Legal Entity Identifier – review of progress and outlook

Abstract

There have been many calls and initiatives in the past to establish unique, global, universal identifiers for businesses, and more recently for financial products and transactions. Nearly all of them failed. The Legal Entity Identifier (LEI) is the notable exception, to date.

The article presents the vision and concept of the Global LEI System (GLEIS) and recalls the history of the LEI’s emergence. It introduces the organisational design and business model of the GLEIS and reviews the current state of progress at the cross-roads between the pre-LEI System, in which over a dozen pre-Local Operating Units have registered over 300,000 entities in over 180 countries, and the establishment of the formal GLEIS, which will be built around a Central Operating Unit to be created as a start-up.

The author provides an overview of the challenges the GLEIS must overcome if it is to succeed and explains why, even now, it is too early to take the success of the LEI for granted. He also explains why the success of the LEI is a necessary condition for many post-crisis reforms to become effective and for other similarly important data initiatives to even have a chance to take off beyond calls. Finally, the article looks at the possible role business registers could take up in the GLEIS.

The present article reflects the personal views of the author only.
New strategies for a changed context

The digital age demands a fresh look at our context

“We can't change anything until we get some fresh ideas, until we begin to see things differently.”

James Hillman, Psychologist, 1926 - 2011

Slightly different chemical bonds among atoms can give a material very different macroscopic properties, for instance being hard instead of soft. Small differences in genes can make for a new species. All the same, a few “smaller” technical innovations of the last twenty years changed our world radically and continue to change it, even where we still might be tempted to perceive stability. Many practices we had become used to are now obsolete; replacing them with successful habits will be difficult, but we should do so while we have time to adapt. Casting a fresh eye on our world can help.

Like it or not, radically changed space and time will change what we do and how we do it.

“You may not be interested in war, but war seems to be interested in you.”

Attributed to Leon Trotsky, Soviet Politician, 1879 - 1940

Moore’s law and the internet have changed our world, enabling real-time, long-range interactions between humans and also, automated, between machines. De facto, they have changed our space and time by annihilating distance and borders, increasing speed by orders of magnitude and creating an explosion in the number and volatility of connexions across the world, among people and machines. Our elders would have rated those new powers as quasi-magical. But there is a downside: now, events far remote from us, we are not even aware of, can have immediate influence on what happens to us. This all applies to the financial system and, like it or not, impacts business registers and statisticians alike.

Whereas many, mainly physical, industries have adapted by changing profoundly, finance (both private and public sector) has let many of its practices become shockingly inefficient and obsolete to the fresh eye. Three reasons could explain the faster progress of physical industries:

• Visibility: real-world failure in a physical process is immediately visible to many, it can’t be ignored.
• Scale: physical industry processes tend to be large-scale. Failure or inefficiency costs a lot, at once, hence it won’t be tolerated - it is interesting to see that also some large-scale financial processes, such as payments, have also become very efficient and up-to-date in their use of technology.
• Culture: physical industries tend to be dominated by more practically focused, technically aware types. Finance remains dominated by people less sensitive to technical or organisational constraints.

Scale has come to finance, partly through the back door

Scale is a powerful driver of change, great opportunity for some, deadly threat for others. Scale in industry has been at the heart of the economic growth that gave us many of the good things we can afford today. Remember Henry Ford?
Finance has scaled as well, in two ways. We all see very large banks, some too big to fail. It is less easy to see that pervasive long-range, high-speed and real-time, partly automated interactions have de facto integrated finance into a single global system. Yet, seeing that is critical, for we must understand that system on a global scale if we want to act responsibly and society to stay in control. A necessary first step is to measure finance on a global scale and at a speed commensurate with the speed of the system.

**Measuring a complex world to serve limited human capabilities**

Our world has become infinitely more complex, very quickly. “Complex” means that much more information than ever, from many more sources across our universe becomes relevant for making an informed decision. “Complex” also means that compressing those huge volumes of information into aggregates before analysis, as we usually do, destroys its value for understanding the complex reality: much of that information must remain available for analysis as is, i.e. in huge volumes. “Complex” is also a measure of the ratio between the volume of information required to address a situation and the capacity of the human brain to absorb and process it; that ratio has exploded and continues to grow.

This contrasts with a simpler world where a little local information and a broader idea of what happens beyond suffices to assess a situation and to make an informed decision for effective action.

High-speed, real-time, long-range interactions have also made the financial system much faster than human timescales. The shift towards complexity and speed has considerably affected human capability to stay in control: our brain has limited capacity to absorb and process information and we need time to react well. Even in a simpler world, we need more time when we first must agree with others on what is happening and on what should be done. The more complex our environment, the more our capability as individuals and society to react effectively and timely, i.e. to stay in control, comes under challenge.

