Title: Leveraging Smart System Technologies in National Energy Data Systems: Challenges and Opportunities

Authors:
Eric Masanet
Luis Munuera
Roberta Quadrelli
Duncan Millard

Publication Date:
(insert date)

Series:
EEG State-of-Knowledge Paper Series

Permalink:
(insert link)

Keywords:
Electricity, economic development, economic growth, energy policy

Abstract:
Effective energy policies rely on credible and comprehensive national energy data systems, in developed and developing economies alike. Smart system technologies will play a central role in the clean energy transition—including applications in smart homes, factories, transport systems, and renewable electricity grids—and their ability to compile and communicate point-of-use energy data presents new opportunities for improving national energy data systems. This paper reviews the growing importance of energy data systems for energy policy in the developing country context, identifies key characteristics a national data system needs to have in order to be robust and viable, discusses the potential role of smart system technologies in national energy data systems moving forward, and recommends several future research areas to better understand their potential, and in developing countries in particular.

Energy and Economic Growth
Applied Research Programme

Series Name: EEG State-of-Knowledge Paper Series
Paper No.: Data
Issue Date: (insert date)

Leveraging Smart System Technologies in National Energy Data Systems: Challenges and Opportunities

Eric Masanet, Luis Munuera, Roberta Quadrelli and Duncan Millard

EEG State-of-Knowledge Paper Series

Oxford Policy Management
Center for Effective Global Action
Energy Institute @ Haas

UKaid from the British people

Oxford Policy Management

The Applied Research Programme on Energy and Economic Growth (EEG) is led by Oxford Policy Management in partnership with the Center for Effective Global Action and the Energy Institute @ Haas at the University of California, Berkeley. The programme is funded by the UK Government, through UK Aid. Over the course of five years, EEG will commission rigorous research exploring the links between energy economic growth, and poverty reduction in low-income countries. This evidence will be specifically geared to meet the needs of decision makers and enable the development of large-scale energy systems that support sustainable, inclusive growth in low income countries in South Asia and Sub-Saharan Africa.

The EEG Working Paper Series showcases the 18 State-of-Knowledge papers produced in Phase 1 of the EEG programme. These papers address gaps in knowledge and data in six critical themes related to energy and economic growth: the links between electricity supply and growth, finance and governance aspects, large scale renewables, sustainable urbanization and energy efficiency, the role of extractives, and design of energy systems. Cross-cutting themes of gender, climate change, and data are also addressed in dedicated papers.

Leveraging Smart System Technologies in National Energy Data Systems: Challenges and Opportunities

Applied Research Program on Energy and Economic Growth State-of-Knowledge Paper

Eric Masanet, International Energy Agency, Eric.MASANET@iea.org
Luis Munuera, International Energy Agency, Luis.MUNUERA@iea.org
Roberta Quadrelli, International Energy Agency, Roberta.QUADRELLI@iea.org
Duncan Millard, International Energy Agency, Duncan.MILLARD@iea.org

Abstract

Effective energy policies rely on credible and comprehensive national energy data systems, in developed and developing economies alike. Smart system technologies will play a central role in the clean energy transition—including applications in smart homes, factories, transport systems, and renewable electricity grids—and their ability to compile and communicate point-of-use energy data presents new opportunities for improving national energy data systems. This paper reviews the growing importance of energy data systems for energy policy in the developing country context, identifies key characteristics a national data system needs to have in order to be robust and viable, discusses the potential role of smart system technologies in national energy data systems moving forward, and recommends several future research areas to better understand their potential, and in developing countries in particular.

The Applied Research Programme on Energy and Economic Growth (EEG) is led by Oxford Policy Management in partnership with the Center for Effective Global Action and the Energy Institute @ Haas at the University of California, Berkeley. The programme is funded by the UK Government through UK Aid. The views expressed in this paper do not necessarily reflect the UK Government's official policies.

I. The importance of robust national energy data systems

The provision of reliable, affordable, and clean energy services is a fundamental goal of many national energy policies, and is also critical to the development agendas of the world's emerging and developing economies. It has long been established that effective energy policies rely on credible and comprehensive national energy data systems. Such data can help governments and other stakeholders track a country's progress toward its energy goals, enable energy policy decisions that are evidence-based, and provide accountability and transparency on a nation's energy supply, usage, and overall energy system trends (OECD 2015).

The importance of robust national energy data systems is reflected by major international initiatives and resources that aim to promote greater consistency, quality, and utility of energy data and statistics across countries, and among stakeholders from both the public and private sectors. Key examples include the United Nations' *International Recommendations for Energy Statistics* (IRES), the Joint Organisations Data Initiative (JODI), the International Energy Agency's (IEA's) *Energy Statistics Manual* and *Energy Efficiency Indicators Manuals*, the Partnership in Statistics for Development in the 21st Century (PARIS21)[1] , and the United Nations' work to develop indicators for its Sustainable Development Goals (SDGs) (UNDESA 2016; IEA 2014 a,b; IEA and EUROSTAT 2004; JODI 2016).

Given the pivotal role that national energy data can play in developing countries—which at the same time often lack the institutional, financial, and knowledge capacities for robust national systems—several international initiatives have been aimed at capacity building, training, and knowledge sharing to governments and statistical agencies in emerging and developing countries. Recent examples include, the IEA *Energy Statistics Courses*, the energy statistics initiative of INOGATE, which is focused on capacity building in Eastern Europe, the Caucuses, and Central Asia and is now taken forward by the IEA EU for Energy Programme, and the United Nation's Global Assessment of Energy Statistics and Balances (INOGATE 2016; IEA 2016a; UNSD 2008).

