Wide Web Consortium (W3C) standard, is an ontology for taxonomies and thesauri [33].

2.2.7. Frameworks

Any taxonomy of public and infrastructure sector KPIs should align with or at least avoid conflict with a variety of frameworks produced by NIST and other organizations active in the relevant standards space. There is an obvious need for alignment with the NIST Smart Cities H-KPI framework, [2], the precursor and companion document to this Technical Note. For this reason, discussion turns to the matter of what constitutes frameworks and why they are useful.

Schwarz et al. [34] offers the summary that “a framework includes the assumptions, concepts, values, and practices of prior literature to aid practitioners or academics to interpret the world around them”. Their study develops a list of abstracted purposes for framework articles from a review of examples in the information systems, social sciences, and related management fields [34]. A first and most obvious set of abstract purposes for framework and review articles involves the integration of previous studies across disciplines or theoretical perspectives. These articles may also have the purpose of helping with the interpretation of data through promoting understanding or the systematic collection, organization, and analysis of findings. Other purposes for frameworks and reviews can include the collection of data for methodological assessment or the search for structure underlying a field of literature in hopes of distilling the assumptions at the core of a research stream. A framework may also focus attention on a particular direction the authors believe a research stream should take for maximum scientific or societal benefit. Perhaps most relevant to this work in developing a taxonomy of KPIs for the public and infrastructure sectors, frameworks can serve to improve understanding with respect to how theoretical concepts in a field are related to policy, operations, and service offerings. Frameworks may help to synthesize a body of relevant literature in a way that makes relatively esoteric research streams actionable to analysts in other institutes and industry. Finally, frameworks and review documents may sketch out the legitimate frontier of a research field or organize the work that has already been done.

2.2.8. Hierarchies and Facets

Taxonomies may be hierarchical or faceted in structure, and elements of both structures may be present, as well. Where each thing to be organized in the taxonomy can be conceived as a broader or narrower example of other things, the hierarchical approach is sensible. With key performance indicators, an instance of a measure may not adhere to a strict hierarchy. Consequently, the taxonomy advanced in this manuscript is faceted. Denton defines a faceted classification as “a set of mutually exclusive and jointly exhaustive categories, each made by isolating one perspective on the items (a facet), that combine to completely describe all the objects in question” [35]. As will be discussed below, facets may be drawn directly from the metadata accompanying key performance indicators. The choice of a faceted approach to the KPI taxonomy is further supported by the literature on municipal systems. One observer notes: “[t]he greatest difficulties to any scientific approach to cities have resulted from their many interdependent facets, as social, economic, infrastructural, and spatial complex systems that exist in similar but changing forms over a huge range of scales" [5]. While the interdependence of facets may make them difficult to isolate into distinct dimensions for this taxonomy, an iterative approach and real data are used to ensure that each identified facet contributes additional explanatory power for the end user.

2.2.9. Citation Order

The *citation order* is the order in which facets are combined to describe an item in a classification system and influences how someone searches for and retrieves information. The citation order advanced in this manuscript reflects a mixture of common theoretical arguments found in the scholarly literature and consideration of end user needs. Broughton [36] points to three main theoretical arguments in support of the citation order for the 13 facets advanced by [37]: the order flows from concrete to abstract; facets depend on preceding facets; and compound subjects are gathered into useful groupings.[2] This citation order echoes that of the roughly contemporary classification scheme developed by [38] and known by the abbreviation PMEST: personality, matter, energy, space, and time. The citation order advanced below for the public and infrastructure sector KPI taxonomy seeks alignment with these existing formulations.

2.3. Challenges and Considerations

One challenge facing the development of a KPI taxonomy for use by many different communities is that each one has a unique IT and governance structure and therefore the requirements for a taxonomy management tool will vary across users. A consequence of this reality is that a search for common denominators must precede the selection of the tools used to develop the controlled language through which communities might collaborate to achieve a better understanding and implementation of KPIs. See [28], page 26 on selecting a taxonomy tool.

Considering the variation in community systems, the faceted taxonomy developed in this manuscript uses a spreadsheet rather than a more advanced or specific tool. All organization personnel and public stakeholders should have access to such a software tool. To avoid introducing an unnecessary barrier to exploration and implementation at an early stage, the selection of a bespoke or purpose-built taxonomy tool is left to future work. Follow-on iterations which incorporate additional objects (KPIs) may introduce additional complexity best managed with more advanced tools.

Factors that are key to the performance of a community or infrastructure system evolve over time with advances in technology, shifts in demographics, and changes to political and economic institutions. Access to fresh horses at regular intervals along trade routes was once key to the performance of communications networks between population centers. The density of cell towers or prevalence of underground fiber optic cable are now far more important as

[2] The citation order from the Classification Research Group is as follows: thing – kind – part – property – material – process – operation – patient – product – by-product – agent – space – time.

indicators of the state of infrastructure. A taxonomy of key performance indicators should seek to create lasting value in an evolving setting by identifying the facets of performance measurement that persist through technological and social change. An example of such a facet, adopted from Portland, OR is *MeasureType,* which sorts measures into the categories of workload, efficiency, output, and outcomes (see Sec. 4.3.1.1 and Appendix C; Table 8).

One important facet of a key performance indicator is how contested or subjective it is. Even objective measures, such as the cost of a business process, may experience changing definitions of what is considered “in scope”. The greater the complexity of a given measure, the greater the opportunity to design the definition to produce favorable results. As will be discussed in the succeeding section, metadata for Portland, OR key performance measures include a categorical variable for the reliability of the process producing a given measure. Though what differentiates medium and low categories for reliability in measure production is difficult to determine, suggesting a role for personal or organizational judgement.

In addition to addressing variety borne of changing social and physical phenomena, a taxonomy of performance measures can’t avoid the reality of strategic interactions and countervailing interests of different stakeholder groups or government bureaus. Performance indicators reflect the lives and work of real people and are often chosen by those individuals and the interests they describe. No one is immune from the urge to put their best foot forward, including through the data that they publish about themselves, their colleagues, and their organizations. This aspect of measures and attempts to game them, both conscious and subconscious, has a long history in social sciences. One seminal formulation of this concept is attributed to Campbell who notes in [39] that, “[t]he more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor”. The choice of measures published with the label of “key to performance” is destined to be a matter contested by the factions they describe.

Many taxonomies are developed in an ad-hoc manner and subject to insufficient testing and evaluation [40]. The taxonomy developed in this manuscript does not pretend to escape this challenge completely. Instead, the audience is invited to test the taxonomy for themselves and evaluate whether they gain greater clarity or understanding of the opportunity to take KPIs that were developed by peer communities or adjacent infrastructure systems and adopt them for their own applications. As the approach discussed here is fundamentally empirical, moving from actual KPIs that are currently implemented, it is up to the reader to discern whether each facet creates value for their own understanding and applications or could be excluded from implementation. The facets that are helpful to the largest set of communities can thus form the common denominators of classification for performance measurement in public and infrastructure sectors. The theoretical discussion contained in this manuscript, including in this section, is present to offer a window into the conceptual challenges and concerns, and not an enumeration of requirements to be imposed on the variety of potential end users.

3. Methodology

The approach taken to develop the taxonomy in this manuscript represents an adaptation of the methods described in [23] and [41]. The research contained in those two articles formulates the problem of defining a procedure for taxonomy development; offers a working definition and parsimonious notation for taxonomies; and outlines some attributes of useful taxonomies in the information systems domain. The desired attributes include that the taxonomy be concise, robust, comprehensive, extendible, and explanatory. The chosen methodology also uses the expansions developed in [25], which begins with three steps that help to refine the taxonomy as a solution to the underlying problems of complexity and disorder that emanate from unstructured or unknown phenomena. The first three steps, discussed in Sec. 3.1 to 3.3, are noted for grounding the taxonomy development methodology in the tradition of design science research (DSR) [42].

3.1. Phenomenon

The taxonomy development method begins by specifying the observed phenomena to be explained and organized. When confronting this task, one can appeal to constructs that describe the characteristics and relationships of perceived importance to the phenomena. One helpful definition clarifies that "*[c]onstructs* or concepts form the vocabulary of a domain" [43]. This vocabulary constitutes the initial interface between the researcher and the subject matter. From these constructs are drawn many of the dimensions and characteristics to populate this taxonomy, as well as the identities of potential end users and the roots of the value propositions which give the taxonomy purpose.

3.2. End Users

The next step in taxonomy development is to determine the intended user group. Here, the developer asks a pair of questions: who interacts with the phenomena under consideration, and who would benefit from a useful taxonomy? The user group may grow or shrink as the taxonomy changes through development. Understanding how the set of end users co-evolves with the scope and substance of the taxonomy may itself be a useful source of explanation concerning the specified phenomena.

3.3. Purpose

The third step in development is to specify the purpose of the taxonomy. The taxonomy should present a value proposition to the user group with the clarity to signal the importance of the underlying phenomena and the net benefits gained through the taxonomy's use. Depending on the phenomena, a taxonomy may serve the purpose of providing structure and mapping relationships between known constructs, or in the case of emergent phenomena, create value through identifying important new constructs [25]. The emergence of new constructs will generally also introduce new relationships of import to understanding a phenomenon.

3.4. Meta-characteristic

The fourth step, that of meta-characteristic determination, is arguably the most significant component of the development process. Kundisch et al. [25] describes the meta-characteristic “as the taxonomy’s angle on the phenomenon under consideration”, which “defines what is relevant for the specific taxonomy design and what is not. Consequently, all the following characteristics and dimensions of the taxonomy must relate to this meta-characteristic" [25]. The meta-characteristic can be adjusted or redefined after the initial iteration of taxonomy development, as the process is likely to reveal important constructs and relationships from beneath the surface of a domain. Indeed, the language used to describe the meta-characteristic evolved throughout the taxonomy development process as it was refined.