**Human society faces a complexity shock**

Here, a society of humans with limited individual ability, needing more time to collectively understand and act, reacting ever more slowly to events in the environment. There, a more complex environment that can develop sudden, large-scale, high-speed and high-energy, turbulent behaviours, giving humans ever less time to react before a potential tipping point to disaster.

**What strategies are possible?**

“**Control your own destiny or someone else will.”**

Jack Welch, Corporate leader, 1935 -

The symbolic graph in Fig. 1 supports a simple idea: if we can shift the two curves to intersect at higher complexity we will be safer. This paper focuses on shifting upwards the “human capability” curve, increasing society’s capability to react faster to events in our environment. It shows how the Legal Entity Identifier helps and why it is a good starting point.
We must enable faster, better-informed decisions and faster, more accurate execution

The following goals for improvements could help and guide us towards useful strategies:

1. **Better measurement**: gives us better, timelier information on what is going on out there;

2. **Better analysis**: faster, more precise, more versatile. Will require quick, ad-hoc additional measurement of things not seen before and large, flexible computing capabilities;

3. **Better decision-making**: through better, faster information and the ability to test and calibrate ideas;

4. **Better execution**: faster and more accurate; might require creating new means for action, and

5. **Better infrastructure**: to enable the above, which otherwise might remain theory or wishful thinking.

The statistician’s focus is on better measurement (goal 1 above). Measuring fast, turbulent events in a globally integrated, digital-age financial system is largely out of reach of classical statistical practices, which were developed for monitoring slow developments in more placid times, usually in a national context, and which still do a good job serving those needs.

To better measure today’s faster, global finance, we must also envisage a more fundamental overhaul of parts of the infrastructure underlying the current ways to measure finance (goal 5 above).
Cannot see, smell or touch finance - need a proxy: data.

None of our senses perceives the abstract objects of finance; we need tools. Concepts, labelled with words and numbers are our tools; we combine them into formal statements we call “data”, e.g.: the “bond” (“ISIN”) has a “face value” of “100” “Euro”. Data is a proxy for “real world” objects, enabling us to “handle” them without touching them. Machines that process data for us produce more data, which we then interprete as proxies for other “real world” things. More powerful analytical data processing, embracing wider areas of finance in more detail, demands ever more and better data. It fails utterly if data is sub-par. Has data generation, the interface between the “real world” and the first layer of “proxy-data”, evolved to the industrial scale required by that exploding hunger? In short: not really.

How does data scale?

A financial system whose “real world” objects interact at global scale and high speed must be understood at global scale and high speed. How does data perform as our tool at that scale and speed?

Our history and culture of data is largely one of ad-hoc generation for specific local use: many uses of data are “vertical”, for instance in an administrative or operational process related to the individual “real world” object. Global measurement demands “horizontal” use, where data on many “real-world” objects is used as a pool to reflect, measure, assess the behaviour of large populations of such objects. Such population often show decoupled behaviour, hence near-source, pre-analysis aggregation, a legacy of pen and paper times, extended by the assumption that a population behaves in lockstep, simply won’t work. That means in principle micro-data end-to-end, hence massive data operations. And potentially massive investment, beginning with research. And perhaps significant re-skilling. And time...

Vertical use, e.g. registration in a business register, supports operations or compliance specific to that entity. Even a free-format paper register could deliver that, with a little patience, although when fewer humans must treat more cases, automation becomes necessary for speed and efficiency here too. For instance in payments, straight-through processing is essential for low-cost, fast, high quality service.

Data quality is key to the performance of automated “vertical” processes, but demands are even higher for “horizontal” uses, especially across large populations. Building larger pools of data by merging smaller ones looks the easy way, but generally falls far short of demands. For instance, one will find many data sets representing a same “real world” object, created by data vendors and users, often telling different stories. “Data cleaning”, the recourse for users left to themselves, yields yet more data sets on a given object. Mapping the many formats, another attempt, resulted in “mapping hell”. Data cleaning and mapping have outlived their usefulness. Why? Because they’re slow and craftmanship doesn’t scale. In large scale data, inefficient simply means ineffective, i.e. failure of the functions that data serves.

“State-of-the-art demand” is, among others, for complex analysis, such as on “exposures” among large banking groups, each one made up of thousands of legal entities based in dozens of countries, transacting in millions of instruments issued by any entity among millions, including in derivatives that depend on the behaviour of other instruments or entities, on indices or on natural events. And each
user needs that same data for ever more core functions, just to survive. Ever more users in the financial system need data qualities our legacy cannot deliver, even if we make the smartest use of it.

