ACS APPLIED NANO MATERIALS

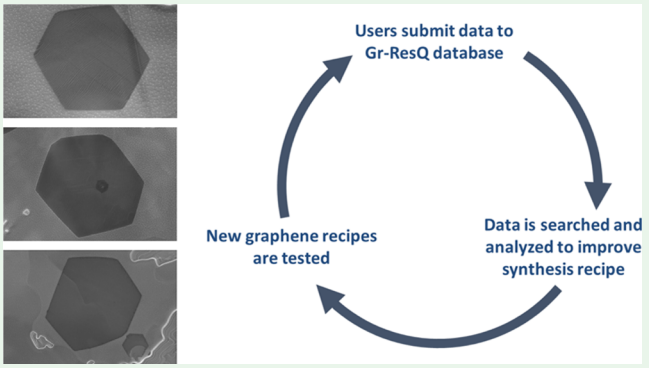
www.acsanm.org Article

Crowd-Sourced Data and Analysis Tools for Advancing the Chemical Vapor Deposition of Graphene: Implications for Manufacturing

Joshua A. Schiller, Ricardo Toro, Aagam Shah, Mitisha Surana, Kaihao Zhang, Matthew Robertson, Kristina Miller, Kevin Cruse, Kevin Liu, Bomsaerah Seong, Chae Seol, Ian T. Foster, Ben J. Blaiszik, Ben Galewsky, Darren Adams, Daniel S. Katz, Placid Ferreira, Elif Ertekin,* and Sameh Tawfick*

Cite This: *ACS Appl. Nano Mater.* 2020, 3, 10144−10155 Read Online

ABSTRACT: Industrial production of graphene by chemical vapor deposition (CVD) requires more than the ability to synthesize large domain, high-quality graphene in a lab reactor. The integration of graphene in the fabrication process of electronic devices requires the cost-effective and environmentally friendly production of graphene on dielectric substrates, but current approaches can only produce graphene on metal catalysts. Sustainable manufacturing of graphene should also conserve the catalyst and reaction gases, but today the metal catalysts are typically dissolved after synthesis. Progress toward these objectives is hindered by the hundreds of coupled synthesis parameters that can strongly affect CVD of low-dimensional materials and poor communication in the published literature of the rich experimental data that exists in individual laboratories. We report here on a platform, "graphene recipes for synthesis of high quality material" (Gr-ResQ; pronounced graphene rescue), which includes powerful new tools for data-driven graphene synthesis. At the core of Gr-ResQ is a crowd-sourced database of CVD synthesis recipes and associated experimental results. The database captures ∼300 parameters ranging from synthesis conditions such as a catalyst material and preparation steps, to ambient lab temperature and reactor details, as well as resulting Raman spectra and microscopy images. These parameters are carefully selected to unlock the potential of machine-learning models to advance synthesis. A suite of associated tools enable fast, automated, and standardized processing of Raman spectra and scanning electron microscopy images. To facilitate community-based efforts, Gr-ResQ provides tools for cyber-physical collaborations among research groups, allowing experiments to be designed, executed, and analyzed by different teams. Gr-ResQ also allows publication and discovery of recipes via the Materials Data Facility, which assigns each recipe a unique identifier when published and collects parameters in a search index. We envision that this holistic approach to data-driven synthesis can accelerate CVD recipe discovery and production control and open opportunities for advancing not only graphene but also many other 1D and 2D materials.



