Emerging Datasets and Analytics Opportunities in Metals Additive Manufacturing

Paul Witherell
National Institute of Standards and Technology
Gaithersburg, MD 20877 USA

Abstract

Additive manufacturing (AM) technologies continue to mature, evolving into stalwarts of high-end production lines, particularly with metals AM. Technology maturation has been facilitated by efforts in materials characterization, process sensing, and part qualification, among others. Advancements have been accompanied by a proliferation of AM data that is creating many new learning opportunities that have yet to be realized, hindered by a lack of curation and sharing. Data is often being generated in silos; associated with a specific time, process, material, location, etc. This manuscript investigates the state of data curation and analytics in AM. It begins by investigating AM data types and how this data is currently generated, curated, and shared. It then looks toward the future, where improvements in data curation will support emerging analytics. Finally, short-term needs and long-term opportunities are discussed, outlining future directions in data analytics for AM.

1 Introduction

In recent years, metals additive manufacturing (AM) technologies have established themselves as a stalwart of advanced manufacturing efforts. While AM technologies have long shown promise in realizing design freedoms, it has not been until recently that these technologies were seriously considered as production alternatives. Success stories from early adopters are demonstrating practicality and continuing to drive AM technologies into mainstream manufacturing as viable, profit-driven production technologies. Industry investments in metals AM have increased substantially in recent years and show no signs of stalling [4].

While initial investment has long been an entry barrier to metals AM, machine investment is only a portion of this cost. Much of the upfront costs can be attributed to “ramping up” activities, including the training and experience necessary to realize production-quality parts. The development of in-house expertise has been a distinguishing factor in determining maturity levels of AM practitioners. Improved understanding of the design-to-product transformation is necessary to overcome these barriers that hinder market entry and the capability to manufacture original and one-off designs.

Towards a better common understanding of AM processes, many efforts have focused on standards development, establishing specifications and communicating best practices in AM [5]. In parallel, significant investment has been made in measurements to reduce overall process uncertainty in AM, including design, material, process, and part characterization. Here we focus on the offshoots of these measurement, as AM characterization is introducing large amounts of new datasets from disparate sources. However, generating data and gaining knowledge from the data generated are quite separate matters.

Data analytics are critical to managing process variability and maturing AM processes. However, processing data can be a tedious task, and often becomes a limiting factor of the value of datasets produced. Studies have shown that only 30 percent of production data collected is also analyzed [6]. With new datasets being generated with domain-specific context, AM provides a unique opportunity to significantly increase that number. Recent advancements in software and informatics technologies are well positioned to capitalize on this proliferation of data.

This paper investigates new trends in AM data generation and methods to improve how AM data is leveraged. To increase data usage we must understand what the data sources are, what processing is necessary, and where the opportunities lie. Linking datasets across a build will provide new insight into the design, manufacture, and qualification details of an additively manufactured part. Compiling this data across builds promises to provide new insight into process control and process improvement.

2 Background

A signature trait of today’s advanced manufacturing processes is the ability to incorporate data into decision making. This incorporation includes measuring process performance, establishing performance baselines, integrating predictive analytics, defining performance metrics, and assessing quality, among others. AM, as perhaps the most “digital” of these advanced manufacturing technologies, stands to significantly benefit from advancements in how data is captured, curated, managed, and incorporated into decision making.

A single, unifying, characteristic of almost all AM data-driven activities is the desire to better understand performance at each stage of an AM lifecycle [7]. Digital threads have been explored to establish provenance of part or process behaviors at given points of time, and their aggregates [8]. However, AM measurements are creating new data management challenges with increased levels of detail and new representation requirements of data types such as time series data and image registrations. The lifecycle stage-driven perception of a digital thread is being challenged as datasets increase in complexity.

New AM data requirements are often introduced from the bottom-up as new measurements are taken, diverging datasets across stages of the design-to-product transformation. Silos of disparate data become representative of ongoing data collection, with types of data varying from voltage signals to light intensities to images to statistical analyses to vectors to graphs to voxels. While the large amount of data clearly has much to offer, extracting value is the challenge now faced. What has yet to be established, but is emerging, is an emphasis on top-down configurations of these new sets of data.