3.5. Ending Conditions

The iterative nature of the taxonomy design method followed here requires the formalization of ending conditions to inform the decision of when to stop development efforts. There are both objective and subjective conditions that the taxonomy must meet to be complete. Fundamentally, the objective ending conditions reflect the convergence of the taxonomy’s basic structure of dimensions and characteristics upon a stable form, and the subjective conditions point to the satisfaction of design goals intended to promote the usefulness of the taxonomy. For a collection of possible objective and subjective ending conditions, see Table 2 and Table 3 in [23]. Generally, the purpose, user group, and meta-characteristic can be used to help triangulate on what constitutes usefulness, when designing the subjective ending conditions. The strength of alignment among components of the taxonomy design process contributes to the coherence of the taxonomy produced. The subjective ending conditions may have thresholds for satisfaction that are unclear, vary by researcher, and change over time. This challenge can be met with clear qualitative discussion of the degree to which the subjective conditions are met.

4. Development of a Taxonomy for Public and Infrastructure Sector KPIs

This section details the development process for the public and infrastructure sector key performance indicator taxonomy. The top-level headings for Sec. 4.1-4.6 closely follow the language and structure of the outline presented in Fig.1 of [25]. These method headers are originally adapted from Peffers et al. [44].

4.1. Identify problem and motivate stakeholders

The observed phenomena are city-reported key performance indicators. The city-reported aspect is important as the interest is in classifying KPIs that are presently in use and readily accessible to the public. While linked through the moniker KPI, the objects classified with the taxonomy can and do differ immensely in context and content. The intended empirical approach to KPI classification informs the focus on extant measures rather than ideal archetypes of H-KPIs found in Sec. 7 of [2].

The primary target user groups include economists, social scientists, data scientists, and to a lesser extent, the decision-makers who rely on these empirical analysts. Many public data series suffer from a lack of explainability around important matters of origin, context, and content. Professional data wranglers and analysts alike must spend non-trivial amounts of time to develop confidence in the data they employ. Students, early career data professionals, and community members make up a second set of target user groups. Often the communities most directly affected by a given policy choice, technology deployment, or infrastructure system do not have the resources—including time and personnel—to credibly wield data in support of their interests. This taxonomy classifies existing KPIs along dimensions that are integral to the conduct of cost-effective empirical work, helping to inform sound choices around data, and can work as a scaffolding to assist stakeholders as they craft their own KPIs. While the *agency* dimension in the taxonomy that follows initially refers to municipal bureaus in the sample data, any organization that seeks to develop and refine KPIs could be considered an agency for present purposes.

The intended purpose of this taxonomy is to help researchers, and decision-makers confidently and cost-effectively employ performance measurement data in their program/project analysis and model development efforts. Ultimately, the taxonomy is to be a tool in support of public and infrastructure sector performance measurement and sense-making. Furthermore, the taxonomy should help analysts select the best data inputs to facilitate replicable research. The taxonomy is also to be agnostic with respect to the size of analytical endeavor it supports. That is, the taxonomy should create value through clarity to those producing everything from univariate charts to digital twins. The taxonomy is intended to aid the composition of digital twins through improved understanding of the quality, cost-effectiveness, and salience of performance metrics in real cities and infrastructure systems.

4.2. Define objectives of a solution

4.2.1. Meta-Characteristic

The taxonomy's meta-characteristic addresses the *structural determinants of transaction costs that shape the accessibility, reusability, and analytical utility of KPIs in public and infrastructure sectors*. Alternatively stated, the meta-characteristic captures the main features of KPI transaction cost structures. These transaction cost structures determine how readily and cost effectively a given KPI can be of service to decision-makers seeking to improve systems performance. This meta-characteristic guides the selection of dimensions that differentiate KPIs based on their structural features rather than attempting precise measurement of transaction costs themselves. These structural features serve as conceptual waypoints that orient decision-makers toward productive lines of empirical investigation.

At the heart of this meta-characteristic is the recognition that the measurement of performance is not a costless activity. Effective management of systems performance requires empirical evidence to evaluate alternatives while balancing measurement costs against potential insights. This fundamental tradeoff shapes how performance data can serve as ingredients for modeling, analysis, and decision support. Prohibitively high transaction costs are commonly associated with the low level of observed reuse and consequent adoption of many KPIs within and across settings. The ability to classify a KPI based on its associated transaction costs should help decision-makers to select and improve cost-effective sources of insight.

4.2.2. Ending Conditions

The development process uses eight objective ending conditions obtained from Table 2 in [23] and listed below for ease of reference. The five subjective ending conditions, also listed below, are drawn from Table 3 in [23].

Objective Ending Conditions

i. All objects in sample have been examined.
ii. No merging or splitting of objects in last iteration.
iii. At least one object is classified under every characteristic of every dimension.
iv. No additional dimensions or characteristics in last iteration.
v. No merging or splitting of dimensions or characteristics in last iteration.
vi. All dimensions are unique.
vii. All characteristics are unique within each dimension.
viii. Each cell (combination of characteristics) is unique and is not repeated.

Subjective Ending Conditions

i. Concise
ii. Robust
iii. Comprehensive
iv. Extendible
v. Explanatory

4.2.3. Evaluation Goals

The evaluation phase of taxonomy development involves making a determination as to "whether the taxonomy is applicable and useful for providing structure to the phenomenon under consideration" [25]. A summary description of five evaluation goals (describing, identifying, classifying, analyzing, and clustering) can be found in Table 4 of [25]. A key evaluation goal is that the taxonomy can readily classify published KPIs from other settings which did not contribute data to the design and development process. Ideally, the taxonomy reduces the complexity creep associated with organizing the performance measurement data of a growing set of municipalities and infrastructure systems.

4.3. Design and development

4.3.1. Iteration 1 – Empirical to Conceptual

The design and development method begins with the choice of an empirical-to-conceptual approach. Published performance data and metadata from the City Budget Office of Portland, OR is chosen for classification with this taxonomy.[3] Interest lies in deriving possible dimensions (facets) for the taxonomy from this list of extant metadata on performance measurement data. The variable containing actual measurement data is, of course, not a candidate for a taxonomy dimension. Data from any municipality or infrastructure owner and operator could fill this roll well given sufficiently robust metadata. It is a point of this article to demonstrate that the methods explored here could be productively applied in a variety of settings.

A brief note about the dynamic nature of public data resources is warranted here. This project began with a version of the performance measurement dataset from October 4, 2022, but has moved to use a more recent dataset from February 2, 2024. The datasets differ in several ways, but the most consequential change for the taxonomy development process, is probably the decline in self-contained clarity to the data analyst due to program name metadata in natural language no longer being present. Effectively, this meant that a crosswalk of program names and acronyms was obtained from the original data source and merged into the newer data to enable sampling specific programs. An alternative approach that only used the most recent data would have had to broaden scope completely due to reduced visibility into the nature of the programs reporting performance measures. The changes encountered serve to demonstrate the role of good metadata in mitigating transaction costs for the analyst.

The sample inputs are restricted to focus on measures associated with public safety and resilience. The performance data from Portland, OR is published as an .xlsx file with 50 variables containing performance measure data and metadata.[4] Table 1 presents the steps used to

[3] The code used to develop the taxonomy and documentation of the input data can be found at the URL: https://datapub.nist.gov/od/id/mds2-3982. The raw data employed can be obtained from the following URLs.
https://public.tableau.com/app/profile/portland.city.budget.office/viz/FY2021-22PerformanceDashboard/PerformanceDashboard
https://www.portland.gov/cbo/documents/fy-22-23-performance-dashboard-dataset

screen the full dataset down to the sample employed in taxonomy development. Each row reflects the cumulative effect of all previous filters. The raw data contains 7805 records. The first screening step is to remove all records not associated with Key Performance Measures (KPMs). For Portland OR, measures are considered KPMs if they are indicators of core service delivery for the relevant bureau. Taken together, the set of KPMs constitute "a layer of Citywide indicators and guide Council decision-making" [45]. Once those measures not deemed as key to performance by their owners (publishing bureaus) are removed, the sample is reduced to 1753 records. Next, data for years other than 2022 is screened out, further reducing the sample to 228 records. To identify records relevant to the public safety and resilience focus, the sample only include records associated with a subset of bureaus (Table 2). There are 80 records associated with relevant bureaus in 2022. Finally, only 40 records pertaining to actual KPM data are considered. These 40 records cover 13 programs presented in Table 3 as well as an unassigned category.

Table 1. Portland (PDX) KPM Data Sample

Data Set Name	Filter	Record Count
PDX	N/A	7805
PDX_KPM	KPM == "YES"	1753
PDX_KPM_2022	Year == 2022	228
… Focal_Bureaus	Bureau ∈ Focal Bureaus	80
… FActual	PostingCode == "PMACTUAL"	40

Table 2: Focal Bureaus

Focal Bureau Code	Focal Bureau Name
EC	Emergency Communications
EM	Emergency Management
ES	Emergency Services
FR	Fire & Rescue
PL	Police Bureau
PN	Planning & Sustainability

Table 3: Focal Programs

9-1-1 Operations	Human Resources Development
Administration & Support	Policy, Research & Innovation
Comprehensive & Strategic Planning	Pollution Prevention
Emergency Management	Unassigned
Emergency Operations	Waste Reduction & Recycling
Emergency Response & Problem Solving	Wastewater
Employee Performance	Watershed

Once the record (observation) filtering process is complete, the first iteration of the taxonomy begins by establishing each of the 50 variables in the sample as a potential dimension. To this number, is added the *ProgramName* variable joined to the data from a previous year's published data set. Next, many dimensions that offer no immediate value to the taxonomy are removed. From the original 51 candidate dimensions, the initial culling reduces the number to 11, plus six variables which contain actual data or can be used to identify the performance measurement record in the dataset. Table 4 contains the subset of dimension candidates that survive this initial reduction using the raw variable names.