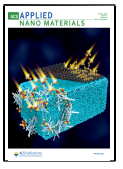
KEYWORDS: *graphene, robotic synthesis, materials data, recipes, machine learning, Raman spectroscopy, nanomanufacturing*

INTRODUCTION

Owing to its versatile electronic properties and atomic thinness, graphene has applications spanning electronic devices, sensors, and transparent electrodes.[1−5] Most such applications require high-throughput, controllable manufacture to enable industrially relevant use. To meet the requirements of electronic device fabrication, graphene synthesis needs to be very tightly controlled, leading to predictable and repeatable layer count and defect density. For example, synthesis via chemical vapor deposition (CVD) uses metal catalysts that are typically dissolved after synthesis, which is not only uneconomical but also unsustainable because of the acids used in dissolving. Moreover, the transfer of graphene from the metal catalyst to dielectric substrates, which is needed for electronic device functionality, creates many defects, hindering its use in commercial applications. The ideal synthesis process should produce high-quality graphene on dielectric substrates, an outstanding objective notwithstanding the current progress.[6−8]

Yet despite focused efforts from the research community, breakthroughs in synthesis are infrequent. New methods are needed to accelerate scientific discovery around CVD growth

Received: August 3, 2020
Accepted: August 27, 2020
Published: August 27, 2020



https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

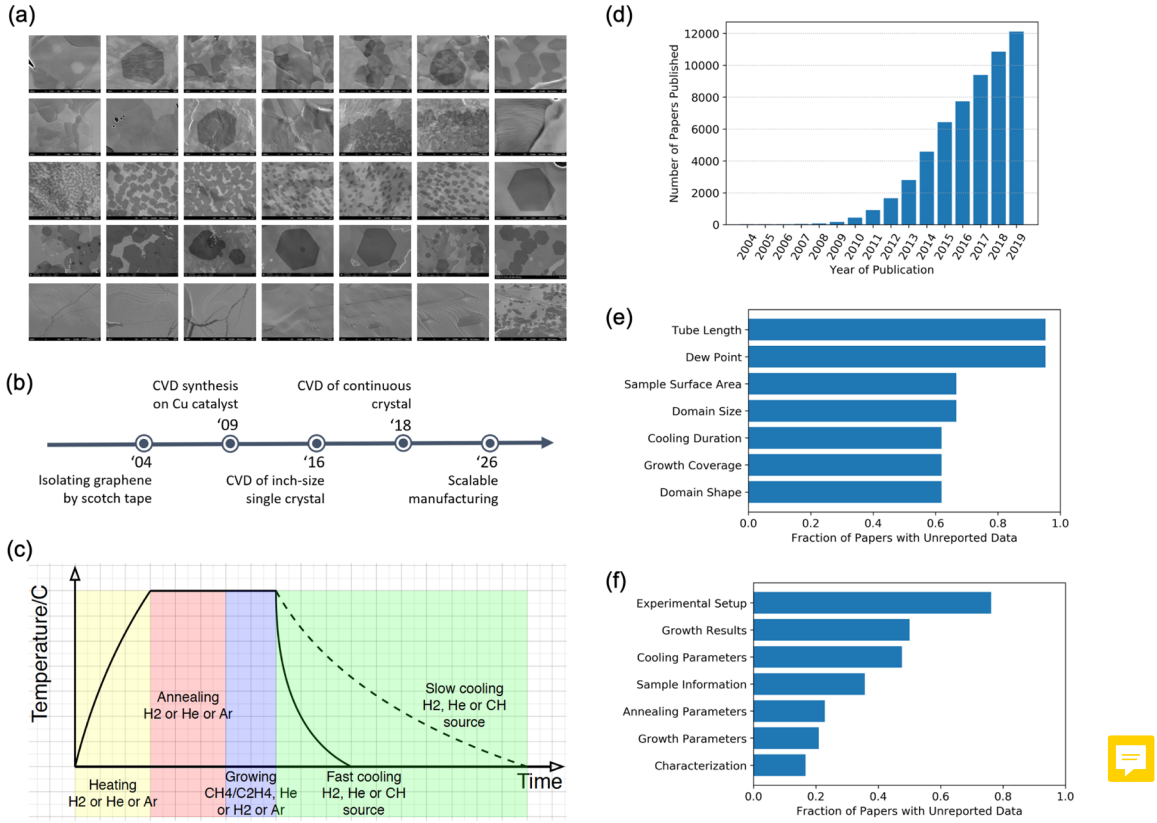


Figure 1. (a) Examples of SEM images of graphene synthesized by CVD (images acquired by the authors). (b) Past graphene milestones and future goals.[10−12] (c) Schematic illustration of a typical graphene CVD synthesis recipe. (d) Number of relevant publications since graphene's discovery. (e) Percentage of unreported information pertaining to seven phases in the synthesis process, collected from a list of 21 highly cited publications related to graphene synthesis. (f) Top seven most unreported parameters from same publications as (e).