Big data and related concepts are teaching us that as we collect new information there is always knowledge to be gained, even in the most unexpected places [9]. By emphasizing the learning potential of available datasets, new opportunities for gaining knowledge emerge. Significant hurdles must be overcome to effectively curate incoming AM data and fully exploit analytic opportunities. The next section discusses the datasets being generated across the lifecycle of an additively manufactured part, from raw material to final part.

3 Emerging Data

This section focuses on data collected from metals AM, particularly metal powder bed fusion. The datasets are separated into three subsections: feedstock material, in situ measurements, and ex situ measurements. This data may be collected through experiments, or in part production, to establish a greater empirical foundation.

3.1 Feedstock Material

The characterization of AM feedstock material has become increasingly advanced. Common powder measurements include powder size distribution and various flowability, morphology, and rheology measurements. New measurements and measurement techniques continue to emerge.

Often described globally, powder size distribution and powder bed density are increasingly studied at various locations of the build volume [10]. Powder flow measurements, to study how powder spreads during the build, include the use of optical cameras mounted in a build chamber or on a recoater blade. Digital image correlation (DIC) measurements are capturing spread direction and powder velocity. Other powder spreading measurements include stiffness, coulomb damping, rolling friction, coefficient of restitution and angle of response during a powder sweep.

Characterization of powder particles includes chemical analysis (mass spectrometer) and morphology (X-ray Computed Tomography). Rheometer measurements have expanded to include metrics such as: total energy, permeability, torsion, normal force, and apparent density. Laser flash systems are measuring thermal properties of powder, including thermal diffusivity. Humidity and moisture measurements are studied during storage and processing.

Each of these measurements play a key part in characterizing feedstock powder, but their interrelationships, and how they affect the quality of a build, are not yet clear.

3.2 In Situ Measurements

Various in situ measurements are taken during the AM process, including those related to the material, process, and part. Measurements may come from both part production and experimental builds.

Material Layer- Layer-by-layer material monitoring provides insight into the state of feedstock material immediately before it is processed. In powder bed fusion, layerwise optical imaging using cameras above the build platform provides insight into the powder spread before processing. During spreading, measurements of the powder surface and of spreading angles are available with a profilometer. Layering instruments may also be monitored, such as the acceleration and vertical displacement of a recoater blade.

Melt Pool- Melt pool monitoring techniques are sought as a means for evaluating process parameters and providing insight into the final part during process time. Common melt pool measurements include: melt pool temperature, melt pool cooling rates, melt pool size, and melt pool shape.

High-speed thermal imaging cameras are used to measure light intensity for given spatial correlations, a simplification of temperature measurements (rather than true temperature). Supported measurements include melt pool temperatures, cooling rates, and generalizations of melt pool dimensions. Experimental studies (Figure 1) have investigated the effect of varying power, velocity, and scan strategy on melt pool dimensions and melt pool shape [1].

Photodetectors are used to detect light intensity from the build chamber (radiance over build volume) during the build and outputting voltage readings. Melt pools are being monitored with light-sensing cameras equivalent to a photodetector array. These cameras, coaxial with the laser path, are able to generate real time melt pool images and can provide data on melt pool shape, length, width, and location [11].

Measurements to help understand melt pool behavior are becoming increasingly advanced. Reflectance and emittance measurements of a melt pool are being measured with spectral direction emissivity methods. Emissivity, in combination with reflectance and radiance, has been measured to calculate true temperatures of a melt pool [12]. Emissivity values are specific to the AM process and measurement taken, with values dependent on material, measurement wavelength, melt pool temperature (energy density), direction (angle of observation) and surface characteristics (surface roughness).

Part- Part measurements taken in process are being directly correlated with final parts. Thermal measurements are providing insight into part cooling behavior, and optical measurements are being adopted for early defect detection.

3.3 Ex Situ Measurements

Measurements on a final part are critical for testing and qualification purposes. Additional measurements are often made on a witness specimen or witness coupons. These measurements may be destructive, as seen with mechanical testing, or non-destructive, as seen with scanning.

Surface Measurements- Surface measurements are used to qualify against specifications, correlate surface quality with process performance characteristics, and provide ex-situ characterization of melt pool surfaces and tracks, among other applications.

