Legal Entity Identifier or the utility approach to shared, high-quality basic micro-data

Francis Gross[1]

Containers, bar code, global trade: does technology or policy drive change?

A recent study[2] demonstrates that the standardisation of the shipping container contributed significantly more to globalisation than all trade agreements taken together. The radical reduction of transportation cost brought by the shipping container created a powerful drive for whole industries to seek the benefits of low production cost in faraway, then poor countries and enabled those countries to at last draw benefit from their poverty in the form of competitive advantage and grow, some of them to now being wealthy. Trade agreements merely removed barriers to that drive, but they usually came behind technology.

The advent of the bar code pushed that development further, enabling international trade to develop sophisticated global, just-in-time supply chains for the production of sophisticated goods, adding to international trade in raw materials and finished goods a complex web of trade in semi-finished goods and services in which a single component can travel through many countries before the finished product is goes into is delivered to a customer. Bar code, combined with information technology and cheap transport combined to profoundly change the structures of whole industries in very few years. Apple doesn't manufacture a single Ipad; retail is now dominated by global giants whose global data-integration creates dominant buying and marketing power.

How do globalisation and the digital age affect statistics?

Statisticians can be seen as "measurement engineers", middlemen who collect information about the "real" world they measure and produce information about that world for users who act in and on that world.

Statisticians put values on categories (e.g. exports, GDP, IIP) defined by users on the basis, for instance, of economic theory, that builds the models, mental or computer-based, with which the users structure their analysis and action. Those categories tend to be relatively stable over time, giving the long time series so prized by users. Limited by technology and driven by the need to keep cost low, statisticians have become used to collecting data specifically tailored to the

1 Directorate General Statistics, European Central Bank, francis.gross@ecb.europa.eu.

2 Quoted in the article: http://www.economist.com/news/...-trade-humble.

categories they are asked to measure. The system that results is highly stable, and more so since it is largely anchored in legislation, which is slow to change.

In a context of fast, deep and global change, categories used by statisticians and their users, and/or data collection schemes serving them can become obsolete or irrelevant, see settlement data for Balance of Payments. In such times, new categories emerge and trigger new data requirements. Such changes can be resisted by force of habit or be addressed through proxies based on strong assumptions, due to lack of resources and skills to obtain adequate data.

Globalisation and the advent of the digital age have de facto united our world by creating a dense web of long range, real-time economic relationships, which are changing the dynamics of the economy and finance in every single country and community. That change continues at high speed and is likely to continue for many years. The increasing complexity and velocity of developments is accompanied by disaggregation and de-integration of previously large, easy to understand economic units. For instance, large companies can now be clouds of thousands of legal entities without a real "home base" or networks of changing, competing suppliers. The world has become more "granular" with many more elements interacting globally, and ever less visibly, at high speed.

In physical systems, what ensues form such developments is turbulence. As an example, whereas a laminar flow of air around a wing can be described and predicted with a few simple equations and a few data elements (air temperature and pressure, speed, attitude), a turbulent flow around the same wing requires high performance computing and thousands of measurement points at high frequency. That also means that measurement and analysis made for laminar flows fail at precisely the moment where the behaviour of the wing becomes critical. The analogy comes back to statistics if, with the wing moving very gradually towards a turbulent regime, measurement is adapted a just little and complemented with a few additional observations, the pilot adding ever more instinct to the instruments. Until the plane stalls, that is, and then flying becomes a very different exercise.

The financial crisis has been the stalling of a good part of financial statistics. In a very short period of time, the financial system has become very turbulent, affecting the real economy and society, and that our measurement instrument was not up to the challenge at the worst moment. Luckily, turbulence has been contained, but we now know that we have a very different system from what our measurement instrument was built for. The system has gone past a threshold, into a new phase where it can behave in a very different and dangerous mode. Some parts of statistics need to change fundamentally to deliver what is needed: fast measurement that works in turbulence.

How can we know where and when an object thrown up in the air will hit the ground?

- ***Rock***: simple physics and measurement allow predicting precisely place and time of impact.
- ***Gravel***: assume that gravel is a rock and apply (1). Roughly right – some spread in space.
- ***Sand***: assume that sand is gravel, just a little finer, and apply (2). Still roughly right – more spread in space and now also time.
- ***Powder***: assume that powder is sand, just a little finer, and apply (3). It becomes messy.
- ***Nanopowder***: assume that nanopowder is powder, just a little finer, apply (4)…

… and find out that nanopowder floats away in the air, poisons your water, short-circuits your PC and penetrates your lungs where it causes cancer. The system has crossed a threshold, the old assumptions can no longer be applied, even as approximations. ***We need new thinking***.

Measuring a complex, fast, granular, turbulent, global, digital financial system

Users of statistics must be able to respond to unforeseen, hard to predict, rapidly changing and potentially dangerous situations. Faced with turbulence or crisis, they will need new statistical products, tailored to the situation, very close to real time, most of which cannot be delivered by the traditional statistical processes. We need new thinking.