of graphene. Two challenges must be addressed: (i) the sensitivity of the resulting material to growth parameters and (ii) the dispersed knowledge available at the level of individual research groups, combined with a culture of incomplete communication of synthesis parameters when reporting in the literature. Figure 1a shows examples of scanning electron microscopy (SEM) images of graphene growth from a small number of experiments using approximately the same recipe. Understanding these images requires deep expertise, not only to identify features in the images but to establish causality given the extreme sensitivity of the growth process. To address these challenges, we have developed the platform "graphene recipes for the synthesis of high quality materials" (Gr-ResQ, pronounced graphene rescue)[9] to enable community-scale sharing of recipes for CVD graphene synthesis.

Figure 1b shows a timeline of past milestones and future goals for graphene synthesis via CVD. After the initial scotch tape exfoliation and discovery of graphene in 2004,[10] research groups around the world shifted their attention gradually from carbon nanotubes (CNTs) to this new exciting material. Fortunately, the reactor designs, growth conditions, and gases for both are similar. The general idea of the CVD recipe is similar in graphene to CNTs (Figure 1c). Synthesis of multilayered graphene on nickel was soon followed by the synthesis of single layers of graphene on copper in 2009 (Figure 1b). Large amounts of carbon can be dissolved in nickel, with most growth taking place during cooling in the form of precipitation and segregation. On the other hand, copper has low carbon solubility, which results in synthesis that is governed by surface adsorption. Mixing copper and nickel in carefully selected ratios within the catalyst, combined with fluid mechanics engineering in the reactor, which led to the production of large domain sizes exceeding a square inch of high-quality graphene in 2016[11] and 2019.[12] Note that it took around a decade to increase the domain size from microns to tens of millimeters. As of the beginning of 2020, there are already more than 12,000 publications per year with the words "graphene synthesis" in the title, abstracts, or keywords: see Figure 1d.

The graphene CVD process, illustrated schematically in Figure 1c, is typically divided into three stages: annealing, growth, and cooling. During the annealing step, a substrate is heated in order to prepare it for growth. During the growth phase, carbon-containing precursors are added, allowing the carbon atoms to crystallize on the substrate. In the final cooling step, the material is brought back down to room temperature. In some CVD processes, growth also occurs in this latter step. While this procedure may seem straightforward, there are hundreds of tuning parameters involved, resulting in a hyper-dimensional recipe phase space. While this can also be true for other chemical synthesis processes, CVD of atomically thin materials remains among the most challenging. First, the volume of material produced is small compared to the reactor size (ratio of $\sim 10^{-9}$), which makes synthesis sensitive to minor local variations in the growth chamber. When two chemical processes are taking place—such as crystallization and etching, with the latter often associated with the presence of uncontrolled amounts of oxygen or hydrogen—the thin nature

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

of the material can entail greater sensitivity to the deleterious consequences of the competing process. Second, while wet chemical synthesis benefits from mixing of the reaction products, atomically thin materials have limited mobility after they crystallize on the surface. Third, gas flow mechanics and heat transfer can affect the kinetics of the monolayer synthesis, altering the results in unpredictable ways. It is often said that graphene synthesis is so sensitive that it may be affected "when a researcher sneezes in the lab"! Unfortunately, the tedious trial and error needed to develop synthesis recipes is an accepted reality, and synthesis research benefits only qualitatively from physical modeling and simulations.