Table 4: Iteration 1 – Dimension Candidates from Raw Data

Dimension Candidates	
AggregationName	Bureau
CollectionMethod	DesiredDirection
Formula	FrequencyName
MeasureType	PrimaryProgram
ProgramName	Reliability
StrategicPlan	
Record Identifiers and Data	
Data	FiscalYear
FixedData	MeasureTitle
PMTarget	URL

Next, these 11 candidates are renamed and merged down to a set of eight. After conducting a survey of taxonomies, [23] finds "no agreement on what represents an appropriate number of dimensions", and instead offers the guidance of 7 dimensions plus or minus 2, citing [46].

A basic citation order, roughly informed by [36], is established next. The citation order is chosen to reflect a movement from the more concrete to the more abstract dimensions explaining a KPI. The resulting taxonomy iteration uses the notation described in [23] and expressed in Equation 1. Table 5 presents the nine dimensions of the taxonomy's first iteration as well as the raw variables from the Portland, OR data that were used or combined to arrive at the dimensions.

Equation 1: Taxonomy Notation

$$T = \left\{ D_i, i = 1, \dots, n \middle| D_i = \left\{ C_{ij}, j = 1, \dots, k_i; k_i \geq 2 \right\} \right\}$$

Table 5: Iteration 1 – Refined Dimensions

Dimension	**Dimension Name**	**Variables**
D_1	Agency	Bureau
D_2	Program	PrimaryProgram, ProgramName
D_3	Measure Type	MeasureType
D_4	Frequency	FrequencyName
D_5	Direction	DesiredDirection
D_6	Strategic Plan	StrategicPlan
D_7	Explainability	AggregationName, Formula, CollectionMethod
D_8	Quality	Reliability

An important aspect in which public and infrastructure sector KPIs differ from other measurements that do not concern organizations operating in the public interest is that KPIs are not immune from the agendas of those who produce and wield them. Peter Drucker notes that "any measurement in a business enterprise determines action—both on the part of the measurer and the measured—and thereby directs, limits and causes behavior and performance of the enterprise" [47]. This simple truth about measurement in and of business enterprise is no less applicable to public and infrastructure sector organizations. If KPIs are to be included in digital twins and other models where outputs will influence policy, then tools are needed to clarify and explain matters regarding both the measurer and the measured. Some institutional progress is needed before KPIs can be credibly incorporated into decision support tools with two-way feedback between models and policy. Analysts, policy makers, and the public need tools for evaluating when a KPI is chosen for its merit with respect to guiding policy and when it is chosen merely to win intramural budget battles. This taxonomy is intended as one such explanatory tool for elevating the quality of empiricism and discussion informing public decision-making processes. Several of the dimensions that emerge over the following development iterations explain aspects of a KPI's origin and documentation that help the analyst understand possible sources of transaction costs rooted in institutional agenda and bias. In the following section, dimension names are highlighted once in bold faced italic font and then written in italics thereafter. The characteristic names within those dimensions are displayed in italics when referenced explicitly.

4.3.1.1. Initial Dimensions

The ***agency*** dimension contains the identity of the government entity or nongovernmental organization (NGO) that owns the KPI. That is, *agency* contains a string indicating the organization responsible for collecting and reporting the data. The characteristics collected in this dimension should communicate which government entity, infrastructure owners and operators, local NGO, or academia could be sought out to learn more about the KPI and its production process. This dimension points to who you can talk to if you have questions, reducing the transaction costs of working with a specific KPI. It is possible that the *agency* variable could be specified at alternative levels of organizational hierarchy. The *agency* dimension is also crucial to the taxonomy for the simple reason that it makes explicit which organization owns the problems under measurement. Knowing who earns the credit for success or responsibility for failure is important for any analyst seeking to understand the appropriate and defensible use of public performance information, as this information will inform some of the structure of transaction costs.

The ***program*** dimension sorts the sampled KPIs into the broad programmatic categories relevant to public safety and community resilience. The *program* dimension contains the largest number of initial characteristics (categories). This number could be reduced through merging categories if the resultant information loss was deemed acceptable for purpose. All programs are kept for the first iteration, including the "unassigned" response. Organizational identifiers like *agency* and *program* provide institutional coordinates for KPIs that help a researcher understand the organizational and institutional architecture through which transaction costs materialize.

The ***measure type*** dimension is drawn directly from performance measurement documentation for Portland, OR [45]. While alternative categorization schemes are also plausible, the Portland categories are incisive for present purposes. Each KPI in the data may be tagged as one of the following: *efficiency*, *outcome*, *output*, or *workload*. *Efficiency* measures capture the use of inputs to produce *outputs*. *Outcome* measures convey the public benefits resulting from programs and services, including benefits over which managerial control may be limited. Service quality measures may be categorized as outcomes. By comparison, municipal documentation [45] explains that “Outputs are activity-oriented, measurable, and usually under managerial control”. *Workload* type measures reflect the scale of demands placed on bureau resources. *Measure type* captures the fundamental complexity of the phenomena under measurement and thus informs the associated transaction costs. The transaction costs of *workload* measures are likely the lowest as the task of measuring direct inputs to production is straightforward. *Output* measures, which track activities under managerial control may be more complex to produce and thus should exhibit moderate transaction costs. *Efficiency* measures must combine the transaction costs associated with both *workload* and *output* components. Finally, *outcome* measures must account for the complexity of potentially numerous causal pathways through which action is converted into results. *Outcome* measures thus entail the highest transaction costs that can be reasonably expected.

The fourth dimension contains information on the ***frequency*** of the measure's collection. The most useful KPIs are more frequently collected and disseminated to support action by responsible parties. Annual data is historically common due to the lower total workload associated with that frequency of collection. Characteristics in this dimension include the following collection intervals: *annual*, *quarterly*, *monthly*, *weekly*, *transactional*, *intermittent*, and *not applicable*. The *frequency* dimension directly effects how often KPI data could be updated for a digital twin or other model. Models harnessing annual data will have a harder time capturing seasonal or short-term dynamics than models using more frequently collected KPIs. *Frequency* is fundamental to transaction costs as repeated activities offer the opportunity to gain efficiencies through learning effects and reduced uncertainty. Per unit transaction costs generally fall as frequency increases. It is possible that total transaction costs are higher with KPIs reported at a higher frequency.

The ***direction*** dimension reflects the desirability of the underlying phenomena measured by the KPI. This dimension contains some judgement on the part of the publishing *agency* and data analyst. Some measures capture clearly desirable phenomena and are assigned a *direction* characteristic of *up*. Other measures capturing negative phenomena are typically assigned a value of *down*. As some phenomena are neither inherently good or bad, the categories of *none* and *not applicable* are also made available through the taxonomy. An example KPI with *direction* coded as not applicable is “Number of community complaints of officer misconduct”.

While misconduct is clearly not desirable, neither would be a downtrend due to a lack of reporting. Within the data, some KPIs such as reported crime counts and percentages of employees retained annually are coded with a *direction* of *none*. The *none* and *not applicable* characteristics may warrant combination in following iterations. *Direction* is useful to the analyst as it removes the cost to them of determining what the desired outcome is from the perspective of the publishing *agency* and its stakeholders. That is, the transaction costs associated with interpretation are reduced when *direction* is clearly documented. When desired direction is left unreported, it is a signal to the analyst that additional resources will be needed to interpret changes in the KPI. The transaction costs of replication, reuse, and decision support are also increased when the *direction* of the KPI trend desired by its publisher is not disclosed.

The sixth dimension contains a small set of characteristics indicating whether there is an explicit reference in the data documentation linking the measure to a ***strategic plan*** or comparable document (in whole or in part). The potential responses for whether a strategic plan is linked to the KPI are coded as follows: *none*, *vague*, *specific*. In practice, what separates the two non-empty responses is whether an analyst can find the specific documentation referenced in the data in a reasonable amount of time without additional identifying information. The *strategic plan* dimension lets the analyst or modeler know something of the level to which a KPI factors into strategic planning decisions. KPI with *specific* strategic plan documents are more likely to be actively incorporated into an organization's actions to drive improved service delivery to the public. A *specific* strategic plan reference leads to the lowest transaction costs for the analyst with respect to determining the strategic context of the KPI as it minimizes effort to verify and understand how the KPI fits into organizational strategy. Vague or missing characteristics for the strategic plan dimension can suggest similar effects on the researcher as do vague or missing citations in scholarly work products.

The ***explainability*** dimension harnesses the largest number of variables from the input data. It is also among the most abstract of the dimensions. Three inputs are used to inform the explainability dimension: text descriptions of the collection methods, formula, and form of aggregation such as sums or averages. This dimension captures the analyst's answer to the question of "how well do you know what goes into the methodology and practice of producing the KPI?" In an informal sense, this is a question of how confident the analyst can be when using a given KPI about its fitness for purpose. Transaction costs fall for the analyst when published KPIs are highly explainable. However, the large number of component variables informing this dimension as currently stated will likely reduce the capacity for the dimension to offer clear signals regarding transaction cost structures.

The ***quality*** dimension is initially determined by the reported reliability value of KPIs. In the raw data, the values obtained for the reliability variable include *low*, *medium*, and *high*. Unsurprisingly, the number of low-reliability KPIs in use is proportionally small. This dimension contains subjective information that may group together measures of substantially different quality. While quality matters to transaction costs structures, how it matters remains less than obvious with the available characterization.