Outside of revisiting existing measurements, new surface characterizations are also being explored. Profile measurements, using techniques such as Scanning Electron Microscopy (SEM), allow detailed comparisons between modeled and produced parts. Datasets include surface images, surface height/ profile measurements, and information about other surface

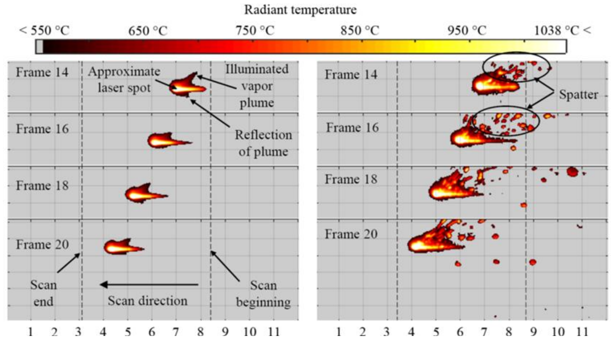

Figure 1: Time-series melt pool data from [1].

characteristics such as cracks and partially melted particles [13].

Microstructure Measurements- Ex situ measurements of part microstructure provide data for correlation with material and process measurements and can provide insight into part performance. These measurements, using methods such as SEM (Figure 2), support the evaluation of microstructure shape including grain orientation, grain size, and grain morphology of sample specimens.

Electron backscatter diffraction (EBSD) and neutron diffraction techniques are used to measure lattice spacing in atomic structures. The shape of a stressed unit geometry can be used to calculate stress and strain tensors from which residual stresses can be derived. Phase-specific stress/ strain measurements provide insight into how changes in crystalline phases affect part properties. The study of phase changes provides insight into the effectiveness of post processes used for stress relief, including heat treatment-induced phase transitions.

X-ray Computed Tomography (CT) measurements- X-ray CT imaging measurements are in the form of grayscale, voxel-based images (voxels located through coordinates) that can be translated to other formats such as STL. X-ray CT imaging allows for measuring

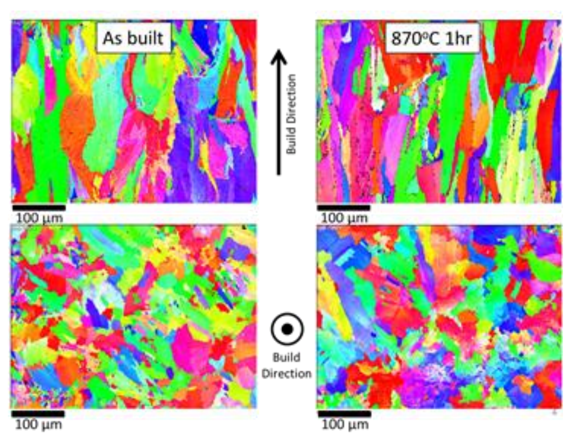

Figure 2: Microstructure images from [2].

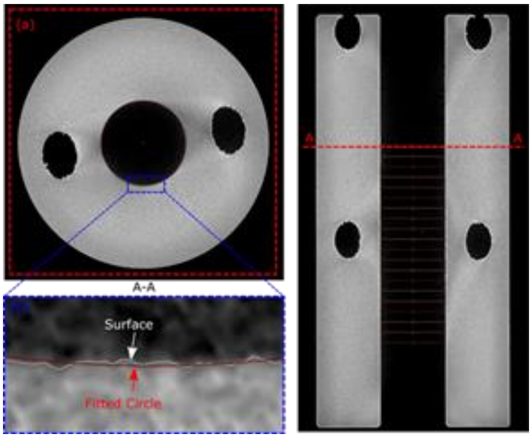

Figure 3: X-ray CT images from [3].

internal surfaces and volumes. Obtained images will vary based on material and scan parameters (e.g. voltage, line, magnification, face).

X-ray CT images (Figure 3) are used to identify and characterize part defects, including voids, porosity, features, and microcracks. Some characterizations are more challenging than others, such as with a “spider-web” like porosity with no well-defined shapes.