The incomplete nature of synthesis parameters reported in the published literature hinders progress toward manufacturing goals in several ways. First, it makes replicating a given growth experiment challenging because all parameters needed to precisely repeat the recipe are not available. For instance, the furnace setup itself affects the graphene crystal size and nucleation density by influencing the fluid flow, heat transfer, and growth kinetics. However, the total flow rate is often not reported (typically, only ratios between reaction gas flow rates are); nor are the tube dimensions, heated zone length, or flow resistance across the tube. Other important yet often ignored parameters are the reactor dew point and oxygen content. The dew point in particular is typically an "uncontrolled" parameter, and usually not measured despite its critical role. Its role has first been discovered for CNTs: within a narrow range, varying dew points can produce single-walled or multi-walled CNTs, or even completely inhibit CNT growth.[13] For graphene, water vapor and oxygen in the furnace or reaction gas affect nucleation, defect density, and growth rate.[14−17] Minute oxygen content can be present in the form of impurities/surface oxides in the catalyst, or neighboring oxides in the reactor. While measuring local variations in oxygen or hydrogen content on the catalyst surface during the reaction is currently not tenable, dew point can be measured upstream or in the ambient lab environment itself. Yet not only are recipe parameters such as dew point or oxygen content often ignored in the scientific literature, so too are characterizations of the products of recipes such as graphene coverage area, nucleation density or domains per area, and growth rate. Even those growth conditions that are reported are scattered within articles and their supporting online information, making it easy to miss important parameters. These incomplete data in the scientific literature also prevent the use of data-driven or machine-learning tools across consolidated sets of growth experiments. Almost without exception, only the conditions that successfully yielded graphene growth are reported. Negative results are important, because there are narrow windows of synthesis conditions that give rise to the best results, with deviations from these windows dramatically affecting the products.

To quantitatively illustrate the dearth of CVD parameters in published articles, we tabulated 29 parameters of synthesis and characterization from 21 of the most cited publications. These articles were identified via query from the ISI Web of Knowledge for "graphene synthesis using CVD",[18−38] and analyzed to quantify missing data for each of the 29 parameters (analysis presented in Supporting Information). Figure 1e displays the top seven parameters crucial to reproducibility that were most frequently unreported. Figure 1f groups the missing data into seven categories of parameters affecting graphene synthesis:

- experimental setup (e.g., furnace design details and ambient conditions)
- growth results (description of graphene produced including sample coverage, domain size, nucleation density)
- cooling parameters (flow rates, cooling rates)
- sample information (catalyst preparation, oxygen content)
- annealing parameters (exact flow rates vs time, temperature and pressure)
- growth parameters (exact flow rate vs time, temperature and pressure)
- characterization (details about growth uniformity and sample morphology based on imaging and Raman spectra)

The most missing data appear in the experimental setup category and include parameters such as the furnace design, whose characteristics are known to have an effect on flow velocities, heat transfer, and ambient humidity.

The Gr-ResQ platform introduced here has been designed to overcome several of these challenges. The platform is accessible online through nanoHUB,[39] and consists of several features: (i) a repository containing recipes and their associated experimental results that users can search and query, visualize, analyze, and submit to, in order to learn from the aggregate knowledge of the community; (ii) a set of tools for analysis of Raman spectra (to determine graphene quality) and SEM images (to estimate growth coverage and domain orientation); (iii) the capability to apply machine learning methods to the aggregate data; (iv) an interface with the "Operating System for Cyberphysical Manufacturing" (OSCM, pronounced awesome), which allows users to make requests for collaborators to run user-generated experiments and characterize results; and (v) full Gr-ResQ datasets that are published periodically to the Materials Data Facility (MDF)[40,41] where they are archived and assigned a digital object identifier (DOI), and where collected parameters are logged in a search index to promote discovery. As a whole, the platform enables community networking and facilitates an iterative, community-driven learning procedure. Similar platforms for disseminating materials data have already seen success in other fields. In the area of polymer nanocomposites, NanoMine has been established as a platform for documenting and analyzing material properties.[42] In the theoretical space, Materials Project[43] and OQMD[44] provide access to large troves of material properties derived from high-throughput density functional theory simulations. Within an individual research group, Gr-ResQ can be used to track and document synthesis conditions and results. Across the community, it can augment the collective talent and resources of researchers across the world.