National energy data vary greatly with respect to geographical coverage (e.g., national, regional, and local scales), temporal coverage (monthly, annual, and multi-annual), sectoral coverage (energy supply, transformation, and demand by end use sectors), reporting level (e.g., by fuel type, by fuel sub-type, by end use, and by technology), and the coverage of various other energy system indicators. Exemplary national energy data system resources can often include, for example:

- **Energy balances**, which typically cover the production, import, and export of primary energy sources, transformation into fuels for final consumption, and final consumption by various high-level demand sectors, including the residential, commercial, industrial, agricultural and transport sectors.

- **Energy efficiency indicators**, which express the relationship between end-use consumption and the various physical and economic activities in a country that drive that consumption within each major demand sector (IEA 2014a,b).

[1] The Partnership in Statistics for Development in the 21st Century (PARIS21, www.paris21.org) is a global partnership of statistical producers, users, donors and technical partners, from both developed and developing countries. PARIS21 works on improving national statistical capacity and co-ordination while advocating for the better use of statistics in the decision-making process at national, regional and international levels.

- **Energy-related environmental emissions**, which can include carbon dioxide (CO2), other combustion-related emissions (such as criteria air pollutants), and non-CO2 greenhouse gas (GHG) emissions.

- **Fuel prices and taxes**.

- Public energy technology research, development, and demonstration (**RD&D**) investments.

- **Development indicators**, such as national and local levels of clean energy access or use of traditional biomass for home heating and cooking purposes.

In developed and developing countries alike, national energy data systems serve multiple energy policy activities involving both public and private stakeholders. For example, policy makers need a detailed understanding of the current state of national energy supplies, energy demand by sector, usage trends, prices, and energy efficiencies to craft policies that deliver energy security, energy access, and efficient, affordable energy services; these data subsequently enable tracking progress toward these goals and assessing the effectiveness of policies over time.

Energy modellers rely on accurate country-, sector-, fuel-, and technology-related data for analysing technology and policy options for meeting national energy, climate, and environmental goals over different time scales. Investors and development banks require credible and reliable national energy data for identifying technology RD&D opportunities that are likely to have the greatest impacts on energy and economic goals, and, therefore, the greatest chances of market success, whilst the public have a need for good data to make inferred choices and to see the impact of government policies. Barriers to robust national energy data systems in developing countries, which are discussed further in Section II, present significant obstacles for sound and efficient decision making by these important policy making, energy modelling, and investment communities.

Robust national energy data systems are also becoming more important for tracking the progress of developing countries toward major international initiatives and agreements aimed at sustainable development and international climate action. For example, the United Nations' SDGs aim to end poverty, protect the planet, and ensure prosperity for all (UN 2016). In particular, SDG7 (i.e., Ensure access to affordable, reliable, sustainable and modern energy for all) relies on indicators related to energy access, renewable energy shares, energy efficiency levels, and energy system and technology investments derived from national energy data systems. According to the OECD, "achieving the millennium development goals will require a "data revolution" and, to achieve this, the international community needs to agree on a global statistical strategy and a global partnership to co-ordinate and deliver it" (OECD 2015).

The monumental Paris Agreement has strong provisions for tracking and reporting national GHG inventories and progress toward nationally-determined contributions as part of its compliance mechanisms.[2] The required GHG inventories and progress indicators involve energy-related emissions that must be derived from national energy statistics and balances (UNFCCC 2015). Finally, national energy data and statistics are an important part of measuring and monitoring the overall concept of "green growth" in developing countries, for which organizations such as the OECD and United Nations have proposed national energy-related metrics for credible tracking (OECD 2014; UNDESA 2014).

While the comprehensiveness and level of detail of national energy data vary by country, historically a general principle has been that, as the complexity of data collection, reporting scales, and reporting detail increases, so, too, does the required institutional capacity and costs of the national data system. This relationship is depicted schematically for the case of national energy efficiency indicators in Figure 1, which depicts various levels of data disaggregation required for different levels of energy efficiency indicator design, and associated data requirements (i.e., the IEA energy indicators "pyramid").

Timeliness is another important determinant of the utility of national energy data in supporting energy policy decisions. For example, credible modelling, progress tracking, and reporting all rely on having accurate data that reflect the most recent state of an energy system, especially in emerging and developing countries where the energy system can evolve quite rapidly. However, the process of national data collection, processing, validation and dissemination requires an inherent time lag; whilst increasing the frequency of data collection and reporting would add complexity and cost to a national energy data system.

However, the growing use of so-called smart system technologies[3] across the energy system—with notable examples including smart homes, commercial buildings, factories, transport systems, and cities, to name a few—may present opportunities for generating myriad point-of-use energy data that can enhance and improve national energy data systems. While many technology components used in smart energy systems, for example, smart meters in homes or power meters in industrial facilities (O'Driscoll and O'Donnell 2013) are undoubtedly already generating data used in traditional national energy data collection reporting systems (e.g., surveys), the rapidly growing presence of smart systems for energy and economic productivity may lead to vast troves of new point-of-use data.[4] As discussed further in Section III, these technologies could potentially play an important role in building up energy data collection capacity in developing countries while achieving other important development goals, such as improving clean energy access through smart grids.