4.3.2. Iteration 2 – Conceptual to Empirical

The second iteration of taxonomy development follows the conceptual to empirical approach. The first iteration has provided strong concepts for the more concrete dimensions of KPI classification. The first four dimensions concern objective facets of KPIs that are unlikely to foster disagreement between and within stakeholder groups. By contrast, the more abstract KPI facets, concerning the desired direction and strategy for progress, as well as the clarity of documentation and quality of measures necessarily contain subjective judgements. The conceptual to empirical approach starts with the meta-characteristic and then moves to the conception of dimensions that are central to that perspective. The second iteration begins where the first iteration arrived at vague, abstract concepts in need of better alignment with the meta-characteristic.

The first iteration of the taxonomy attempts to cover a lot of conceptual ground with the two dimensions of explainability and quality. While explainability and quality certainly touch on the transaction costs of working with KPIs, they are too vague to effectively manage and mitigate the problems of KPI complexity. Some other concepts that are related to these dimensions include: transparency, calculation complexity, definitional clarity, benchmarking, limitations or caveats, actionability, and the potential for gaming or manipulation of public information. These concepts have subjective components that complicate their implementation within the taxonomy, thus offering little resolution to the problem of vagueness.

Instead of introducing several more dimensions capturing aspects of these nuanced and subjective topics, the final two dimensions are replaced with one labeled ***Formula*** and one labeled ***Collection Method***. At the root of this decision is the need of the taxonomy to meet the intended users at a point of crucial concern to them. Applied economists and other analysts need to create reproducible work. That is, the data and code necessary to replicate findings needs to be accessible to others who might want to learn from a given approach or check others' work. The documentation accompanying empirical analysis needs to manage the complexity of the process that produced the research outputs in a manner that leaves actionable instructions for others to follow. When formula and collection data documentation is weak, the transaction costs of replication are considerably higher, all else equal. These two dimensions have a direct relationship back to the meta-characteristic and are conceptually accessible to the intended end user.

Ideally, researchers want to use KPIs in their analysis and model development that are well-documented, facilitating the replication of results. Because raw data is collected in the field and analysis is conducted using software and hardware stacks often unique to specific researchers, ensuring replication requires documentation and communication of the particulars associated with both phases of research. Some overlap is likely present between the *Collection Method* and *Formula* dimensions. Roughly stated, the *Collection Method* dimension captures the level of documentation associated with data collection in the field, and the *Formula* dimension explains the level of mathematical specificity associated with the calculation or statistical estimate of a KPI.

The two new dimensions were contained as characteristics within the previous explainability variable along with *Aggregation*. *Aggregation* is now conceptually incorporated into the broader *Formula* dimension.

Characteristics are developed within the Collection Method dimension by clustering the responses found in the sample data. A characteristic labeled *none* emerges to reflect KPIs for which no information is offered regarding collection. Another cluster of responses contains *vague* or unhelpful strings of text. These are called out as distinct from the *none* category to help identify opportunities where clarification could improve public data products. The third cluster of responses of collection method contain *pointers to databases* maintained or owned by named vendors or other organizations. The fourth and fifth clusters to become characteristics are *simple* and *detailed* text descriptions of the organizational structures and processes involved in the production of the KPI. A main feature differentiating these last two groupings is the complexity of the measure and its acquisition process. Simple procedures get *simple* descriptions. An initial implementation of this demarcation relies on establishing the length of the text response and its percentile. To keep matters tractable at this point in the taxonomy development process, above-median length responses are categorized as *detailed*. Future work incorporating additional KPI for classification may utilize more sophisticated approaches to clustering.

The *Formula* dimension is also applied to the sample of data from Portland, OR. The first cluster forms around *null* responses. The second cluster is around *vague* responses including to other entities. The third and fourth clusters respectively form around plain text descriptions of *simple* and *complex* mathematical equations for calculating or estimating a given KPI. Simple and complex responses to the Formula dimension follow the same logic of response length percentile as was discussed for Collection Method.

A final dimension added in the second iteration of taxonomy development is labeled **URL**. This dimension is binary. Some KPIs report no link to internet resources through a URL. Of those KPIs that do report a URL, most report a general reference to a related landing page. A sparsely populated category labeled precise was not warranted by the sample data, but a future version of the taxonomy might beneficially add a category for KPI with a link to more relevant and specific web documentation.

As the meta-characteristic focuses attention on the *structural determinants of transaction costs that shape the accessibility, reusability, and analytical utility of KPIs in public and infrastructure sectors*, it is worth making explicit the connections between the new dimensions and the meta-characteristic. There are both content and documentation components to the transaction cost implications of these new dimensions. In each case, clear, concise and complete documentation of the formula and collection methods is associated with the lowest transaction costs. Less effective documentation, or even a complete lack of documentation will impose greater transaction costs on the end users of a KPI. In the case of the *URL* dimension, a single string of text that can direct a researcher or interested reader to relevant resources reduces transaction costs relative to the unguided search that may occur in the absence of such information. However, even if effective documentation is present, the complexity of the functional form (*formula*) or data production process (*collection method*) directly effects transaction costs. Characteristics in these dimensions that indicated increasing levels of complexity may be helpful in identifying the costliest of competing alternatives on the grounds of transaction costs. With this information in hand, the researcher can orient their efforts to more cost-effective approaches to empirical investigations. Beyond increasing the transaction costs associated with data use, poor documentation also increases the chance that KPIs will be misunderstood and misapplied.

The second iteration of the taxonomy concludes here. Clearly, the changes made in this iteration imply that the ending conditions have not been met.

4.3.3. Iteration 3 – Empirical to Conceptual

Iteration 3 begins by rearranging the dimensions slightly, bringing the *URL* binary forward in the citation order to precede the *strategic plan* dimension. The characteristics are also reordered in several dimensions so that they progress from a null response to the conceptual opposite. The *direction* dimension is simplified through the merging of *none* and *not applicable* responses.

4.3.4. Iteration 4 – Static

Table 6 presents the evolution of taxonomy dimensions over four iterations. The row order does not reflect the final citation order, which can be found in **Error! Reference source not f ound.**. No Changes were made in Iteration 4. Attention now shifts to the demonstration of the taxonomy.

Table 6: Taxonomy Dimension Change Matrix

Dimension	Iteration 1	Iteration 2	Iteration 3	Iteration 4
Agency	Added	Present	Present	Present
Program	Added	Present	Present	Present
Measure Type	Added	Present	Present	Present
Frequency	Added	Present	Present	Present
Direction	Added	Present	Modified	Present
Strategic Plan	Added	Present	Present	Present
Explainability	Added	Removed	-	-
Quality	Added	Removed	-	-
Collection Method	-	Added	Present	Present
Formula	-	Added	Present	Present
URL	-	Added	Present	Present

4.4. Demonstration

A demonstration of the current taxonomy is conducted by reviewing how it orders the 40 objects in the Portland, OR KPI sample, which is defined in Table 1. The first task of the demonstration is to check the objective ending conditions from Sec. 4.2.2. **Error! Reference source not found.** contains all the dimensions and characteristics in the taxonomy with the number of objects from the sample possessing each characteristic presented in parentheses. The individual key performance measures are presented in Table 8 found in Appendix C.

Table 7: Sample Summary – Dimensions and Characteristics

	D01 Agency	D02 Program	D03 Measure Type	D04 Frequency	D05 Direction	D06 URL	D07 Strategic Plan	D08 Collection Method	D09 Formula
Ch.1	EC (4)	9-1-1 Operations (1)	Efficiency (2)	Annual (29)	None (10)	No (16)	None (24)	None (18)	None (17)
Ch.2	EM (10)	Administration & Support (1)	Outcome (27)	Quarterly (3)	Down (9)	Yes (24)	Vague (10)	Vague (5)	Vague (0)
Ch.3	ES (5)	Comprehensive & Strategic Planning (2)	Output (4)	Monthly (1)	Up (21)		Specific (6)	Database Pointer (3)	Simple (4)
Ch.4	FR (2)	Emergency Management (10)	Workload (7)	Weekly (0)			Integral (0)	Simple Description (1)	Complex (19)
Ch.5	PL (14)	Emergency Operations (2)		Transactional (4)				Detailed Description (13)	
Ch.6	PN (5)	Emergency Response & Problem Solving (4)		Intermittent (3)					
Ch.7		Employee Performance (3)		NA (0)					
Ch.8		Human Resources Development (1)							
Ch.9		Policy, Research & Innovation (1)							
Ch.10		Pollution Prevention (2)							
Ch.11		Unassigned (10)							
Ch.12		Waste Reduction & Recycling (1)							
Ch.13		Wastewater (1)							
Ch.14		Watershed (1)							

Having examined all 40 KPIs in the sample, no objects were merged or split in the last iteration. No sample objects are classified in the *weekly* or *not applicable* characteristics of the *frequency* dimension. No KPIs are classified as having the *contested* characteristic for the *direction* dimension. The *integral* characteristic of the *strategic plan* dimension and the *vague* characteristic of the *formula* dimension are also unpopulated. The ending condition that at least one object is classified under every characteristic of every dimension is thus not met explicitly. However, this condition may be met by removing these empty categories from the taxonomy or ingesting more KPIs from other settings.

No additional dimensions or characteristics were developed, merged, or split in the last iteration. All dimensions and sets of characteristics within them are unique. The final suggested objective ending condition—that each combination of characteristics held by an object is unique within the sample—is not met. There are unique measure titles (KPI names) associated with the same collection of taxonomy characteristics. This last condition is disregarded because the fact that two different KPIs have the same taxonomy profile does not reduce the taxonomy's fitness for purpose.