X-ray CT allows for the study of failure due to defect propagation initiated by voids. Mechanical testing (e.g. axial loading) of defect-induced specimens can be performed within an X-ray CT to provide insight into how pockets/shapes fail, essentially resulting in time series X-ray CT imaging.

3.4 Analytics from the Bottom-Up

Bottom-up measurements are often taken with specific requirements and analytics tasks. These tasks range from assessing equipment performance to assessing process behaviors to establishing fundamental AM correlations.

Model validation is a common application of measurements driven by bottom-up approaches. Thermal measurements are used to validate melt pool models. DIC measurements can validate powder spreading models. EBSD measurements help predict geometric errors and distortion due to residual stresses while validating distortion and stress models.

Often with multiple ways to measure a specific type of process behavior, measurements are sometimes desired to have insight into how different measurement types, or equipment types, compare. For instance, comparison of a thermal camera signature to a photodetector signature may be desired to see how well the measurements correlate.

When comparing in situ measurements with ex situ measurements, analytics provide insight into process controllability. For instance, in investigating scan strategies, correlations can be made between in situ (size, shape, temperature) and ex situ melt pool measurements (chevron, depth, microstructure, porosity). Surface measurements are being mapped to process characteristics, such as laser power and scan path. Ex situ track images provide information on scan patterns, track paths, cooling, texture, and melt pool profiles. Powder rheology measurements are correlated with powder spreading behaviors. Each of these examples provide insight into process control.

Perhaps the most sought after analytic opportunity, and one that requires both bottom-up and top-down approaches, is establishing expanded correlations between materials, process, part, and geometry. For instance, mechanical-microstructure property relationships are studied to identify correlations with part performance. Correlations between thermal history and microstructure lead to better understanding of how changes in microstructure occur.

While analytics already play an essential role in understanding AM processes, we are far from realizing the full learning potential of the data being collected.

4 Making Sense of it all: The Top-Down Approach

The requirement of a more top-down approach to AM data curation and analytics is based on the premise that to effectively utilize the enormously populous, yet greatly disparate, datasets, actions must be taken prior to the onset of data creation and curation. The goal is to ensure that data is curated in a manner that enables actionable, homogenously characterized data types to support increased analytic opportunities.

Reconciling disparate data types can be a daunting task. When working with many unknowns, asserting common references is problematic. The top-down approach offered here seeks to provide baseline references driven by domain nomenclatures, the domain being additive manufacturing. By initially constraining datasets, though abstractly, we can begin to label [14] the types of data being generated from the top-down. This approach seeks to quickly assert nomenclatures and associativity with datasets as the process dictates. Support for data characterization is perhaps best provided by advanced informatics techniques and algorithms, such as those associated with machine learning and neural networking, and enhanced representations, such as those provided through semantics and category theory.

Using the concepts just discussed, a top-down methodology for curating and analyzing AM data is proposed (Figure 4).

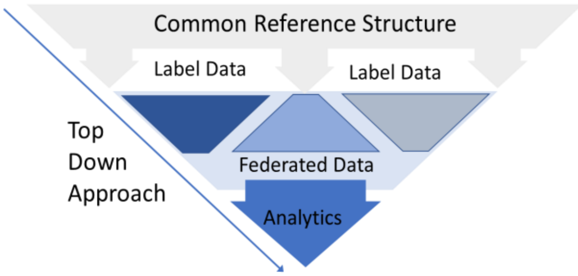

Figure 4: A top-down framework for AM data analytics.

Establishing a Common Reference Structure- A major challenge in effectively analyzing AM data is first establishing common references for how data is stored and accessed. With the data types varying significantly, challenges may come from software capabilities and data structure. The homogenization of heterogenous datasets requires some commonly shared identification mechanisms so that data can be effectively identified and labeled. A key to finding any common reference is utilizing structures that provide domain context and allow for controlled data label convergence.

One method for creating such structures is the adoption of ontology, where concepts can be explicitly defined and related. The explicit formalization of ontology supports well-structured definitions, and the contextualization of data, although the characterization of the data types remains a challenge. The hierarchical nature of the ontology supports the abstraction and subsumption of data labels, and thus the reconciliation of data types through labels.