Furthermore, smart system technologies are emerging as a cornerstone of integrated and clean urban energy system designs, and may therefore see wide usage in developing countries

---

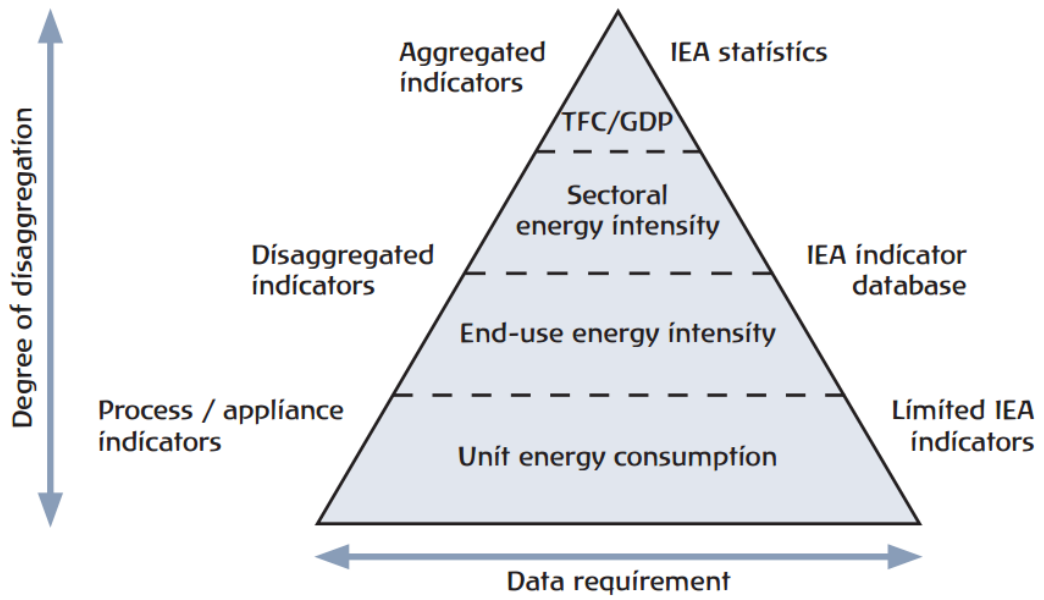
[2] The Paris Agreement text states that each party shall regularly provide "Information necessary to track progress made in implementing and achieving its nationally determined contribution under Article 4." (UNFCCC 2015).

[3] While there are numerous definitions of smart systems, in this paper the term "smart system technologies" refers to information and communication technologies (ICT) applied to the specific purpose of continuously controlling the state a physical system to deliver a desired service, inclusive of energy services. ICT can be further defined as systems whose fundamental functions are anchored in the generation, processing, storage, communication, and/or presentation of digital information (Masanet and Matthews, 2010).

[4] For example, adoption rates of control systems for building environments and energy-intensive process equipment in U.S. industrial facilities have grown significantly, mainly driven by energy efficiency, cost savings, and productivity considerations (U.S. EIA 2013, 2010, 2004, 2000). However, their deployment also enables ready access to detailed plant and equipment-level energy use information for reporting purposes at U.S. industrial plants.

(whose populations are urbanizing rapidly) as they seek to redirect urbanization trends toward new models of more sustainable, interconnected, and liveable cities (IEA 2016).

Figure 1. Schematic representation of the IEA energy indicators pyramid



Notes: unless otherwise indicated, all tables and figures in this publication are derived from IEA data and analysis. TFC = total final consumption.

Source: IEA 2014a,b

II. Elements of robust national energy data systems and requirements for data quality

While smart system technologies may hold potential for improving national energy data systems, developing countries can face multiple barriers and challenges to implementing robust and viable national energy data system infrastructures. This section provides a brief summary of how national energy data systems are constructed, their key elements, and primary challenges. It also provides a brief overview of statistical design and data quality characteristics, which may present both challenges and opportunities in the use of smart system data to strengthen the national data systems. In Section III, some possible opportunities for leveraging smart system technologies are discussed. In Section IV, several research priorities are proposed to explore how smart systems might play a role in developing and strengthening robust national energy data systems in the developing country context.

Historically, each country has developed its own specific statistical system to respond to the various requirements (e.g. priority policy questions from government, business, investors and the public; international or regional data requirements). The way that national statistical frameworks—and their energy components—are organised varies a lot, both across developed countries and more globally. A key upfront question is how energy fits within the overall policy landscape, and how energy statistics fits within the overall national statistical system.

At the national level, the effective functioning of the energy statistics system would depend on a sound **data governance model** for overall statistics (i.e., the existence of a legal framework, with statistical acts, mandatory reporting requirements, etc., although this may not be enough, especially in developing countries, where enforcing it may be difficult), the existence of standardised methodologies and definitions and the adoption of statistical standards; transparent dissemination of data; and protection of confidentiality. All these, together with professional independence of statistical agencies, scientific competence of their staff and impartiality, are prerequisites for official statistics to be **trusted** by business and the wider public (UN 2013).

Clearly, **resources** are essential to the establishment and to the maintenance of a national energy statistics system. A number of questions are relevant to an effective resource allocation: how energy data are prioritised vs other statistical areas; how synergies can be exploited between energy and other policy areas; what priority is given to the monitoring component within the policy formulation process; and how resources are allocated across areas within the energy domain. The adequacy of resources is likely to be more effectively assured if strategic planning and prioritisation, as well as standardisation of procedures and tools, are in place.

Within the energy statistics field, the institutions with key responsibilities at the national level may be National Statistical Offices, but also Ministries (of Energy, Economy, Development, Finance), Regulators, Energy Agencies and Energy Efficiency Agencies (the latter, generally of more recent establishment). Solid **institutional arrangements** among the key stakeholders are a central point to ensure the robustness and flexibility of the energy data system. Under this area, there is a value in presenting country good practices on the benefits/barriers of different types of structure (e.g. centralised/decentralised/provincial); on the roles and responsibilities between producers/users of specific information, and how additional value is generated from the establishment of cross-cutting working groups and the sharing of information across government entities (by MOUs, etc). Depending on the national situation, the collection of new data may require the expansion of an existing institutional framework or the development of a new one.

The type of institutional arrangements and energy statistics system in place will also affect the **flexibility** to react and provide data to changing policy priorities, and to absorb the adoption of innovative technologies when they become available.