Today’s demands require very large scale pools of basic data strictly standardised at source. But data as it is practised today doesn’t scale. It is in our common interest to build a better, large (global) scale data infrastructure. Together. And for all of us.

**What goals should we have for data?**

In finance, only data and statistics allow us to see. With today’s data, analysts, accountants, statisticians and engineers might not even agree on the shape or size of what they “see”. If that sounds extreme, remember the days of the Lehman failure or read the press. Whatever long-term goals for data, a reasonably ambitious yet feasible first goal should be to identify every single “real world” object in the financial system in a standardised, unequivocal way. That would deliver tremendous value at once to all sectors, also in operations. It would have transformational power, leading to much further progress not possible without it. Witness bar code that unlocked fundamental change in many industries.

**Statistics: what can stay? What will have to change?**

“21st Century” statistics will be comparable to today’s on some counts and very different on others.

What will remain? Statistics will continue to deliver condensed information produced from large volumes of data collected from large numbers of sources and it will continue to deliver products fit for use, generally trying to serve the needs, expectations and habits of its many users.

What will have to change or to be added?

- Products for real- or near-time, and high-frequency delivery will be added to the slower ones;
- Micro-databases ready-to-use will expand the product range beyond aggregates and time-series;
- Products dynamically tailored to suddenly arising new needs will be added to rigid, pre-set products;
- Global focus must supplant national focus, where relevant, for instance in data collection;
- Very much larger data volumes, as input and as output, as flow and as stock;
- Many more data sources will be tapped, globally, through new organisations and technologies;
- Some compilation processes will scale, when fragmentation becomes inefficient or leads to failure;
- Very-large-scale, radically new data products, such as “live maps” of the banking industry, or “flows of liability” will be compiled and maintained daily, their history documented;
- Automation will become even more dominant, throughout the whole supply chain from data source to database to live product to information for human or machine consumption.
The long way to “data nirvana” starts near you

“Quality is produced, not checked.”

Werner Niefer, former CEO of Mercedes-Benz AG, 1928 - 1993

Building the data infrastructure and statistics of 2100 cannot be planned, but we must get started; we will learn as we progress. A few constraints can guide us, especially bearing in mind very large volumes and high-speed flows of data at global scale:

• Data processing will need to be fully automated throughout the supply chain, end to end;

• Data quality must be given at the input point. Data volumes and speeds as well as scarcity of human time prevent any fixing downstream: right first time will be the iron rule.

Those two constraints deliver a first round of guidance for designing new strategies:

• Data that refers to and describes “real world” objects must be understood in the exact same way by all who handle or use it.

• The most basic layer of data, on which most other data build, should refer to the most basic layer of “real world” objects that can be identified unequivocally and in the same way by all parties involved.

• Basic data becomes better with scale. Data standardisation is ideally best when data is collected, maintained and stored in a single, large-scale machine that imposes language and format. In the internet age, a federated network can, together, operate a (real or virtual) single machine, hence a single system can combine the benefits of centralisation and decentralisation.

Not all data can be subject to such a strategy, of course. However, our migration to a better data world could start from a place where data responds to such criteria. We speak of reference data identifying the most basic elements of the financial system, which are already subject to clear identification we can all agree about: entities, instruments, transactions.

A good showcase will ease many applications, paying for itself very quickly. It will support discovery of the best next steps: easiest, cheapest, accepted by most, that delivers value to all. Next steps can be taken by public sector agents or by businesses seeking competitive advantage through innovation. Where a joint solution is the only feasible or the clearly better one, both should work together.

A start has been made by the Legal Entity Identifier (LEI). That’s where the way to data nirvana starts.

A first step and a showcase: the Global Legal Entity Identifier System

An effective Global Legal Entity Identifier System (GLEIS), though a necessary first step, is certainly not sufficient, but it has become feasible – and it is being built right now. Building the GLEIS will give us a much-needed infrastructure; it will also offer a playground for learning for the next steps. Moreover, building the GLEIS will provide all of us with a data infrastructure that will make our operations more
efficient and should help fixing some that are not really working today. Indeed, the LEI will reduce the waste that comes with low quality data and that grows exponentially with increasing complexity of the financial system. The key to success lies in data standardisation in the strict, rigorous sense of the term.

A brief history of the LEI’s emergence

The emergence of the LEI is a result of the commitment of a few individuals from all sectors who came together and were lucky to have the crisis as the opportunity for the idea to emerge on the policy stage. Early on, the idea won enthusiastic support from industry associations and some regulators. It was debated at many industry conferences and took off after high-level leaders endorsed it publicly, first among them Jean-Claude Trichet, then President of the ECB, in a speech on 23 February 2009.