Considering the removal of empty categories and the decision to disregard the uniqueness condition for the combination of all characteristics across dimensions, all objective conditions for ending the iterative taxonomy development process, laid out in Sec. 4.2.2, have now been met. More iterations could be conducted as part of future research if the incorporation of additional KPIs leads to the need for alterations to the structure of dimensions and characteristics. The discussion now moves to the subjective evaluation of the taxonomy artifact.

4.5. Evaluation

Next, the taxonomy artifact is evaluated with respect to the subjective ending conditions introduced in Sec. 4.2.2. Following the guidance of [23], the taxonomy artifact should be concise, robust, comprehensive, extendible, and explanatory.

While the taxonomy artifact has 9 dimensions, enough of those dimensions are restricted to a small number of characteristics to maintain some parsimony overall. Ordering the sample of 40 KPIs along the 9 dimensions was not an onerous task. The resulting artifact is sufficiently concise for its purpose.

The taxonomy is robust enough to produce classifications that meaningfully differentiate between those KPIs that are and are not well-documented. This helps in determining which KPIs are amenable to replicable empirical analysis of infrastructure programs and interests. The taxonomy can capture the difference between KPIs that are good candidates for incorporation in an analysis or modeling exercise and those that should be left out in favor of more-defensible alternatives.

The taxonomy artifact is comprehensive in that it was able to classify all the KPIs in the sample. Furthermore, the taxonomy captures much of what is of interest to the user group of analysts, by outlining the dimensions important to understanding the transaction costs associated with KPI use for the purpose of producing replicable research on public and infrastructure systems.

The extendibility of the taxonomy should be evident from the removal of conceptually plausible, but unpopulated categories from several dimensions. If KPIs of distinct frequency, collection methods, or formula types arise, they can easily be incorporated into the taxonomy. Also, the relatively concise nature of many dimensions (with few characteristics) makes it possible to expand the set of characteristics without making the taxonomy unwieldy. If more dimensions were needed, some of the dimensions early in the citation order might need to be reconsidered and merged, but that would not pose any great problem.

Finally, the taxonomy meets the condition of being explanatory. The dimensions included in the taxonomy provide useful explanations regarding the transaction cost and fitness of a KPI for employment in analysis and modeling. An analyst can use the taxonomy to obtain an explanation containing a set of basic facts about a potential data input. With the results of this classification scheme, they can make an informed decision on how best to employ or disregard a piece of public data.

The taxonomy artifact performs acceptably with respect to the five subjective ending conditions. That is, the taxonomy is concise, robust, comprehensive, extendible, and explanatory.

4.6. Communication

Communication regarding the taxonomy development process and outcomes entails three main design recommendations according to [25]. The evolution of the taxonomy and the approaches taken for each iteration must be documented. The taxonomy needs to be visualized with a mixture of tables, figures, and textual components that are fit for the intended purpose and target user groups. Descriptions of each dimension and its characteristics are needed as well. These three recommendations have been met in Sec. 4. An additional visualization method, similar in form to a radar plot, is presented in Sec. 4.6.3.

4.6.1. Taxonomy Reporting

This manuscript constitutes the communication component of the taxonomy development process. The input KPI data, and script needed to replicate the outcomes are also part of the communication effort. The URL for the input data is reported in Sec. 4.3.1. The MATLAB script used to filter the input data, produce tables, and visualize KPI characteristics with a spider plot is available at https://datapub.nist.gov/od/id/mds2-3982.

4.6.2. A Note on KPI Data Sources

The format and availability of Portland, OR KPI data made it a useful choice of input material for the taxonomy development process. However, any infrastructure owner and operator publishing KPIs with sufficient metadata in a single spreadsheet file could just as easily have been organized. The communication of data and metadata in a single file allowed us to avoid making the subjective determination of when enough data had been collected on a measure to be sure about its level of documentation. In other words, for organizations that report KPI data without directly and adjacently supplied metadata, not only is data collection more time consuming, but consistency in the efforts associated with collection must be reasonably uniform to avoid the introduction of bias. When good metadata is not available, the transaction costs of working with data can be considerably higher. The evaluation of this taxonomy begins with the discussion contained in this document. Ultimately, a more thorough and lasting evaluation will require that the taxonomy be subjected to testing through the incorporation of additional KPI data from many other municipalities and infrastructure owners and operators.

4.6.3. Visualizing KPI Adoption Using this Taxonomy

The taxonomy structure is intentionally visualized as a simple table to maintain accessibility. However, other visualization methods can offer better clarity into the objects classified (KPIs). The taxonomy artifact produced here can be harnessed to visualize the adoption of a KPI by an organization or stakeholder community using a modified radar plot, also sometimes referred to as a spider plot.

The code to implement this visualization is prototyped with MATLAB R2022a and presented in Appendix C. The fourth through ninth dimensions of the taxonomy each form a potential axis or spoke of the radar plot. Where this visualization differs from others is that the axes capture categorical rather than monotonic numerical values. The radar plot line-loops are color coded to distinguish between measure types. For a dimension such as Frequency, annual KPIs are marked proximally to the center of the plot while more frequently collected KPIs are positioned distally. For a dimension such as Collection Method, KPIs reporting the characteristic of none are scored proximally, while those with complex descriptions would be marked at the maximal distance from the plot center. **Error! Reference source not found.** presents a separate radar p lot for each *agency*. The same could be done for different programs. The resulting graphics allow for the visual comparison of the state of KPI development, documentation, adoption, and fitness for analytical applications across programs, agencies (bureaus), cities, and infrastructure systems.

Error! Reference source not found. shows that outcome measure types are reported by all six a gencies (bureaus) in the sample. The plot lines are transparent so that darker lines indicate multiple KPIs with the same characteristics within a given taxonomy dimension. The overall shape of these lines communicates the differences in approach to KPI development and use across organizations and their respective operating domains. While both the Emergency Communications and Planning and Sustainability bureaus only report outcome KPIs, they do so in different ways. The former uses database pointers to indicate the systems housing their data while the latter relies on detailed text descriptions. It is worth reiterating that the visualizations are constrained by the level of detail and fidelity provided by the bureaus regarding their contributions to the central performance measure dataset made available as a spreadsheet file.

The taxonomy captures those characteristics of KPIs that are centrally reported and therefore do not require additional subjectively managed data wrangling.

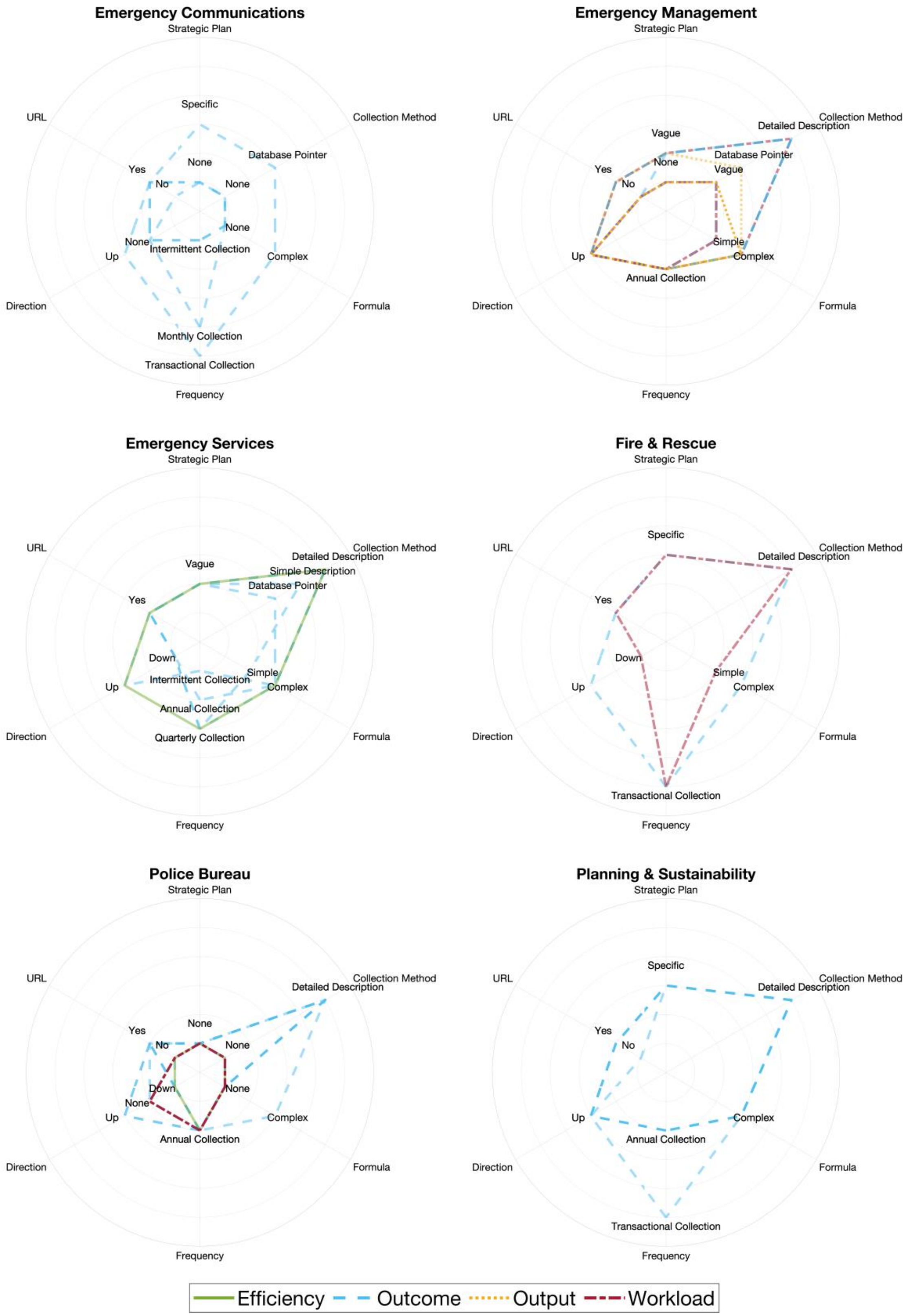

Fig. 1: KPI Taxonomy Radar Plots by Agency (Bureau)

The radar plot visualization makes it clear which organizations have KPIs directly linked to specific strategic plan documents, which make only vague references, and which ones do not offer any explanation of what guides their actions and approach to performance measurement. KPIs for which greater documentary detail exists display as larger polygons on the radar plots, so over time, improvements to documentation and KPI quality will be visible in the expansion of these polygons. Future versions of these radar plots may benefit from the incorporation of a third dimension passing thought the origin of the plot that allows for the visual separation of each polygon. The sample data contained in **Error! Reference source not found.** may also be f ound located in Table 8 of Appendix C.

