Chapter 2
A Big Data Platform for Enhancing Life Imaging Activities

Leila Abidi
Université Paris 13, France

Hanene Azzag
Université Paris 13, France

Salima Benbernou
Université Sorbonne Paris Cité, France

Mehdi Bentounsi
Université Paris Descartes, France

Christophe Cérin
Université Paris 13, France

Tarn Duong
Université Paris 13, France

Philippe Garteiser
Université Sorbonne Paris Cité, France

Mustapha Lebbah
Université Paris 13, France

Mourad Ouziri
Université Sorbonne Paris Cité, France

Soror Sahri
Université Sorbonne Paris Cité, France

Michel Smadja
SISNCOM, France

ABSTRACT

The field of life imaging spans a large spectrum of scientific study from mathematics and computer science to medical, passing by physics, biology, etc. The challenge of IDV project is to enrich a multi-parametrized, quantitative, qualitative, integrative, and correlative life imaging in health. It deals with linking the current research developments and applications of life imaging in medicine and biology to develop computational models and methods for imaging and quantitative image analysis and validate the added diagnostic and therapeutic value of new imaging methods and biomarkers.

DOI: 10.4018/978-1-5225-4963-5.ch002

1. INTRODUCTION

The healthcare industry is a large generator of biomedical data. For instance, the U.S. healthcare system expected to reach the zettabyte (10^{21}) scale from electronic health records, scientific instruments, clinical decision support systems, or even research articles in medical journals (Raghupathi & Raghupathi, 2014).

In the last decade, we have witnessed the increasing resolution of imaging technologies which are considered as one of the most promising medical and health areas example and application of big data (e.g., NIH Brain initiative, n.d.) transforming case-based studies to large-scale, data-driven research (Luo, Wu, Gopukumar, & Zhao, 2016) and (Serrano, Blas, Carretero, & Desco,2017).

Interdisciplinary research in the field of imaging in the life sciences is essential. It requires the implication of different clinical and preclinical imaging departments yielding easy access to the state-of-the-art imaging equipment and patient data. Cooperative projects, including physicians, mathematicians, computer scientists, and physicists who are working closely together with bio scientists and clinicians are then launched in order to *(i)* develop computational models and methods for imaging and quantitative image analysis, and *(ii)* validate the added diagnostic and therapeutic value of new imaging methods and biomarkers.

Imaging is characterized by a large diversity in the types of data. Indeed, the data can originate from many different acquisition device, i.e., modalities, and the data format convention are quite loose with an important diversity in file formats and in completeness of annotation. The data themselves also strongly differ in their dimensionality, scale, size, and finality.

In such context, the life imaging project "IDV" (for Imageries Du Vivant) funded by University Sorbonne Paris Cité (USPC) launched the "Atlas IDV" initiative, which is a typical use case for data volume, variety and veracity in big data. The Atlas IDV initiative aims at *(i)* providing an integrated and agile environment supporting cooperation between scientists, and *(ii)* enabling to augment the research perimeter of imaging scientists and the extraction of new knowledge (data-driven research and images analytics) from the big multi-modal and multi-scale clinical and preclinical images available within the university.

A lot of studies in small animal imaging are hampered by small number of subjects, to the detriment of statistical quality of the findings. The junction of imaging data from a wide perimeter enables researchers to analyze a larger number of subjects, and hence to improve the statistical quality of their reports. Two use cases can be cited: