



# Digital Continuum in Materials Research

## Digital transition applied to Material Nano-Characterization.

Keywords: Materials – Characterization – Data Management and workflow – DFT

Authors: Thierry Deutsch, Luigi Genovese (CEA – Grenoble, France)

### 1- Challenges in the field of materials simulations

---

Materials simulations, quantum-mechanical in particular, have become dominant and widely used tools for scientific discovery and technological advancement: since they are performed without any experimental input or parameter they can streamline, accelerate, support or replace actual physical experiments. The field of computer-assisted materials modelling, discovery, and engineering is extremely vital.

Enhancing the impact of quantum mechanical modelling on materials research will stretch our simulation capacity towards systems of increasing size and complexity and our predicting abilities towards chemical accuracy, thus requiring the deployment of ever more sophisticated (and expensive) modelling and theoretical methods. More and more frequently, the whole process relies on extensive, database-driven searches, where millions of calculations are deployed to probe unexplored vistas in materials' space, in turn accumulating **an ever-increasing treasure trove of curated, high-quality computational data**.

Community codes are often a jumble of mathematical libraries, equation solvers, and property calculators, resulting from uncoordinated efforts of generations of students and postdocs, usually without specific training in IT, and aiming at quick results rather than at structured, easy-to-maintain and to-port software. As a consequence, porting those codes among hardware architectures has always required extensive recoding in the past. This cannot be sustained any longer in view of the considerable level of complexity reached by both community codes (several hundred thousands code lines each) and the forthcoming diverse, heterogeneous, and rapidly evolving hardware architectures.

The solution we have identified is to refactor the code base into multiple software layers, resulting from the assembly of weakly coupled components (modules and libraries), to be maintained and enhanced independently from each other, shared among different codes, and ported across different architectures. In the European project MAX, we will now leverage our past efforts **to design solutions that will work across entire classes of codes**, and deploy them on a larger and even more representative number of community codes that will be ready for production on pre-exascale machines and beyond by the completion of the present programme. This will allow us to achieve a substantial economy of scale and the entire scientific community to exploit solutions that can be adopted by community codes not yet represented in our consortium.

### 2- Challenges in the field of characterization techniques

---

In other way, progresses in characterization techniques have been constant over the last decades due to the continuous development of radiation sources (e.g. synchrotron), optical elements, nanometer positioning systems, 3D techniques etc... Development of techniques that are capable of **material or device characterization at the nanoscale** are often a basic requirement of many process or device developers. The techniques covered include electron microscopies (TEM and SEM) with associated spectroscopies (EELS, EDX) and imaging



# Digital Continuum in Materials Research

techniques (holography, tomography), ion spectroscopies (SIMS), XPS, NMR, Scanning Probe Microscopies, etc.

However this ever increasing performance in the sensitivity and accuracy of characterization tools has been often accompanied by an increasing amount of generated data. Now it is rather common to have 1 Terabytes per experiment which have to be processed and analyzed by means of models and simulations. The technicality of this processing is often far below that of the characterization one (home-made or generic software on personal computers).

It would then be very valuable to tailor, in the Materials characterization area, some new digital tools as Data Management, Data workflow management systems but also some Artificial Intelligence tools like machine learning ones to classify and organize data or to build models.

**Data management and analysis are keys to fostering materials characterization,** accelerate the use and the exploitation of these results, improving materials but also industrial production by a better quality control.

## **2- Towards a digital continuum between characterization and simulation**

---

A paradigm shift for computational design and discovery is also ensuing, in which massive high-performance computing (HPC) and high-throughput computing (HTC) efforts are combined with high-performance data analytics (HPDA) to identify the most promising novel materials or those with improved or designed properties and performance. Such effort requires simulation codes able to deliver the predictive accuracy needed to sustain or streamline experiments; able to address the complexity of real-life conditions; able to exploit or drive the evolution of the current and forthcoming hardware platforms; and able to leverage the wealth of data that is generated or harvested in computations and characterization experiments alike. **The capabilities of codes need to be standardized in workflows that provide highly curated data on demand,** and that can be exploited easily using cloud technologies by the scientific and industrial community at large.

Moreover, these data workflows should be directly coupled with 3D characterization improving considerably the material analysis. Statistics of the structure of materials could be greatly improved by a systematic and automated analysis using a digital continuum between characterisation and simulation based on these complex data workflows.

**CEA contact: Dr. Thierry DEUTSCH**

[thierry.deutsch@cea.fr](mailto:thierry.deutsch@cea.fr), phone: +33 4 38 78 34 06