Energy statistics at the national level are generally derived through a number of **data collection and compilation practices**, including: surveys (of energy providers, households, businesses, etc.), use of administrative data (e.g. data collected from other institutions or for different purposes, with associated challenges of boundary definitions and data sharing, and modelling tools. One typical example of effective use of administrative data to monitor a specific policy is the data-matching exercise performed in the UK (BEIS, 2013) to track progress in gas and electricity savings at the individual address level due to insulation implementation.

One very practical aspect to consider, especially in developing countries, is the availability/choice of the **IT infrastructure**. Key questions include: are the available technological resources suitable to support the statistical production process? How to design a system to collect/store data? Can a given data collection be implemented online? What tools are used in other countries to be effective?

At the technical level, it is very important to develop and maintain **skills and knowledge** among the statisticians who are responsible to collect, compile, and disseminate the data. The international work of defining **harmonised methodologies** on energy statistics (e.g. the recently agreed International Recommendations on Energy Statistics at the United Nations level) is very relevant. One important example of ongoing international work is better defining the renewable sources (e.g. various forms of bioenergy) for which data need to be collected, including discussing the additional aspect of sustainability of their use, which has traditionally not been included in energy statistics definitions.

Capacity building in the technical area is essential (e.g., through webinars, training events, production of guidebooks, and online material), leveraging on the current work of regional and international organisations in this domain.

Table 1 summarizes these dimensions, with brief summaries of possible challenges faced in the developing country context.

While the above obstacles to establishing robust national energy data systems may be particularly pronounced in developing countries, smart systems may present their own set of challenges as an emerging resource for national energy data in all countries. Table 2 summarizes key statistical design and data quality dimensions identified for national data systems (UNSC 2012, ONS 2013; Statistics Canada 2009), as well as the associated potential opportunities and barriers related to smart systems data. Building an improved understanding of the severity of each barrier—and possible opportunities for overcoming them—across the range of point-of-use data that may be made available through smart systems in myriad applications, is an important area of future research for the smart system technology, energy policy, and energy statistics communities.

Table 1: Matrix of potential challenges for developing countries related to overall energy statistics

Dimension	Possible challenges
Energy within overall policy priority	May be not enough; not in synergy with other policies (multiple benefits)
Overall data governance	Lack of trust; independence from policy makers; lack of legal framework; poor or lacking statistical standards
Resources for energy statistics	May be poor; ineffectively allocated across areas; competing areas
Institutional arrangements/ flexibility	Structural inefficiencies; unawareness of other work; competing priorities for different stakeholders; multiple data requests; lack of data sharing
Data collection practices	Not effective/well designed; lack of training; lack of expertise; not based on innovative tools; inappropriate validation tools; lack of resources
IT infrastructure	General availability (e.g. internet); lack of expertise; lack of resources
Methodologies	Lack of national standards; unawareness of / discrepancy of national system with international methodologies; lack of training

Table2: Key statistical design and data quality dimensions for leveraging smart system energy data into official energy statistics

Dimension		Description	Potential Opportunities	Potential Barriers
Relevance		The degree to which a measure meets current and emerging users' needs	Smart system data would be very granular and frequent, and could meet broad client needs (e.g. high resolution geographical disaggregation, etc).	The risk of too much irrelevant data must be managed through proper system design. Data likely to be end-use driven, so unlikely to be complete (e.g., the transformation sector may be underrepresented).
Accuracy		The degree to which data correctly describe the phenomena they were designed to measure	Well-designed and well-functioning smart systems could generate data without the typical errors of survey-derived information (e.g., respondent errors, non-response bias, data processing errors and imputation errors)	Mechanical errors could introduce measurement errors. As with all Administrative systems, statisticians need to be involved from the start in defining any data needs to ensure relevance.
	Sampling		As smart systems proliferate, over time larger sample sizes may become more practical.	Data from smart systems may be not as well-organized as a controlled sample. An important skill to be developed is in applying statistical theory to less organized data to obtain representative data.
	Coverage	Non-coverage errors arise when segments of a population are not sampled, leading to a sample that is not representative of the population	As smart systems proliferate, over time coverage may improve.	Smart systems may be installed in a non-representative way , either initially or by the nature of their market applications, limiting their potential use in national-level data systems. For example, smart homes may be first deployed in homes within wealthier segments of a population, while smart factory systems may be installed in only large industrial plants with greater financial resources than small industrial plants. When coverage-errors are present, smart system data must be used with limitations in mind and augmented with data collected through traditional methods.

Dimension		Description	Potential Opportunities	Potential Barriers
	Non-response	Non-response errors arise when a portion of the sample does not respond	The potential ubiquity of smart systems might enable very large sample sizes, which may reduce the impacts of non-response errors.	Smart systems may experience reliability issues due to equipment failures, communications issues, or other technical problems.
Timeliness		The frequency of sampling and the duration of data analysis and processing should enable timely use of data	The continuous operation of smart systems may enable high frequency of data collection.	Large volumes of data needing processing, harmonization, and validation might lead to delays in data availability. Upfront design of data collection and communication standards and protocols may improve data quality, leading to reduced processing and validation times.
Accessibility and clarity		The ease with which users can obtain relevant data	As smart systems proliferate, well-designed data systems may offer improved accessibility to a wide variety of energy system data.	Format of output needs to be well designed to fit the purposes. Metadata is essential.
Coherence and comparability		Statistics need to be consistent internally and comparable over time and across sources, to be suitable for multiple uses.		Smart systems can vary greatly in the data they sense, monitor, collect, analyse, and communicate, which may pose significant challenges for designing standards and exchange protocols across millions of installed devices made for wide range of purposes by a diverse group of stakeholders. Similarly, the end uses of national energy data from smart systems may evolve rapidly, as new information leads to new uses and new demands for data types and formats. Managing the complexity of abundant but disorganized data may be a critical challenge for ensuring data consistency at reporting levels required in national energy data systems.