Gr-ResQ also adheres to the FAIR data principles.[45] For data to be considered FAIR, they should be findable, accessible, interoperable, and reproducible. In this work, the individual data and complete data sets are both findable and accessible for free to the public through the Gr-ResQ interface as well as via MDF web and programmatic interfaces. We have also worked to make the data interoperable, reproducible, and indeed extendable by defining MDF schemas that describe and represent key elements of the processing, structure, and property results captured by the experiments. These schemas are openly available for reuse, and we also reuse schemas where

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

possible (e.g., DataCite to represent general author information, descriptions, and literature linkages and NIST material-specific schemas to represent high level aspects of the experiments). We see the work here as critical to achieving fully reproducible work, across labs, in the graphene synthesis community.

OVERVIEW

Figure 2 illustrates the components of Gr-ResQ. It consists of the database (shown on the right) connected to the various

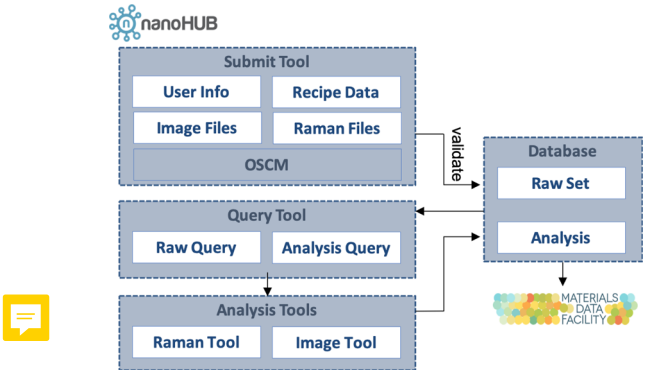


Figure 2. Gr-ResQ platform has three main components: the submission tool, the query tool, and the analysis tools. Raw data are fed into the database with the submission tool and derived data are added using the analysis tools. These data are then validated and the resulting data set is published to MDF. The submission tool is also compatible with OSCM to allow users to generate and test new recipes.

tools (on the left). Users can launch these tools directly from nanoHUB to upload, analyze, and post-process their own graphene data. As illustrated in Figure 3a, we divide the data into six categories: (1) provenance (author, institution, etc. \enleadertwodots), (2) recipe (annealing, growth, and cooling parameters), (3) other synthesis conditions (furnace setup, catalyst sample details, preparation steps, known contamination, ambient conditions), (4) growth results (user-reported graphene domain size, nucleation density, and coverage), (5) characterization data (raw images, Raman spectra), and (5) analyses (useful information extracted using the image and Raman analysis tools such as the in-plane orientation of graphene domains, Raman peak locations, and intensity ratio). The database currently contains both data from the authors as well as data from the literature.[20,22,24,27,28,30,32]

The submit tool handles data ingestion into the database, uploading recipes, raw images (png, tiff, formats), and Raman spectra (txt). The query tool allows users to search and visualize existing data. Query functionality is designed to be more effective compared to a typical literature search engine. It allows users to search for recipes directly by specifying parameter filters, such as a growth temperature higher than a specific value, or graphene with a specific Raman peak (G'/G) ratio. This query functionality can eliminate the arduous process of combing the literature. To establish a consistent analysis of growth results, users can post-process the graphene SEM images and Raman files, using the custom developed analysis tools. The analysis tool set includes an application that can fit Raman spectra in order to determine graphene quality and layer count. It also includes an image tool that can apply