Some communities have crossed the threshold to new thinking in measurement, such as meteorologists with their satellites and large scale computer models. Some aspects:

- Data is collected in very large amounts, permanently.
- Data covers the whole relevant system, for weather the Earth, whereby granularity can vary locally.
- Data is highly standardised.
- Data granularity is increasing with each new generation of satellites, increasing the usefulness of the output, e.g. longer, more precise weather forecasts.
- Data collection has become very cheap, through technology and economies of scale.
- Different data sources (weather, geology, etc.) use the same geographical coordinates, enabling their combination to produce new services for the public and the private sector alike, e.g. agriculture, conservation, etc.
- Data is analysed and transformed into useful information through very large scale computing facilities and, still.
- Output such as weather forecasts come very near real time.
- These processes are facilitated by technically- and scientifically-minded communities that cooperate globally.

However, economy and finance are very different from weather. Weather is a natural system whereas economy and finance are entirely man-made. At the micro-level, the molecules that make up weather follow fairly well-known rules of physics and chemistry, whereas economy and finance are at the mercy of human psychology, at individual and mass level.

Also, it has taken decades for global weather scientists and governments to build these systems, both technical and organisational, and to acquire the funding and the political support they now enjoy.

Weather therefore doesn't offer an easy blueprint for a quick solution for the measurement challenges faced by measurement in economy and finance, but it inspires a few thought that can help to shape our own new thinking:

- Solutions will emerge and evolve progressively, probably over decades.
- Solutions, or at least components, will need to be at the scale of the system, *i.e.* global.
- Solutions will require many communities to cooperate; that demands learning from all.
- A common vision can unite these communities around a shared goal and guide our work.
- We must begin with an ambitious, concrete project that benefits all within a finite time.
- That first project should also have transformational power to ease further progress.

Good, cheap basic data in a global, public Utility: a possible first project

Making basic micro-data for economy and finance good and cheap by building a global Utility as a public good could be such a first project. Basic micro-data is understood as identifiers and reference attributes allowing to identify and describe objects of relevance, for instance a legal entity, a financial instrument or a contract (share, bond, loan, etc.), links among them and changes to them. Basic micro-data can be seen as equivalent to the basic data collected by weather satellites or to bar-code data in supermarket supply chain: it is the layer of data that connects the real world and the data world.

Today, good basic data is so expensive, or even impossible, that many products or processes that need it become inefficient and mediocre, or even gradually ineffective. So, in that context, what would make good basic data? Has data become worse with time? No. The same data that was good fifteen years ago can be bad now, because:

- To be good in a global system, data must be globally standardised, fit for IT;
- To be good in a fast system, data must be real time;
- To be good in a complex system, data must be granular.

Basic financial data must now be all three at the same time: the definition of good has changed, crossing a threshold where the old methods don't work anymore. Just as nanopowder is not a rock, just ground finer. We need new thinking.

The Legal Entity Identifier (LEI)

Luckily, in the field of basic micro-data for finance, all sectors, private and public, facing similar challenges and the same obstacles, have shown that they are ready for new thinking. The experience of developing the initiative for a Global Legal Entity Identifier System (GLEIS), initially for the financial sector, has delivered global agreement that there can only be one Utility. In the words of Jean-Claude Trichet, former ECB President: "Each one needs it. There can only be one. No one can do it for themselves. So we must do it together."[3] Industry and authorities, across all their components, have been cooperating towards a global solution, which is now being implemented under a G20 Charter. The GLEIS will be worth billions to industry through cost savings and better risk control, and invaluable for authorities who couldn't deliver many new functions without. And it has the transformational power to facilitate further steps.

Currently, the Regulatory Oversight Committee (ROC), a global body of now 53 public sector institutions established under a G20-endorsed Charter, is preparing the creation of the central building blocks of the GLEIS, namely:

- The Global LEI Foundation, based in Switzerland (GLEIF)
- The private sector, 15 to 20 strong Board of Director of the GLEIF (BoD)
- The Central Operating Unit (COU), controlled by the BoD / the GLEIF.

The COU will be a start-up and the hub of a network of Local Operating Units (LOUs), that will conduct the registration, validation and other operations of the GLEIS. Beginning from the current "pre-LEI" implementation serving OTC reporting in the USA and the EU, the GLEIS has vocation to grow universal and global. One can hope that it would be successful and ultimately evolve beyond covering entities to become a "Financial Object Identifier".

Statistics and the GLEIS: what strategy for statisticians?

Statisticians will be well-served by the GLEIS as an infrastructure in many respects. However, the GLEIS will reach its potential only if it becomes the operational data infrastructure of the financial industry, worldwide, hence helping industry to solve the data quality problem that costs it so dearly.

Statisticians are by definition well placed to help design standards for basic reference data. They also have an interest to do all they can to help make reporting cheap, large scale, high quality, high frequency, high timeliness as well as highly flexible and easily extensible.

Hence, statisticians should adopt a positioning extending their strategic remit upstream along their supply chain. There they should actively influence the processes in and around the private sector industries, including legislation where they have expertise too, to lead to effective global adoption of data standards in the shape of a global Reference Data Utility as a shared strategic infrastructure, of which the GLEIS will be the first building block.

[3] Said in a meeting with EU leaders at the ECB on 10 Sep 2010.

Is it legitimate for statisticians to influence the world they are supposed to measure? Clearly yes. Indeed, not doing so would see statisticians have a much worse influence on that same world, simply by delivering sub-standard products in fields so critical to the safety of our society.

We need new thinking.