Dimension	Description	Potential Opportunities	Potential Barriers
Other considerations:			
Validation	The quality and conformance of collected data can be validated against desirable criteria to ensure suitability for intended end uses	Smart system data are inherently digital, which may enable efficient and robust validation algorithms with the proper upfront data validation protocols.	Similar to the challenge of ensuring consistency, the complexity of validating the quality of abundant but disorganized data from myriad smart systems may prove challenging to traditional data validation systems and approaches. Significant upfront effort may be needed to design data standards, protocols, and exchange platforms that ensure ease of validation prior to dissemination, perhaps with algorithms and analyses at various steps in the smart system data chain that ensure data delivered to statistical agencies is of sufficient quality.
Confidentiality	Privacy and confidentiality are protected in data collection, processing, dissemination, and storage	Legal measures could protect the confidentiality of smart system data, as they do for other types of collected data.	The collection of point-of-use energy data from public and private entities may present challenges for protecting privacy and confidentiality, and public and private sector concerns and perceptions may prove to be cultural and legal barriers to national data collection. Additionally, storage and archiving of such data may need high levels of data security to prevent inadvertent or malicious access. The risk of cybersecurity breaches may be present across these systems, requiring strong upfront and continuous data protection initiatives.

As clear from the table, integrating raw operational data from multiple sources into national statistics involves a number of challenges at all levels, also including the legal and institutional ones. Research is already developing on the integration of multiple data sources into general official statistics (UNECE, 2016), and more particularly on the potential to integrate information from smart systems into national energy statistics, as deployment of smart meters increases: for example in Europe (ESSnet, 2016); and Canada (Ma, 2014).

This research has formulated initial recommendations on the importance of cultivating relationships between statistical offices and utilities and all relevant smart system data suppliers; to establish partnerships with key stakeholders to facilitate data acquisition; and to design effective data storage and data processing capacity to accommodate new data. More research will be needed to strengthen these findings and to assess the potential benefit of smart system data and how they can be combined with traditional survey data in producing comprehensive energy statistics.

III. Emerging opportunities for energy data systems in developing economies

Traditionally, energy data collection has been restricted to fuel sales and trade data, surveys, or accounting of stocks and flows, and mainly on the supply side. Projections of energy demand for planning and policy development have been carried out using broad relationships to high-level macro-economic data.

Across every sector within the energy system, ICT technologies and pervasive sensing are opening up new possibilities for vast amounts of data flows to enhance planning and operation, to improve information on end-use consumption and to involve consumers to a much greater degree. Energy utility spending on data analytics is set to reach $3.8bn by 2020, with OECD utilities expected to spend as much as $100 per home on ICT analytics in the same period.

For example, utilities have been deploying smart meters in ever growing numbers, spurred by the promise of improvements in operational efficiency, reliability and customer service. With improved demand data, utilities can better incorporate variable generation assets such as renewables and deploy innovative programs that allow customers to better control their energy usage while simultaneously reducing capacity requirements.

At the same time, as monitoring is deployed throughout the energy system, innovative opportunities are created to reduce the cost, decrease losses, and increase the efficiency of energy supply. For example, sensors attached to smart heating and air conditioning equipment are informing operators of consumer behavior when faced with weak or unreliable grids. In India, selective load shedding, such as turning off the load for air conditioners or water pumps in peak hours across nearly 90,000 households is currently on trial (Selvam et al. 2016).

The resulting proliferation of real time data also creates opportunities for improved and expanded collection of energy flow data across the entire energy system.

Data to track policy objectives

Beyond smart meters, increasingly, real-time data are allowing for a better understanding of how technologies operate in a real-world environment, and whether they are fulfilling policy objectives. Across the developed world, field trials of new demand-side technologies are becoming more common. In the power sector, LIDAR, smart inverters or wind farm monitoring systems are sending much more accurate information, which greatly helps utilities understand how renewables operate in real-world conditions and allows deployment at the lowest cost.

A large number of examples exist in the developed world. In the UK, the use of household consumption data and data linking helped refine the incentives given to promote renewable heat and energy efficiency in homes. At the same time, these data are allowing the tracking of energy policy objectives, such as the reduction in residential energy demand, or the efficiency of retrofit insulation (Hawkes et al. 2011). Location-referenced consumption data has allowed for the development of maps to target potential locations for new technologies in urban areas, including rooftop solar PV (ADEME 2014), district heating and cooling (DECC 2013), or targeted building retrofits.

In the developing world, cooking stoves are widely used to prepare meals and provide heat indoors. The IEA's World Energy Outlook estimates 4.3 million deaths a year, around 80% of these in Asia where 1.9 billion people rely on traditional biomass for cooking. Enhanced cookstoves are being deployed through a wide number of projects. The "Global Alliance for Clean Cookstoves" produces guidelines on proper design, and aims to reach 100 million homes adopting cleaner cooking by 2020. As part of the initiative a platform for Integrated Cookstove Assessment (PICA) has been created, where sensors are deployed to gather real-time data on black carbon emissions from cookstoves, to monitor and evaluate their use in real-world conditions. In Sudan, traditional survey data was compared to these advanced data measurement techniques. The study found that the survey reported 1.2 hours of daily cooking higher than real-time measured data, as well as 1.3 daily cooking events higher. The study outputs highlight how such low cost sensors can collect more reliable information than traditional methods.

Data to support new business models

In much of the developing world, grid availability, relative cost of service and high frequency and duration of service loss are key concerns. In addition, loss of revenue through non-technical losses aggravates the issue, and investment in grid infrastructure carries uncertain returns. Data-enabled distributed energy sources can improve energy access and support new business models. In Kenya for instance, companies such as M-Kopa distribute a PV and battery system with a communications tool. The initiative monitors production and storage volumes and can remotely evaluate conditions.