The LEI was enshrined in the US Dodd Frank Act of July 2010. At its Summit on 4 November 2011 in Cannes, the G20 mandated the FSB to develop recommendations for a LEI. A global, public sector “Expert Group” supported by an informal network of industry experts delivered its report to the June 2012 Summit in Los Cabos, Mexico, where the G20 endorsed the recommendations as submitted, asked for their implementation and issued Principles for the LEI. The now “Implementation Group”, again with support by industry experts, delivered the LEI Charter. The G20 endorsed the Charter at a meeting in Los Cabos on 4 November 2012. In parallel, ISO developed the ISO 17442 standard for the LEI code and attributes attached at an unprecedented pace, delivering in months what normally takes several years.

The LEI Charter established the Global LEI System and its supreme governing body, the Regulatory Oversight Committee (ROC). The ROC was formally established on 24 January 2013 in Toronto. The ROC admits as members public sector institutions interested in the development of the LEI; it now counts 65 members from over 40 countries on all continents. For more information, see www.leiroc.org.

Organisational design of the GLEIS and current state of implementation

The ROC guides the development of the GLEIS. Operations of the GLEIS consist of a globally federated network of Local Operating Units (LOUs) led by a Central Operating Unit (COU), itself “operated” by the Global LEI Foundation (GLEIF), founded by the FSB under Swiss law on 26 June 2014.

The GLEIF has a 16-strong Board of Directors selected from the private sector, balanced for regional and sectoral affiliation, covering the scope of expertise and experience required for guiding the successful development of the COU and the GLEIS’ operations. The GLEIF will hold all intellectual property that might be generated by the GLEIS, to avoid any undue appropriation of such rights.

The COU is a start-up, currently being established. The recently recruited founding CEO of the COU will be Stephan Wolf, an industry veteran and successful data entrepreneur. The COU will organise and develop the GLEIS’ operations; it will build and lead the network of LOUs. It will interact with many stakeholders of the GLEIS, where necessary jointly with the ROC.

The LOUs register entities entitled to a LEI; they validate data submitted and maintain their accuracy.
The GLEIS can be better understood if compared to a franchise, such as some hamburger chains.

The ROC is supported for development work, both conceptual and practical, by the public sector Committee on Evaluation and Standards (CES).

**Business model of the GLEIS**

Data held in the GLEIS is a public good; it is freely accessible to all for all uses, without any licensing or commercial constraints. The business model of the GLEIS is designed to support the nature of the GLEIS as a public good and to ensure its sustainability: the GLEIS is financed through registration fees and annual maintenance fees paid by each registrant to their LOU. The COU is strictly non-profit, financed by a levy on fees collected by the LOUs. LOUs should set fees such that they operate under cost-recovery.

**Data Model of the LEI**

The LEI Charter states that the GLEIS would collect two types of data, “Level 1” data and “Level 2” data.

“Level 1” data, defined by the ISO 17442 standard, identify an entity through its unique 20-digit alphanumeric LEI code and provide a few basic attributes describing it, such as name, address of headquarters, date of registration, address of registration.

“Level 2” data represent relationships between entities registered at Level 1. Work is progressing at the CES, supported by industry experts, on designing concepts for relationship data. To enable early useful results, a first generation of relationship data will consist in identification of immediate and ultimate parent in the accounting sense. That first step is currently under development. It goes without saying that good “Level 2” data will require high coverage at “Level 1”.

All data of the GLEIS will be equipped with metadata and full historical data that will support quality management as well as forensic analysis.

Whereas the first generation of relationship data is probably going to remain focused on ownership relations, relationship data should grow to cover many other types of relationships. For instance, each contract, financial instrument or other, establishes a relationship between entities concerned.

Conceptually, it could be envisaged that “Level 2” data will be collected in a database of “relationship elements”, the minimal statement that describes a relationship. A “relationship element” could hold as basic data elements: the LEIs of the entities concerned, the type of relationship, the description of the relationship. “Level 2” data will then be used to build the next layer of statistical and analytical products.

**Group maps, networks, exposures, etc.**

Relationship data will be extremely useful as they will enable tasks such as understanding groups and networks, and calculating exposures. Before the LEI’s advent, all of these products were largely theoretical visions and in practice wishful thinking.
It is to be expected that the LEI relationship data, once they reach critical mass, will give rise to a whole new range of services, perhaps to a few new branches of the financial information services industry, initially in fields related to risk management, compliance and transparency. Indeed, the rise of the LEI might enable the emergence of services that would deliver daily calculation of group and network structures and exposures among them. Such services could flourish amidst competition that should also boost the growth of new theories and products serving the many potential users of such information.