Labeling and Registration- The post processing of raw data often necessitates human intervention and the methodical interpretation of datasets, a time-consuming process. Here we refer to labeling [14] as a mechanism for augmenting raw data with specific attributes or characteristics. The more attributes assigned to a piece of data or dataset, the better characterized the data becomes. Well-characterized data is essential for data analytics, as the characteristics provide the common context on which an analysis can be performed. Without this context, the meaning of the data can become lost, thus making analytics ineffective.

The time commitment of augmenting raw data can significantly limit the availability of data, especially time-dependent data. Recent advancements in computational techniques such as machine learning promise to lessen our reliance on human interpretation. Automation in the labeling process will effectively increase analytics opportunities.

One step beyond labeling is the process of data registration. Data registration allows data, such as images, to be linked through a single coordinate system. Figure 5 depicts what such a concept could mean if comprehensively applied to AM data. The datasets are registered across time (t), from model, to build data, to layer data, to process data, to microstructure data to X-ray CT data as a means to trace the evolution of a single part. This example of exhaustive traceability can lead to a better understanding of how defects in parts are formed.

Data Federation- The centralization of large datasets can quickly become unmanageable. Additionally, disparities in data types sometimes necessitate alternative storage methods. Utilizing data federation techniques is a necessary step to realizing the learning potential offered by AM datasets. Data federation [15] allows data stored in various locations to be accessed as a sole data source, a notion critical for big data analytics.

Given the variability demonstrated by AM processes, collaborative efforts are often sought to reduce individual investments. The concept of data federation facilitates collaboration, as various entities may share dispersed information and localized expertise.

Analytics- While AM data curation has long been a topic of interest, the advanced levels discussed here are not easily realized. However, they are necessary steps to take advantage of advanced analytics techniques that are becoming increasingly available. The larger and more diverse the datasets, the greater the potential to learn from them. While analytics opportunities exist within the objective-driven, bottom-up approach, greater opportunities are available. A top-down approach to data curation and analysis enables more objective learning approaches, where patterns may initially be sought irrespective of labels and characterization. Deep learning and neural networking techniques support such approaches, and properly preparing for such approaches is essential to realizing the upcoming opportunities.

Each of the steps described in Figure 4 are attainable with today's tools. While much upfront investment is required to set the stage for new analytics opportunities, the payoffs are significant. Training methods with the ability to label and classify datasets will provide the instantaneous enrichment of new datasets. Data patterns will provide new insight into AM correlations and material-process-part relationships. Experimental and physics-based data can become complementary, feeding into new process control opportunities.

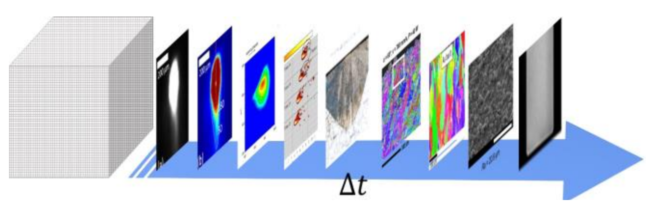

Figure 5. Data registration across time t.

5 Discussion

Until now, much of the contents of this paper have focused on emerging datasets and upcoming analytics opportunities often associated with big data. Big data techniques have demonstrated to be broadly applicable and extremely informative when correctly applied. Such approaches have the potential to revolutionize how data is used in manufacturing. However, most of the techniques discussed are susceptible to failure when data is mischaracterized or mislabeled. Methods are needed to guide, govern, and control the application of advanced analytics. Methods are also need to control data feeds, or to separate the collection of useful data from what may be considered valueless.

Closing the loop on the framework described in Section 4 requires determining and restricting behaviors at multiple scales. Consistently monitoring the conformance of datasets within the set constraints is important. The methods described can be applied in both sequence and in parallel, where errors can easily propagate and can become difficult to identify. Such outcomes restrict real time applications of data analytics. Desired scenarios support the generation and curation of data from the bottom-up while ensuring conformance from the top-down. Ideally a metalanguage provides the constructs with the ability to check for conformance across each of the steps described in Figure 4.

Metalanguages such as category theory can be applied to formally constrain interpretations of information, including those represented as disparate data and datasets. As a metalanguage based on mathematical formalisms, category theory has the ability to mathematically restrict the data and information flow outlined in Figure 4, thus constraining and controlling these transitions. Establishing such trust is crucial to taking the next steps in AM data analytics, where automation could be leveraged in applications such as inline control and programmed material behavior.