Throughout the developed world, data is also supporting the emergence of new business models. Among these, smart micro-grids, aggregation models and Virtual Power Plants (VPPs) hold particular promise. Under these ICT-enabled solutions, an energy utility or local system operator is able to aggregate information from many users, and regulate or shift their use by either controlling devices in homes remotely, learning to anticipate patterns of demand better from the vast amounts of information gathered, or sending price signals at the right times to incentivise behavioral change. In developed countries, this can increase efficiency and reduce peak capacity needs. In developing economies, VPPs have potential applications to reduce load shedding (Fadaeenejad et al. 2014).

Data to support efficient, sustainable urban planning and transport systems

Data is at the core of urban planning, but new data sources are dramatically improving the possibilities for planning and operating transportation systems in developing countries. Sensor-connected mini-buses in Kenya are being fitted with a device that records information on location, ignition, route, distance, speed, acceleration and deceleration, and sends it in real time via mobile network to an IT platform, which then aggregates, analyzes and visualizes the data, and allows an urban planner to optimize utilization and direct investment where needed. Owners can use this data to track their vehicle's location and productivity, and to receive alerts when unsafe driving events such as hard braking, speeding and dangerous acceleration occur.

Electric vehicles are projected to grow from 1 million on the road today, to 150 million in 2035 (IEA 2016). Over 40 million electric motorcycles are sold per year around the world, the vast majority of them in the developing world. These can help move emissions from private transport and its concurrent health impacts away from densely populated areas. The possibility for emerging and developing economies to benefit from this rapid transformation of private transport requires a radical change in transport and energy networks. Traffic and battery status data can inform power demand predictions associated with e-bikes and EVs, while pricing information can influence consumer decisions to support load balancing.

Such examples are already in place in Germany and the United States. In Munich, Germany, with the use of a mobile app, owners of electric vehicles can access real time pricing data and charge them with a minimum impact on the grid. In the United States, the utility SDG&E is piloting aggregated EV charging to participate in demand response markets. Lessons learned from data-enabled EV deployment programmes are being translated to emerging economies looking to up-scale their low carbon mobility plans. India, Indonesia and China, driven by environmental concerns, are deploying similar measures to support uptake of electric vehicles with a minimum impact on grids which are often weak or underdeveloped (Fadaeenejad 2014).

Urban planning models are also being changed thanks to new data sources. In Tokyo, the Harmonious Mobility (Ha:mo) initiative merges multiple modes of transportation and real-time data to support urban planning and efficient transportation. Data is at the core of such initiatives: these programmes leverage data on vehicle status, traffic, and likely times of arrival at stations to support a highly-interconnected network. With this information, better predictions of transport demand could enhance efficiency of urban planning itself, helping direct investments where they are most needed.

Urban planning can often be qualitatively different in more developed economies, and data solutions are emerging tailored to the densified conditions of developing cities. For instance, urban slums and crowded informal settlements are at a high risk of fires, particularly from household cooking with traditional biomass. RF-mesh networked technologies can detect and track fires in these areas, and are already being deployed in Kenya and South Africa.

Data to improve the planning and operation of energy systems

Data is helping improve the planning and operation of gas, fuel and electricity networks. In particular, new technologies hold promise for developing countries to leapfrog traditional grid build-out, and many of these options are seeing deployment in such environments.

In transmission grids, Phasor Measurement Units (PMUs), widely deployed in developed countries, are becoming more commonplace in a developing country context. India, Thailand, Indonesia and China have recognized the cost savings that can be achieved through investments in such smart data analytics. PMU data provides accurate information on power system status, allowing operators to improve grid operation and to operate the transmission lines and other components closer to their true limits. In a developing country context, where equipment is often operated close to technical limits, the savings are potentially very significant. Such technologies also provide early detection of potential failures or disturbances, and aid in restoring systems during black-outs or transitory failures. The cost of such measures, and the data centers needed to store the data they produced, has been high, but at present other barriers (awareness, local capacities) are preventing their wider use in developing countries (Nangia et al. 2016).

Low voltage (distribution) networks in urban and rural areas have traditionally been designed with a ‘fit-and-forget’ approach, which with smarter technologies, distributed generation and solar PV is becoming obsolete (Lopes et al. 2016). In distribution grids, owners and operators are installing ICT and supervisory control and data acquisition (SCADA) systems to monitor voltages and power flows along distribution feeders, and to enable the integration of renewable generation in distribution grids.

Now monitoring can be more readily integrated into a robust data system and network model, complete with detailed load estimations and dynamics. Increasingly, monitoring enables system control to adjust voltages in real time (sometimes called “distribution automation”) such as tap-changing transformers. Historically, such monitoring and control was neither necessary nor economical at the distribution level. Examples abound in the developed world. In the United States, the utility Commonwealth Edison has invested significantly in distribution automation since 2011, and the data analytics help isolate outages and reduce duration for customers.

IV. Priority research questions and proposed research strategies

The emergence of smart system technologies presents both opportunities and challenges for improving national energy data systems in both developed and developing countries alike. However, exploring their potential for energy data collection and communication is particularly critical for developing countries as a possible means to accelerate progress toward policies that can enable the fulfilment of sustainable development, clean energy, and climate goals. Below are three priority research questions and proposed research strategies that can take important first steps in this direction.

1) **How can we maximize the potential utility of smart systems to national data systems in developing countries, and through which specific actions and policies?**

As discussed in this paper, there is significant potential for deployment of smart system technologies in developing countries to meet multiple social, environmental, and economic goals including those related to energy access, energy efficiency, renewable energy generation, and low-carbon development. However, as reflected by the information in Table 2, the mere existence of smart system technologies does not ensure that data with sufficient quality, formats, context, energy system coverage, and spatial and temporal characteristics will be available for credible and ready use in national energy data systems.