5. Conclusion

This manuscript presents a taxonomy development process and resulting artifact intended to help organize public and infrastructure sector KPIs in support of cost-effective empirical analysis and modeling activities. At the margin, data production costs have fallen, but analyzing and drawing insights from this data still comes at costs that reflect documentation quality. KPI adoption could accelerate with a focus on options enjoying favorable transaction cost structures. Well-documented KPIs create advantages for decision makers by reducing analytical costs of performance management and enabling more effective policy formation. The taxonomy artifact allows the sorting of KPIs by their dimensions and characteristics, offering a framework for constructing comparative estimates of transaction costs.

The intended end users for the public and infrastructure sector KPI taxonomy includes economists, social scientists, data scientists, and to a lesser extent, the decision-makers who rely on these empirical analysts. Professional data wranglers and analysts may employ this taxonomy in their own evaluations of which data is cost-effective for producing insights and fit for their purposes. This taxonomy is also intended to help inform the decision to integrate a given KPI into an emerging class of models such as digital twins of communities, infrastructure systems, or regions. For a taxonomy of digital twins, see [48]. The complexity of these emerging models ensures that a failure to account for the effects of inefficient transaction cost structures can seriously limit their value propositions. Taxonomies can evolve with the phenomena they address. While not included in the 2020 taxonomy [48], in time, the nature of KPI integration may become an important dimension of digital twins. Whether public and infrastructure sector KPIs are used in simple calculations, incorporated into regression specifications, or integrated into high-fidelity, high frequency synchronized models of the real environment, the purpose of this taxonomy is to provide clarity regarding the nature, quality, and costs of data inputs.

The taxonomy's meta-characteristic, or angle on the phenomena of KPIs, concerns the ***structural determinants of transaction costs that shape the accessibility, reusability, and analytical utility of KPIs in public and infrastructure sectors***. This meta-characteristic informs the selection of dimensions that differentiate KPIs based on their structural features and is not intended as a precise measurement of transaction costs themselves. These structural features serve as conceptual waypoints that orient decision-makers toward productive lines of empirical investigation through cost-effective approaches to performance management.

Through an iterative process, real KPI data from the city of Portland, OR, is employed to develop a simple taxonomy of public and infrastructure sector KPIs with a focus on public safety and community resilience (Appendix A). While the initial focus is narrow, it concerns activities that Americans count on every day, in every community, large and small, urban, suburban, and rural. Future work may readily extend this taxonomy to the classification of KPIs and other public performance data from other domains. The taxonomy artifact presented in this manuscript is complete and fit for the purpose of the initial public safety and resilience focus. In the longer run, the value proposition of the taxonomy will evolve as it is tested with data from additional settings and domains.

References

[1] Black PE, Scarfone K, Souppaya M (2008) Cyber security metrics and measures. *Wiley Handbook of Science and Technology for Homeland Security*:1-15. https://doi.org/10.1002/9780470087923.hhs440

[2] Serrano M, Griffor E, Wollman D, Dunaway M, Burns M, Rhee S, Greer C (2022) Smart cities and communities: A key performance indicators framework. *NIST Special Publication* 1900:206. https://doi.org/10.6028/NIST.SP.1900-206-upd1

[3] Gabaix X , Laibson D (2006) Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics* 121(2):505-540. https://doi.org/10.1162/qjec.2006.121.2.505

[4] Ellison G , Ellison SF (2009) Search, obfuscation, and price elasticities on the internet. *Econometrica* 77(2):427-452. https://doi.org/10.3982/ECTA5708

[5] Bettencourt LM (2013) The origins of scaling in cities. *science* 340(6139):1438-1441. https://doi.org/10.1126/science.1235823

[6] Varian HR (2010) Computer mediated transactions. *American Economic Review* 100(2):1-10. https://doi.org/10.1257/aer.100.2.1

[7] Parmenter D (2015) *Key performance indicators: developing, implementing, and using winning KPIs* (John Wiley & Sons).https://doi.org/10.1002/9781119019855

[8] Schrage M , Kiron D (2018) Leading with next-generation key performance indicators. *MIT Sloan Management Review*. https://sloanreview.mit.edu/projects/leading-with-next-generation-key-performance-indicators/

[9] Schroeder K, Trinh H, Pillitteri V (2024) Measurement Guide for Information Security: Volume 1—Identifying and Selecting Measures. (National Institute of Standards and Technology). https://doi.org/10.6028/nist.sp.800-55v1

[10] Härting R-C , Sprengel A (2019) Cost-benefit considerations for data analytics-an SME-oriented framework enhanced by a management perspective and the process of idea generation. *Procedia Computer Science* 159:1537-1546. https://doi.org/10.1016/j.procs.2019.09.324

[11] Williamson OE (1975) Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

[12] Allen DW (1991) What are transaction costs? *Rsch in L & Econ* 14:1.

[13] Alchian AA (1965) Some economics of property rights. *Il politico*:816-829. https://www.jstor.org/stable/43206327

[14] Williamson OE (1980) *Organizational innovation: the transaction cost approach* (University of Pennsylvania, Center for the Study of Organizational Innovation, Philadelphia, PA),https://books.google.com/books?id=eP8wMQAACAAJ

[15] Commons JR (1936) Institutional economics. *The American Economic Review* 26(1):237-249.

[16] Coase RH (2012) The nature of the firm. *The Roots of Logistics,* (Springer), pp 317-333. https://doi.org/10.1007/978-3-642-27922-5_22

[17] Hayek FA (2013) The use of knowledge in society. *Modern understandings of liberty and property,* (Routledge), pp 27-38. https://doi.org/10.4324/9781315053844

[18] Geyskens I, Steenkamp J-BE, Kumar N (2006) Make, buy, or ally: A transaction cost theory meta-analysis. *Academy of management journal* 49(3):519-543. https://doi.org/10.5465/amj.2006.21794670

[19] Macher JT , Richman BD (2008) Transaction cost economics: An assessment of empirical research in the social sciences. *Business and politics* 10(1):1-63. https://doi.org/10.2202/1469-3569.1210

[20] Cuypers IR, Hennart J-F, Silverman BS, Ertug G (2021) Transaction cost theory: Past progress, current challenges, and suggestions for the future. *Academy of Management Annals* 15(1):111-150. https://doi.org/10.5465/annals.2019.0051

[21] Grady D, Thiemann C, Brockmann D (2012) Robust classification of salient links in complex networks. *Nature communications* 3(1):864. https://doi.org/10.1038/ncomms1847

[22] Riley J (2017) *Understanding Metadata: What is Metadata, and What is it For?: A Primer* (National Information Standards Organization, Washington DC, United States),http://www.niso.org/publications/press/UnderstandingMetadata.pdf

[23] Nickerson RC, Varshney U, Muntermann J (2013) A method for taxonomy development and its application in information systems. *European Journal of Information Systems* 22(3):336-359. https://doi.org/10.1057/ejis.2012.26

[24] Ulrich H, Kock-Schoppenhauer A-K, Deppenwiese N, Gött R, Kern J, Lablans M, Majeed RW, Stöhr MR, Stausberg J, Varghese J (2022) Understanding the nature of metadata: systematic review. *Journal of medical Internet research* 24(1):e25440. https://doi.org/10.2196/25440

[25] Kundisch D, Muntermann J, Oberländer AM, Rau D, Röglinger M, Schoormann T, Szopinski D (2022) An Update for Taxonomy Designers. *Business & Information Systems Engineering* 64(4):421-439. https://doi.org/10.1007/s12599-021-00723-x

[26] Jones MB, Berkley C, Bojilova J, Schildhauer M (2001) Managing scientific metadata. *IEEE Internet Computing* 5(5):59-68. https://doi.org/10.1109/4236.957896

[27] Wand Y, Monarchi DE, Parsons J, Woo CC (1995) Theoretical foundations for conceptual modelling in information systems development. *Decision support systems* 15(4):285-304. https://doi.org/10.1016/0167-9236(94)00043-6

[28] Lippell H (2022) *Taxonomies: Practical Approaches to Developing and Managing Vocabularies for Digital Information* (Facet Publishing, London).https://doi.org/10.29085/9781783304837

[29] National Information Standards Organization (2010)– *ANSI/NISO Z39.19-2005 (R2010) Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (National Information Standards Organization, Baltimore, MD). https://doi.org/10.3789/ansi.niso.z39.19-2005R2010

[30] Fayen EG (2007) Guidelines for the construction, format, and management of monolingual controlled vocabularies: A revision of ANSI/NISO Z39. 19 for the 21st century. *Information Wissenschaft und Praxis* 58(8):445. https://www.b-i-t-online.de/pdf/iwp/IWP2007-8.pdf

[31] Clarke D , Stella G (2011) ISO 25964: a standard in support of KOS interoperability. *Facets of knowledge organization: proceedings of the ISKO UK Second Biennial Conference, 4th-5th July 2011, London*, pp 129-134. https://doi.org/10.1108/el-04-2013-0059

[32] Gruber TR (1995) Toward principles for the design of ontologies used for knowledge sharing? *International journal of human-computer studies* 43(5-6):907-928. https://doi.org/10.1006/ijhc.1995.1081

[33] Miles A , Bechhofer S (2009) *SKOS simple knowledge organization system reference*. Available at https://www.w3.org/TR/skos-reference/.

[34] Schwarz A, Mehta M, Johnson N, Chin WW (2007) Understanding frameworks and reviews: a commentary to assist us in moving our field forward by analyzing our past. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 38(3):29-50. https://doi.org/10.1145/1278253.1278259

[35] Denton W (2003) *How to make a faceted classification and put it on the web*. Available at https://www.miskatonic.org/library/facet-web-howto.html.

[36] Broughton V (2006) The need for a faceted classification as the basis of all methods of information retrieval. *Aslib proceedings*, (Emerald Group Publishing Limited), pp 49-72. https://doi.org/10.1108/00012530610648671

[37] Classification Research Group (1964) Some Problems of a General Classification Scheme, report of a conference held in London, June 1963.

[38] Ranganathan SR (1963) *Colon classification: basic classification* (Asia Publishing House), Vol. 26

[39] Campbell DT (1974) Assessing the impact of planned social change. *Psychology*. https://doi.org/10.56645/jmde.v7i15.297

[40] Oberländer AM, Lösser B, Rau D (2019) Taxonomy research in information systems: A systematic assessment. *European Conference on Information Systems (ECIS) 2019*, (Association for Information Systems). https://aisel.aisnet.org/ecis2019_rp/144/

[41] Nickerson RC, Muntermann J, Varshney U (2010) Taxonomy development in information systems: a literature survey and problem statement. https://doi.org/10.1057/ejis.2012.26

[42] Hevner A , Chatterjee S (2010) Design science research in information systems. *Design research in information systems: theory and practice,* (Springer), pp 9-22. https://doi.org/10.1007/978-1-4419-5653-8_2

[43] March ST , Smith GF (1995) Design and natural science research on information technology. *Decision support systems* 15(4):251-266. https://doi.org/10.1016/0167-9236(94)00041-2

[44] Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S (2007) A design science research methodology for information systems research. *Journal of management information systems* 24(3):45-77. https://doi.org/10.2753/mis0742-1222240302

[45] City Budget Office (2022) FY 2021-22 Budget Manual, City of Portland, Oregon. https://www.portland.gov/budget/2021-2022-budget/development/preparation

[46] Miller GA (1956) The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological review* 63:91-97. https://doi.org/10.1037/h0043158

[47] Drucker PF (1959) Long-range planning—challenge to management science. *Management science* 5(3):238-249. https://doi.org/10.1287/mnsc.5.3.238

[48] Van der Valk H, Haße H, Möller F, Arbter M, Henning J-L, Otto B (2020) A Taxonomy of Digital Twins. *AMCIS*. https://www.researchgate.net/publication/341235159_A_Taxonomy_of_Digital_Twins

[49] Williamson OE (1979) Transaction-cost economics: the governance of contractual relations. *The journal of Law and Economics* 22(2):233-261. https://doi.org/10.1086/466942

[50] Williamson OE (2010) Transaction cost economics: The natural progression. *American Economic Review* 100(3):673-690. https://doi.org/10.1257/aer.100.3.673

[51] Alchian A (1997) Some implications of recognition of property right transactions costs. *The Economic Foundations of Property Rights,* (Edward Elgar Publishing), pp 7-28. https://doi.org/10.4337/9781035303441.00010

[52] Allan S, Buckley AR, Meacham JE (2001) *Atlas of Oregon* (University of Oregon Press), Second Ed.

[53] Merriam PG (1979) The "Other Portland": A Statistical Note on Foreign-Born, 1860-1910. *Oregon Historical Quarterly* 80(3):258-268.

[54] Abbott C (2001) Portland: civic culture and civic opportunity. *Oregon Historical Quarterly* 102(1):6-21.

[55] Richmond HR (1970) The History of the Portland District Corps of Engineers, 1871-1969.

[56] Helgeson J , O'Fallon C (2021) Resilience dividends and resilience windfalls: Narratives that tie disaster resilience co-benefits to long-term sustainability. *Sustainability* 13(8):4554. https://doi.org/10.3390/su13084554

Appendix A. The Portland, OR Test Case

The public and infrastructure sector KPI data used for initial taxonomy development in this document is obtained from publicly accessible performance management data published by the City Budget Office of Portland, OR. This appendix presents a discussion of the merits and liabilities of using this data in the taxonomy development process. Some additional background information on the city and region is also presented for context. This appendix begins with the aspects of the chosen methodology that make the taxonomy development process robust to the choice of input data. Next, discussion turns to how the taxonomy meta-characteristic aligns well with the specific format and content of the Portland data. Brief historical sketches of Portland and the Pacific Northwest region follow. Census data is then used to profile the current demographics of the city.

Methodological Considerations

The approach to taxonomy development presented in this manuscript could have used KPI data from one of many other communities or infrastructure systems. There is sufficient detail in the description of the taxonomy development process for the reader to construct a similar taxonomy artifact from alternative data sources. While outcomes would vary with a different choice of initial input data, there is reason to expect a degree of harmony between the resulting artifacts. At the root of this expectation is the maintenance of the chosen meta-characteristic—*structural determinants of transaction costs that shape the accessibility, reusability, and analytical utility of KPIs in public and infrastructure sectors*—guiding taxonomy development to the same end.

Furthermore, the selection of a unified data resource in common spreadsheet format offers consistency of inputs to process. So long as the chosen input data contains ample metadata on the methodology, documentation, and strategic application of real (not theoretical) KPIs, the taxonomy artifact produced should be robust to the selection of a given system and setting.

The choice of Portland, OR to provide the initial KPI input data is not intended as an endorsement or criticism of any specific form of governance, performance management or its communication with the public.[5] The data selection is not presented as representative of the extant public data on performance management of public resources or infrastructure systems. Instead, it is acknowledged that the specific publication format and documentation of KPI data was conducive to the study of the chosen meta-characteristic.

There is good alignment between the chosen taxonomy meta-characteristic and the published format of the performance measurement data for Portland. A critical manifestation of this alignment is the minimized transaction costs of obtaining the inputs necessary for taxonomy development. For thoughtful discussion of transaction costs as both a concept and subfield of economic study, see [12, 49-51].

Historical Considerations

While reasonable observers might debate the degree to which Portland captures common characteristics of the present American municipal experience, the historical path of the city's

[5] For transparency, the principal author of this manuscript lives in Portland, OR making data accessibility relatively simple.

development contains threads shared with many other large cities of the Pacific Slope. The context provided by Portland's location in Oregon is formative with respect to the city's history, development, and current condition. One comprehensive atlas of the state captures the essence of the region as follows.

> "Oregon's physical environment is shaped by its position at the edge of the continent. The clockwise rotation of the cold Northern Pacific ocean currents and mid-latitude air circulation from west to east combine to give Western Oregon a milder "oceanic" climate than most places of similar latitude enjoy. The coastal region is part of the largest temperate rain forest on earth. The continuing tectonic collision of the North American and Pacific Plates along with the resulting volcanic activity has created the Cascades, confining the damp moderate climate to a narrow belt and giving the interior a much drier and more extreme "continental" climate. This tectonically controlled climate shift is the single dominant factor influencing nearly every natural system discussed in this book" [52].

National Expansion and Founding Context

A decade before its incorporation as a city on February 8, 1851 at the furthest point up the Willamette River accessible to oceangoing ships [52], the land that would become Portland was covered in Douglas fir forests [53]. Infrastructure access to local forest and agricultural products, as well as distant regional and international markets proved crucial to the Portland's development into the state of Oregon's largest city. By 1908 this infrastructure would grow to include four bridges across the Willamette River [52].

Peer Municipalities

One profile of Portland, OR found that with respect to many social and economic indicators the city shares much with Seattle, WA [54]. That this similarity carriers over to the midwestern cities of Indianapolis, Kansas City, and Columbus, reflects the region's history of settlement. "Most of the early settlers of the Willamette Valley came from the Ohio and Missouri valleys, making Oregon a far finger of the Middle West" [54].

Development

The Army Corps of Engineers early work to improve the navigability of the Columbia River enhanced the economic opportunities of a growing Portland [55]. In time, the large-scale dam and other infrastructure projects would shape many aspects of the region's interaction with the significant water resources that makes it so lush and prosperous. These water endowments created the opportunities for the timber industry which contributed to the growth of cities and infrastructure across America's west coast.

Current Census Demographics

As of the 2020 Decennial Census, Portland, OR is a city of 652,503 people in 302,034 housing units spread over 133.4 square miles.[6] The 2023 American Community Survey (ACS) 1-Year Estimates place median household income at $86,057 with 55.3% of people holding bachelor's

[6] https://data.census.gov/profile/Portland_city,_Oregon?g=160XX00US4159000

degrees or higher. The 2023 ACS estimated median age at 39.4 and the share of the population over 65 at 14.9 percent. The wider Portland-Vancouver-Hillsboro, OR-WA Metro Area has a population of 2.5 million people.[7]

Community Resilience and Public Safety

Matters of community resilience and public safety are highly salient to many living in Portland, though the framing and monikers associated with these subjects differ across stakeholder groups. The city has recent experience with wildland urban interface fires approaching outlying communities and blanketing the streets in hazardous smoke [56]. Civil unrest in the summer of 2020 and the economic and public health challenges of the COVID-19 pandemic influenced public perception and experience of public safety and community resilience. While the specific narratives that emerged from various cities differed, the shared presence of these formative experiences supports the view that for the purpose of taxonomy development, Portland may be sufficiently representative of American municipalities.