One can also speculate about how authorities will use that new data resource. Will they choose scale or silo for their applications? What will that choice mean for their capability to develop faster, high-quality information/statistical products at reasonable cost? What if the cost and complexity of such operations require scale, and if the specific skill profiles and capabilities required are not sufficiently available at any single authority? Perhaps new organisational solutions will evolve, maybe shared technical services? Or perhaps some of those services will have to be sourced from industry?

Part of the answer to those questions might come from the consideration of legal possibilities for offering access to data, mainly a matter of privacy and data protection, briefly treated in the next section. One conclusion appears certain now: using the data in a sub-optimal fashion would defeat the main purpose of the whole infrastructure, namely to increase financial stability and reduce systemic risk.

**Privacy vs. transparency: a societal debate in waiting?**

There is broad consensus that “Level 1” data won’t present legal challenges to access and that access to “Level 2” data will likely present such challenges on grounds of privacy and data protection in a number of jurisdictions. As the solutions might require changes to law, debate is likely, both at the level of society and among experts representing divergent interests.

The societal debate could happen around which of privacy and transparency in financial markets is the higher good, and how best to reconcile both in practice. Years ago, an article in the Economist noted as strange that legal entities enjoy the same protection as granted to people on grounds of human rights.

One front-line in the more discrete expert debate could well be between society’s interest in transparency and the protection of financial operations built on the kind of intransparency that would be lifted. In that debate, there should be shortage of legitimate arguments to prevent change, which should however be no guarantee for a quick or clean conclusion.

**Early stage, the pre-LEI system**

The US Commodities Futures Trading Commission (CFTC) and the European Securities Markets Authority (ESMSA) mandate the use of the LEI as identifier of all counterparties to derivative transactions reported to them under CFTC Rule from October 2012 and under EMIR from February 2014 onwards. Those two reporting schemes started at a time when neither the GLEIF nor the GLEIS existed.

Hence the ROC established the so-called pre-LEI System under which ROC members could sponsor a candidate organisation, private or public, wishing to become a “pre-LOU”. Pre-LOUs had to obtain endorsement from the ROC, subject to meeting set criteria. Those pre-LOUs would then register entities
and issue “pre-LEIs”. On the eve of the formal establishment of the GLEIF, 14 pre-LOUs had, together, issued nearly 300,000 pre-LEIs in 186 countries.

**Current state of progress and challenges ahead**

Following the establishment of the GLEIF, all pre-LEIs issued under the pre-LEI regime became LEIs.

The COU must now overcome the numerous hurdles facing a start-up on day one, including the selection of a location, which must not necessarily be the same as for the GLEIF, and will have to be selected to offer an optimal environment for the operation of the COU and its network of LOUs. All of these steps, including the recruitment and forming of a strong team as well as the shaping of the network of LOUs and the underlying relationship between the COU and the LOUs will determine the initial success of the GLEIS.

The GLEIS will have to overcome numerous challenges. For one, it is essential that all members of the GLEIS be and remain committed formally and in spirit to the values and principles of the LEI as set by the G20, essentially as a global monopoly and an infrastructure dedicated to the public good. There will also be numerous challenges in the fields of technology, design and processes – technical, commercial, legal –, which can all be overcome through solid work and honest, good cooperation among all stakeholders.

Finally, the LEI will need more active support from all who can support its development and who stand to benefit. To that end, it needs to be better known and understood by its numerous stakeholders, including incumbents in the field such as business registers, but also by authorities who will be expected to mandate its use to boost its spread. By its very nature, the value of the GLEIS will increase with coverage and quality; it must strive for global, universal coverage: all entities in all countries.

**Business registers and the LEI**

The mission of the GLEIS can be described as providing a globally unique identifier for “legal entities”, a term whose definition might evolve over time alongside needs. That unique identifier will enable all data sets on a given entity to be attributed to that entity, a capability whose absence became painfully visible when the crisis broke, and for which there is no other known solution.

Whereas the GLEIS stores and maintains a small number of attributes with the LEI of a registered entity, destined to facilitate identification, the full depth of information existing today will best remain where it is, among others at business registers.

Business registers should feel chiefly interested in the LEI, but have no reason to feel threatened by it. The LEI simply adds the globally effective link that is missing in today’s globalised world, and that national, regional or sectoral business registers cannot deliver by themselves.

Actually, some business registers could find out that they are well placed to play a role in the GLEIS, for instance as LOU, a step some such as INSEE or Bundesanzeiger have chosen to take. There might also be an opportunity to study the adoption of the LEI as the business registers’ standard identifier, globally.