Moreover, smart systems are comprised of myriad metering, sensing, control, and communications technologies operating at different scales and for different purposes, and not all of which may be relevant to national energy data needs. Therefore, there is a real risk that smart systems may contribute to a "big data" unmanageable problem of large volumes of unusable data in developing countries unless these systems are supported by strategic national data system policies, protocols, and capacities to make best use of smart system data as they emerge. However, governments face considerable knowledge gaps on how to design such strategic approaches, given the nascent and rapidly evolving nature of smart system technologies.

One way to address these knowledge gaps is to develop a comprehensive technical roadmap for leveraging smart system information in national energy data systems, which would aim to:

- Develop and propose a taxonomy of terms and characteristics that represent the technical components, institutional elements, stakeholders, statistical interpretation methods, and flows of information comprising an idealized nexus between smart systems and national energy data systems, providing a more efficient and consistent lexicon for the research and policy communities;

- Identify (from case studies, literature sources, etc.) and propose a set of new or improved national data opportunities presented by smart systems, as well as ideal technical data characteristics (data types, spatial coverage, temporal coverage, sample sizes, etc.) needed to support national data systems across opportunity areas and within each opportunity area;

- Identify solutions for issues relative to data confidentiality, for example utilization of time lagged or anonymized data.

- Assess trends in major technology components comprising smart systems in key demand sectors (residential buildings, commercial buildings, industry, passenger transport, freight transport, and agriculture) and energy supply networks, and discuss their potentials for meeting ideal technical data characteristics;

- Identify major potential technical, economic, legal, behavioral, and policy barriers to achieving ideal technical data characteristics in real-world smart systems, particularly in the developing country context;

- Propose a list or "roadmap" of priority actions and policies that key system stakeholders (local and national policymakers, statisticians, technology developers, etc.) can take in order to reduce or eliminate barriers and to maximize the synergies between smart systems and national data systems in proactive fashion moving forward.

2) How can developing countries use new data sources to improve energy access?

Advanced metering infrastructure and remote sensing offer new opportunities for energy access. Smart meters are deployed in increasing numbers However, many of these technologies have features designed for the developing world, where 24/7 access to electricity with very low duration and frequency of service interruption is a pre-condition. In developed countries, consumers are accustomed to being unaware of their energy use.

In developing countries vast tranches of the population have no access, or access to low quality of supply electricity, and as a result are highly aware of their own energy use and the issues related to intermittent supply. Small-scale energy production for energy access can have be very different proposition in developing countries as it allows access to information and communications technology.

However, the minimum cost functionalities and requirements necessary for such technologies in a developing country context are poorly understood. Expensive designs are directly transposed, and the IEC, IEEE and ISO are not developing designs and standards with basic functionalities in mind. This research stream would focus on the requirements for smart metering and remote sensing technologies in a developing country context.

- Carry out a systematic review of technical needs for remote sensing equipment in developing countries. This should include the relatively high cost of regular maintenance and the need for predictive maintenance; low cost options for data collection and reporting (e.g. SMS, mobile apps); need for interoperability with outdated or legacy equipment.

- Engage the major standards organizations (IEC, ISO) to produce a white paper on standard needs for advanced metering and remote sensing in developing countries.

Match these needs to a review of the readiness of EEG countries of focus to adopt new data collection technologies across a range of applications (e.g. smart metering infrastructure, remote sensing in transport systems, etc). Need for pre-existing data privacy and security frameworks; institutional capacity; engineering skills.

3) How can we enhance collection and use of good quality energy data for better policy making, particularly in developing countries?

As discussed in this paper, there is strong potential to enhance the use of sound energy data for policy making, in developed and developing countries alike. The entry into force of the Paris agreement is just the most recent driver to highlight the importance of high quality energy data and its application in policy design and monitoring globally, needed now and in the near-term. Country-level NDCs have shown a variety of intended actions on energy, covering renewables and the role of fossil fuels on the supply side and energy efficiency on the demand side, all requiring strengthened or newer efforts to track national progress of the energy sector transformation, with much greater emphasis than previously needed on the energy end-use level.

This proposal is to develop a work program, which - from case studies, surveys to key stakeholders, literature sources, regional workshops - would aim, particularly but not limited to developing countries, to:

- Identify and share good practices in data sharing and in use of data within the overall energy policy context, from the individual policy monitoring and evaluation cycle to the national energy policy strategies.

- Identify and share good practices in national energy data collection globally, with greater focus on energy demand (residential buildings, commercial buildings, industry, passenger transport, freight transport, and agriculture), for which data are generally much weaker (or not even disaggregated), due to the difficulties of tracking a very large number of end-users; and on those where technologies have enhanced the effectiveness of data collection (e.g. surveys, administrative sources, and various techniques for utilising meter data).

- Identify major technical and non-technical barriers that generally limit the development of data collection, and data use for better informing policy making, at the national level, including through the production of energy balances.

- Propose a list of priority actions across the areas identified in Section II that key stakeholders (policymakers, statisticians, but also technology developers, etc.) can take in order to reduce barriers and to maximize the good use of energy data for policy makers.

V. References

Department for Business, Energy & Industrial Strategy (BEIS) (2013): National energy efficiency data framework (NEED). https://www.gov.uk/government/collections/national-energy-efficiency-data-need-framework

European Commission ESSnet Big Data project, Work Package 3, Smart meters https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/index.php/WP3_Report_1

Fadaeenejad, M., Saberian, A.M., Fadaee, M., Radzi, M.A.M., Hizam, H. and AbKadir, M.Z.A., 2014. The present and future of smart power grid in developing countries. Renewable and Sustainable Energy Reviews, 29, pp.828-834.