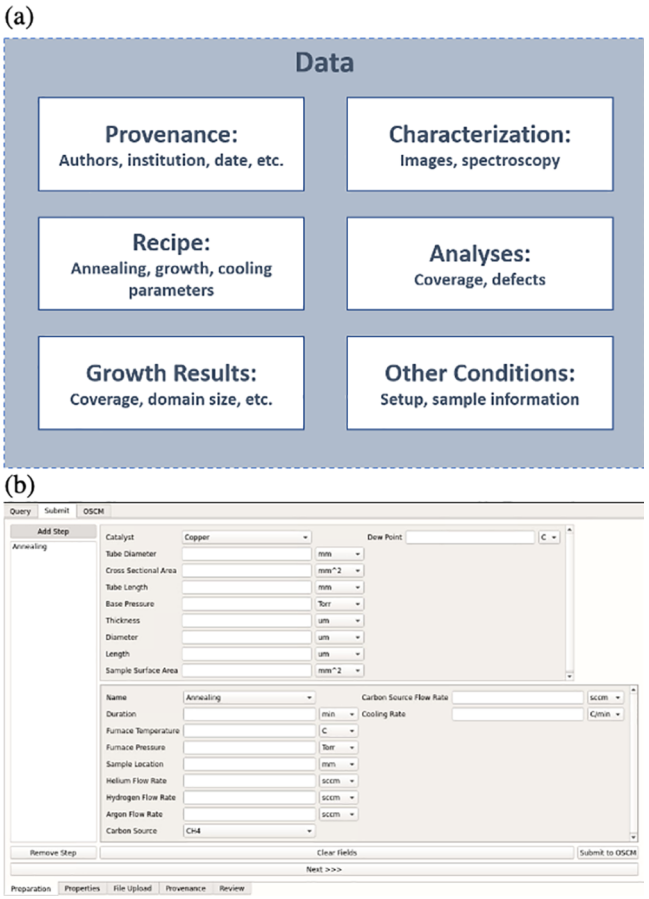


Figure 3. (a) Types of data contained within Gr-ResQ. (b) Recipe submission interface. Users input recipe procedure step-by-step.

masks and image processing algorithms to separate graphene from the substrate.

Gr-ResQ uses the MDF, a set of general materials data publication and discovery services, and the OSCM. MDF provides a stable data archive platform on which to publish periodic releases of Gr-ResQ data and through which users can query and find the Gr-ResQ data. Each data release is a static snapshot of the entire Gr-ResQ database tagged with a DOI. OSCM allows users to crowdsource their experiments by providing instructions to participating scientists, who then manufacture the material and return the associated data. The Gr-ResQ interface allows users to build a step-by-step procedure and submit it to one or more participating labs as an OSCM transaction. An OSCM transaction is completed when the associated experimental data have been uploaded.

The Gr-ResQ platform is developed in Python using the cross-platform PyQt library built on the popular Qt framework to generate graphical user interfaces. PyQt combined with the pyqtgraph plotting package provide the flexibility to develop a host of widgets and tools for data manipulation. On the backend (as of version 1.3.0), we use a MySQL database and the SQLAlchemy toolkit are used to manage recipe data. Finally, our tools are hosted on nanoHUB,[39] a platform for both students and researchers, providing learning materials, software tools, and tools for collaboration. nanoHUB provides a single access point for Gr-ResQ with the computing resources and support to host the tools as well as functionality to manage member access and privileges. Users can run the Gr-

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

ResQ application directly from nanoHUB with no need to install software locally.

We anticipate four types of Gr-ResQ users:

- Querier/analyzer: in the basic use case, this user has open access to the tools and database and has privileges that permit read-only access. This user can search the database for relevant data and visualize the results to facilitate future synthesis experiments. They can also make use of the analysis tool suite to post-process their own data.
- Submitter: a user with writing privileges, who can add new recipes to the database.
- Validator: a core administrator who can verify and validate submitted recipes to ensure data integrity.
- OSCM user: a user who can request experiments through the OSCM interface.