Evolving Governance

Today, Portland, OR is actively evolving its governance structure. Public performance measurement data publication is likely to evolve with its underlying subject matter.

[7] https://censusreporter.org/profiles/31000US38900-portland-vancouver-hillsboro-or-wa-metro-area/

Appendix B. Public Safety and Community Resilience Focus

The domains of public safety and community resilience are chosen as initial foci for the KPI taxonomy development due to their common importance to people across the country. Rural, suburban, and urban communities of all sizes are home to those charged with maintenance of public safety and the resilient operations of critical infrastructure in an environment of evolving hazards. The economic competitiveness of organizations, regions, and the nation itself is directly influenced by the quality of institutions and infrastructure that determine the cost structures of doing business and civic engagement. Low public safety or inadequate infrastructure resilience diminishes value propositions as individual people and organizations must approach large scale societal challenges with solutions cobbled together at inefficient scale. By contrast, when public safety and infrastructure resilience are strong, less value is dissipated on the tasks of securing safe and reliable access to markets and public services. That is, public safety and infrastructure resilience help mitigate the escalation of transaction costs. People and industry seek the low costs of living and doing business in communities where the state of public safety and infrastructure resilience do not present a set of persistent burdens. As such, small improvements in the performance management of public safety and community resilience can have a substantial positive impact on economic competitiveness and well-being.

Ensuring infrastructure and public service value delivery requires a strategy for performance management. In many domains, industry leads government in the formulation and execution of performance management strategy. However, the public documentation and explanation of those strategies is not of uniform quality. This taxonomy has placed an emphasis on the accessibility of such documentation regarding KPIs through the URL, strategic plan, collection method, and formula dimensions. Where KPIs are clearly and publicly linked to resources explaining how they are conceived, implemented, and to what ends, the public benefits. This taxonomy orders KPIs according to how well they are documented within the simplest raw data resources available. Ideally, consolidated KPI data for infrastructure and public services are directly linked to the metadata necessary to understand provenance and context. Without sufficient detail in this regard, it is dangerous to compare similar KPIs across domains, regions, or organizations. Empirical study of infrastructure performance or public service delivery is costlier when such metadata is dispersed and disorganized. If the cost of developing evidence-based policy and operational decisions can be reduced through better inputs to empirical analysis, the cost of getting better outcomes for people may fall. The achievement of conspicuously better outcomes for stakeholders is likely among the fastest paths to wider adoption of any given KPI.

KPIs can be developed and implemented to help guide the decisions of infrastructure owners and operators towards the realization of greater value creation. The expansion of value comes through both increased benefit delivery and effective cost management. This taxonomy is intended to assist in the evaluation of KPIs currently in use and their improvement. This improvement and its visibility to potential adopters is crucial to expanding the effective use of KPIs for public benefit. The visualization method explored in this document makes improvements to the level of KPI documentation immediately perceivable as the profile of the KPI increases in radius.

Appendix C. Taxonomy Applied to Sample KPI Data

The taxonomy was applied to 40 key performance measures from 2022 for Portland, OR. Table 8 presents each of these 40 measures along with the assigned characteristic for each dimension of the taxonomy. For a summary of this data, see Table 7.

Table 8: KPI Taxonomy Sample Data

Measure Title	D01	D02	D03	D04	D05	D06	D07	D08	D09
Percentage of emergency 9-1-1 calls answered within 20 seconds	EC	9-1-1 Operations	Outcome	Transactional Collection	Up	Yes	Specific	Database Pointer	Complex
Percentage of emergency 9-1-1 calls answered within 15 seconds	EC	Unassigned	Outcome	Monthly Collection	None	No	None	None	None
Percentage of overall operations staff (including trainees) retained	EC	Unassigned	Outcome	Intermittent Collection	None	Yes	None	None	None
Percentage of certified operations staff (non-trainees) retained	EC	Unassigned	Outcome	Intermittent Collection	None	Yes	None	None	None
Percentage of completed improvement plan tasks completed within agreed upon timeframe	EM	Emergency Management	Outcome	Annual Collection	Up	Yes	Vague	Detailed Description	Complex
Percentage of PBEM plans that are up-to-date according to their published standards	EM	Emergency Management	Workload	Annual Collection	Up	Yes	Vague	Detailed Description	Complex
Number of new PublicAlerts registrations	EM	Emergency Management	Output	Annual Collection	Up	Yes	Vague	Database Pointer	Complex
Percentage of neighborhoods	EM	Emergency Management	Outcome	Annual Collection	Up	Yes	Vague	Detailed Description	Complex

with active NET teams.									
Percentage of participants who rate PBEM classes and exercises as good or excellent	EM	Emergency Management	Outcome	Annual Collection	Up	No	Vague	Detailed Description	Complex
Number of participants in a Portland Bureau of Emergency Management class and exercise annually	EM	Emergency Management	Outcome	Annual Collection	Up	No	None	Vague	Simple
Number of hours completed by participants in a Portland Bureau of Emergency Management class and exercise annually	EM	Emergency Management	Workload	Annual Collection	Up	No	None	Vague	Simple
NET Program Diversity	EM	Emergency Management	Output	Annual Collection	Up	No	None	Vague	Complex
Outreach to historically underserved communities	EM	Emergency Management	Output	Annual Collection	Up	No	None	Vague	Complex
BEECN Program Deployment Readiness Index	EM	Emergency Management	Output	Annual Collection	Up	No	None	Vague	Complex
Number of sanitary sewer overflows	ES	Pollution Prevention	Outcome	Quarterly Collection	Down	Yes	Vague	Database Pointer	Complex
Watershed Health Index for water quality	ES	Watershed	Outcome	Intermittent Collection	Up	Yes	Vague	Detailed Description	Complex
Number of combined sewer overflow events	ES	Pollution Prevention	Outcome	Quarterly Collection	Down	Yes	Vague	Simple Description	Simple
Average single family household bill	ES	Administration & Support	Outcome	Annual Collection	Down	Yes	Vague	Detailed Description	Complex

as a percent of median income									
Percentage of urgent public health and safety related service requests responded to within two-hour timeframe	ES	Wastewater	Efficiency	Quarterly Collection	Up	Yes	Vague	Detailed Description	Complex
Percentage of structural fires where flamespread was confined to room of origin	FR	Emergency Operations	Outcome	Transactional Collection	Up	Yes	Specific	Detailed Description	Complex
Number of civilian deaths due to fires	FR	Emergency Operations	Workload	Transactional Collection	Down	Yes	Specific	Detailed Description	Simple
Average travel time to high priority dispatched calls in minutes	PL	Emergency Response & Problem Solving	Outcome	Annual Collection	Down	No	None	Detailed Description	Complex
Number of Crime Against Persons offenses per 1,000 residents	PL	Emergency Response & Problem Solving	Outcome	Annual Collection	Down	Yes	None	None	None
Number of Crime Against Property offenses per 1,000 residents	PL	Emergency Response & Problem Solving	Outcome	Annual Collection	Down	Yes	None	None	None
Percentage of sworn members who identify as female and/or a person of color	PL	Human Resources Development	Outcome	Annual Collection	Up	Yes	None	Detailed Description	None
Percentage of investigated complaints that are sustained (excluding use of force complaints)	PL	Employee Performance	Outcome	Annual Collection	Down	Yes	None	Detailed Description	None
Number of community complaints of officer misconduct	PL	Employee Performance	Outcome	Annual Collection	None	Yes	None	Detailed Description	None

Number of community commendations of officer conduct	PL	Employee Performance	Outcome	Annual Collection	Up	Yes	None	Detailed Description	None
Average call queue time until a responding officer is available (high priority calls)	PL	Emergency Response & Problem Solving	Efficiency	Annual Collection	Down	No	None	None	None
All Priority Dispatch Calls for Service Average Response Time (in minutes)	PL	Unassigned	Workload	Annual Collection	None	No	None	None	None
High Priority Dispatch Calls for Service Average Response Time (in minutes)	PL	Unassigned	Outcome	Annual Collection	None	No	None	None	None
Low Priority Dispatch Calls for Service Average Response Time (in minutes)	PL	Unassigned	Outcome	Annual Collection	None	No	None	None	None
Medium Priority Dispatch Calls for Service	PL	Unassigned	Workload	Annual Collection	None	No	None	None	None
Reported NIBRS Group A Person Crime Offenses	PL	Unassigned	Workload	Annual Collection	None	No	None	None	None
Reported NIBRS Group A Property Crime Offenses	PL	Unassigned	Workload	Annual Collection	None	No	None	None	None
Percentage of Portlanders Living in a Complete Community	PN	Comprehensive & Strategic Planning	Outcome	Annual Collection	Up	Yes	Specific	Detailed Description	Complex
Percentage of waste recycled or composted	PN	Waste Reduction & Recycling	Outcome	Annual Collection	Up	Yes	Specific	Detailed Description	Complex
Percentage of seven-county region’s new	PN	Comprehensive & Strategic Planning	Outcome	Annual Collection	Up	Yes	Specific	Detailed Description	Complex

housing that is in Portland.									
Percent of undeveloped floodplains outside the levee system that are subject to flood storage capacity and vegetation replacement requirements	PN	Unassigned	Outcome	Transactional Collection	Up	No	Specific	Detailed Description	Complex
Percentage reduction in per person carbon emissions from 1990 levels	PN	Policy, Research & Innovation	Outcome	Annual Collection	Up	Yes	Specific	Detailed Description	Complex