INOGATE (2016). About INOGATE: In Brief. http://www.inogate.org/pages/1

International Energy Agency (IEA) and EUROSTAT (2004). Energy Statistics Manual. OECD/IEA. https://www.iea.org/publications/freepublications/publication/statistics_manual.pdf

International Energy Agency (IEA) (2014a). Energy Efficiency Indicators: Essentials for Policy Making.
OECD/IEA. http://www.iea.org/publications/freepublications/publication/IEA_EnergyEfficiencyIndicators_EssentialsforPolicyMaking.pdf

International Energy Agency (IEA) (2014b). Energy Efficiency Indicators: Fundamentals on Statistics.
OECD/IEA. http://www.iea.org/publications/freepublications/publication/IEA_EnergyEfficiencyIndicatorsFundamentalsonStatistics.pdf

International Energy Agency (IEA) (2016a). IEA Energy Statistics Course. https://www.iea.org/training/calendarofieatrainingactivities/energystatisticscourseoctober2016/

International Energy Agency (IEA) (2016b). Toward Sustainable Urban Energy Systems: Energy Technology Perspectives 2016. OECD/IEA. http://www.iea.org/etp/etp2016/

Joint Organisations Data Initiative (JODI) (2016). Jodi World Databases. https://www.jodidata.org/

Lopes, J.P., Hatziargyriou, N., Mutale, J., Djapic, P. and Jenkins, N., 2007. Integrating distributed generation into electric power systems: A review of drivers, challenges and opportunities. Electric power systems research, 77(9), pp.1189-1203.

Ma, L. (2014) A Big Data Pilot Project with Smart Meter Data, Proceedings of Statistics Canada Symposium 2014 Beyond traditional survey taking: adapting to a changing worls. www.statcan.gc.ca/eng/conferences/.../program/14274-eng.pdf.

Masanet, E., and H.S. Matthews (2010). "Exploring Environmental Applications and Benefits of Information and Communication Technology." *Journal of Industrial Ecology*, Volume 14, Number 5.

Nangia, V., Oguah, S. and Gaba, K., 2016. Managing the Grids of the Future in Developing Countries.

O'Driscoll, E., and G. E. O'Donnell (2013). Industrial power and energy metering – a state-of-the-art review."*Journal of Cleaner Production*, Volume 41, February 2013, Pages 53-64.

Office of National Statistics (ONS) (2013). Guidelines for Measuring Statistical Output Quality, Version 4.1 http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/method-quality/quality/guidelines-for-measuring-statistical-quality/index.html

Organisation for Economic Co-operation and Development (OECD) (2014), Green Growth Indicators 2014, OECD Green Growth Studies, OECD Publishing. http://dx.doi.org/10.1787/9789264202030-en

Organisation for Economic Co-operation and Development (OECD) (2015), Strengthening National Statistical Systems to Monitor Global Goals. OECD Publishing. https://www.oecd.org/dac/POST-2015%20P21.pdf

Selvam, M.M., Gnanadass, R. and Padhy, N.P., 2016. Initiatives and technical challenges in smart distribution grid. Renewable and Sustainable Energy Reviews, 58, pp.911-917.

Statistics Canada (2009) Statistics Canada Quality Guidelines, Fifth Edition. Statistics Canada – Catalogue no. 12-539.

United Nations Economic Commission for Europe (UNECE) (2016), http://www1.unece.org/stat/platform/display/hlgbas/High-Level+Group+for+the+Modernisation+of+Official+Statistics

United Nations Statistics Division (UNSD) (2008). Preliminary report on the Global Assessment of Energy Statistics and Balances. http://unstats.un.org/unsd/statcom/doc08/BG-Energy.pdf

United Nations Department of Economic and Social Affairs (UNDESA) (2014).Developing green economy indicators in the context of the sustainable development: An overview of the main initiatives during 2008 – 2014. http://unstats.un.org/unsd/greeneconomy/

United Nations Department of Economic and Social Affairs (UNDESA) (2016). International Recommendations for Energy Statistics (IRES). Statistical Papers, Series M No. 93, New York. http://unstats.un.org/unsd/energy/ires/IRES_Whitecover.pdf

United Nations (UN) (2016). Sustainable Development Goals. https://sustainabledevelopment.un.org/

United Nations Framework Convention on Climate Change (UNFCCC) (2015). Adoption of the Paris Agreement. https://unfccc.int/resource/docs/2015/cop21/eng/l09.pdf

United Nations Statistical Commission (UNSC) (2012). Guidelines for the Template for a Generic National Quality Assurance Framework

(NQAF). http://unstats.un.org/unsd/dnss/docs-nqaf/GUIDELINES%208%20Feb%202012.pdf

U.S. Energy Information Administration (U.S. EIA) (2013). Manufacturing Energy Consumption Survey: Number of Establishments by Usage of General Energy-Saving Technologies, 2010. Washington, D.C. http://www.eia.gov/consumption/manufacturing/data/2010/

U.S. Energy Information Administration (U.S. EIA) (2010). Manufacturing Energy Consumption Survey: Number of Establishments by Usage of General Energy-Saving Technologies, 2006. Washington, D.C. http://www.eia.gov/consumption/manufacturing/data/2006/

U.S. Energy Information Administration (U.S. EIA) (2004). Manufacturing Energy Consumption Survey: Number of Establishments by Usage of General Energy-Saving Technologies, 2002. Washington, D.C. http://www.eia.gov/consumption/manufacturing/data/2002/

U.S. Energy Information Administration (U.S. EIA) (2000). Manufacturing Energy Consumption Survey: Number of Establishments by Usage of General Energy-Saving Technologies, 1998;. Washington, D.C. http://www.eia.gov/consumption/manufacturing/data/1998/

United Nations Economic and Social Council (UN) (2013). Fundamental principles of official statistics. http://unstats.un.org/unsd/dnss/gp/FP-Rev2013-E.pdf