Submit Tool. The submit tool enables users to contribute recipes to the Gr-ResQ database. It uses multi-tab organization for users to sequentially add their data. The tool contains tabs for preparation, properties, file upload, and provenance. The preparation tab, as pictured in Figure 3b, allows users to input the experimental conditions for their recipe and add any number of annealing, growing, or cooling steps. To ease workflow, desired input units can be selected for each parameter. The submit interface also permits users to directly construct recipes for OSCM submission. A user simply inputs their desired recipe and the tool will construct and submit the recipe to OSCM (described further in integration with OSCM). Analysis information, such as coverage and number of layers, can be entered in the properties tab. The file upload tab allows users to upload associated SEM images and Raman spectroscopy data. Uploaded Raman data are automatically scanned and G, G', and D peak positions and amplitudes are calculated. These peaks provide information for users to measure the graphene's quality. For instance, the ratio G'/G is correlated to the number of layers. Users also correlate the uploaded Raman spectra with the sample uniformity by specifying the fraction of the sample which is represented by the uploaded spectrum. This allows our database to determine average values of Raman metrics derived from the peak amplitudes for each sample. Lastly, users can provide their name and affiliation in the provenance tab, so that each recipe can be associated with its author. The data are then reviewed by the user and submitted to the database.

Upon submission, the tool automatically checks for invalid entries. The checks include:

- Checking for author name and institution
- Checking for preparation input
- Checking for a base pressure
- Ensuring all preparation steps have a corresponding temperature, pressure, and duration
- Ensuring all growth steps have a carbon source and flow rate
- Checking for proper Raman spectroscopy format and corresponding characteristic percentage

Should any of these tests fail, the user is notified so that they can correct their input. Should the submission be valid, any other user with validation privileges need only provide a final validation for the recipe to be visible to all users.

Query Tool. The query tool is the access point to the Gr-ResQ database, providing search and visualization capability. With the query tool, users can perform searches on attributes related to the synthesis recipes (temperature, catalyst, etc.), the growth results (number of layers, coverage, etc.), or graphene properties (Raman peaks characteristics). An example search is illustrated in Figure 4a: the rows under "results" become

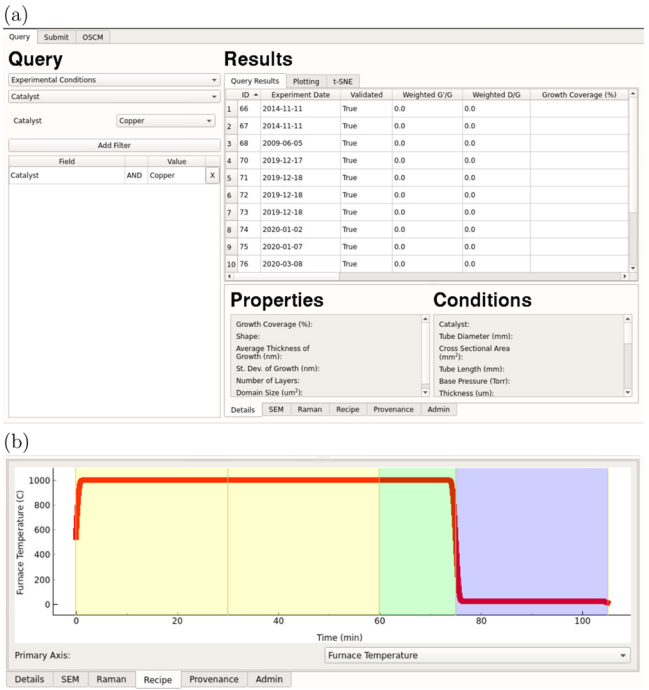


Figure 4. (a) Example search using the query interface to filter recipes that use copper foil as a catalyst. (b) Example recipe visualization of the recipe's furnace temperature.

populated with all database recipes that match the query (in this case, all recipes that use copper foil film as a catalyst), a given recipe can be previewed in the preview pane (Figure 4b) by clicking the entry in the results table. The preview pane is split into six tabs: details, SEM, Raman, recipe, provenance, and admin. The details tab shows information regarding the properties that were entered with the recipe and general information about the experimental conditions (tube size, catalyst, etc.). SEM images and associated masks corresponding to regions of the image classified as graphene or substrate can be found in the SEM tab, while the Raman tab shows plots of associated Raman data. The recipe tab allows users to visualize the synthesis parameters such as temperature, pressures, or gas flow rates versus time. An example recipe visualization is illustrated in Figure 4b. Finally, the provenance tab contains the name and affiliation of the recipe contributor.