6 Summary

Additive manufacturing technologies have made significant advancements over the past five years. No longer viewed as only for toys or prototyping, significant investments have been made into maturing the technologies. One of the byproducts of these investments has been the enormous amount of data generated. This data will play a crucial role as AM technologies continue to mature. This paper investigated the characteristics of this emerging data and proposed a top-down methodology for curating this data in a way that will open new data analytics opportunities in the future.

Literature

[1] J. C. Heigel and B. M. Lane, "The effect of powder on cooling rate and melt pool length measurements using in situ thermographic techniques," in *Solid Freeform Fabrication Symposium*, 2017.

[2] T. Q. Phan, L. E. Levine, M. Stoudt, and J. C. Heigel, "Synchrotron X-ray Characterization of Powder-bed Fusion Laser Melt Traces on Solid Nickel-based Super Alloy Plates," in *TMS2017*, San Diego, CA, 2017.

[3] F. H. Kim, E. J. Garboczi, S. P. Moylan, and J. Slotwinski, "Investtigation of Pore Structure and Defects of Metal Addtive Manufacturing Components using X-ray Computed Tomography," *Additive Manufacturing*, 2017.

[4] T. Wohlers, *Wohlers report 2016*. Wohlers Associates, Inc, 2016.

[5] ANSI, "AMSC Standardization Roadmap for Additive Manufacturing," 2017, Available: https://www.ansi.org/standards_activities/standards_boards_panels/amsc/.

[6] S. Jacobson and R. Franzosa, "Predicts 2016: Opportunities Abound for the Factory of the Future to Reach Its Potential," vol. 25 November 2015, Available: https://www.gartner.com/doc/3172033/predicts--opportunities-abound-factory

[7] D. B. Kim, P. Witherell, R. Lipman, and S. C. Feng, "Streamlining the additive manufacturing digital spectrum: A systems approach," *Additive Manufacturing*, vol. 5, pp. 20-30, 2015.

[8] P. W. Witherell, Y. Lu, S. C. Feng, and D. Kim, "Towards a Digital Thread and Data Package for Metals Additive Manufacturing," *Journal of ASTM International*, 2017.

[9] A. Labrinidis and H. V. Jagadish, "Challenges and opportunities with big data," *Proceedings of the VLDB Endowment*, vol. 5, no. 12, pp. 2032-2033, 2012.

[10] J. Whiting and J. Fox, "Characterization of Feedstock in the Powder Bed Fusion Process: Sources of Variation in Particle Size Distribution and the Factors that Influence them," in *Solid Freeform Fabrication Symposium*, 2016.

[11] J. C. Fox, B. M. Lane, and H. Yeung, "Measurement of process dynamics through coaxially aligned high speed near-infrared imaging in laser powder bed fusion additive manufacturing," in *Thermosense: Thermal Infrared Applications XXXIX*, 2017, vol. 10214, p. 1021407: International Society for Optics and Photonics.

[12] B. Lane *et al.*, "Design, Developments, and Results from the NIST Additive Manufacturing Metrology Testbed (AMMT)," in *Proceedings of the 26th Annual International Solid Freeform Fabrication Symposium–An Additive Manufacturing Conference, Austin, TX*, 2016, pp. 1145-60.

[13] J. C. Fox, S. P. Moylan, and B. M. Lane, "Preliminary Study Toward Surface Texture as a Process Signature in Laser Powder Bed Fusion Additive Manufacturing," in *2016 Summer Topical Meeting: Dimensional Accuracy and Surface Finish in Additive Manufacturing*, 2016.

[14] H.-L. Chen, K.-T. Chuang, and M.-S. Chen, "On data labeling for clustering categorical data," *IEEE Transactions on knowledge and Data Engineering*, vol. 20, no. 11, pp. 1458-1472, 2008.

[15] J. Lee, P. Mazzoleni, J. Sairamesh, and M. Touma, "System and method for planning and generating queries for multi-dimensional analysis using domain models and data federation," ed: Google Patents, 2008.