Analysis and Plotting. To enable statistical analysis and machine learning of Gr-ResQ data, the "plotting" and "t-distributed stochastic neighbor embedding (t-SNE)" tabs within the Query Tool allows users to visualize correlations between parameters in the recipe data. In Figure 5, two methods of visualization are illustrated. The "plotting" tab allows users to construct XY scatter plots from data. For instance, Figure 5a shows how the G' to G Raman peak correlates to the carbon flow rate across the queried data.

The t-SNE tab allows search for clusters of data that correlate to a selected metric. A schematic representation of this can be seen in 5b. t-SNE is a dimensionality reduction algorithm that seeks to preserve, in a low dimensional space, the relative similarity between points in a high dimensional

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

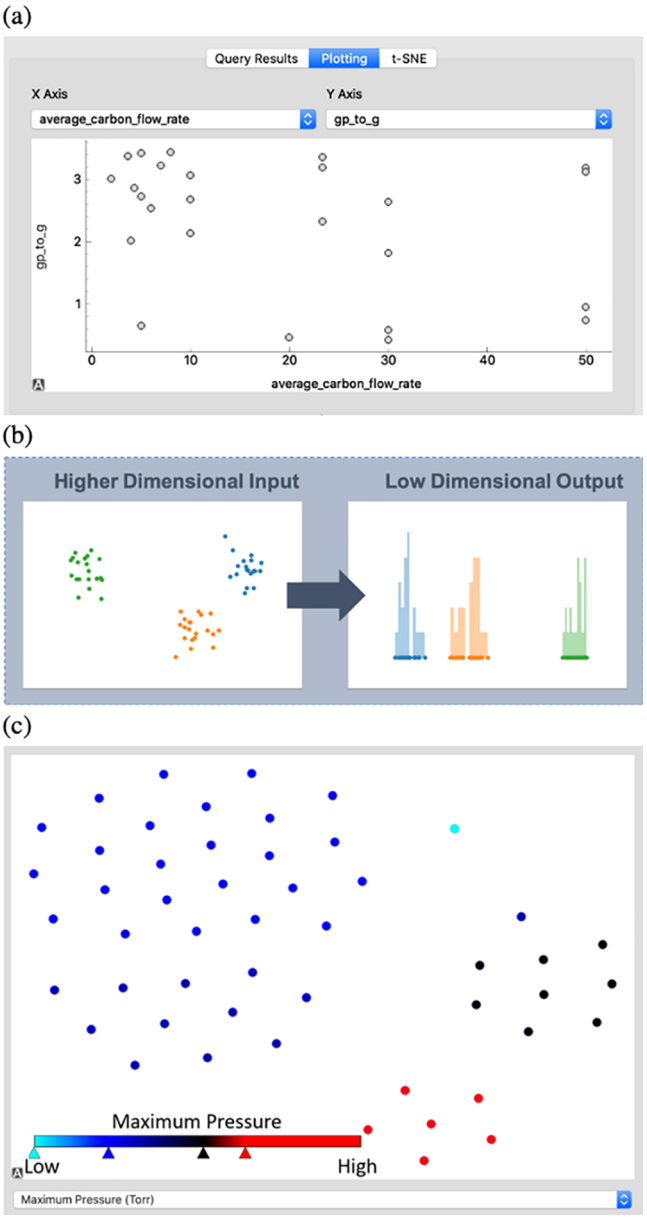


Figure 5. Query tool allows users to visualize sets of recipe data as conventional plots or t-SNE plots (a) example of conventional plotting allowing direct comparison of synthesis criteria to output. (b) Schematic representation of a t-SNE calculation. Low dimensional visualizations can be used to look for clusters of synthesis parameters that correlate with a desired output. (c) Example of the t-SNE plot for 13 parameters colored according to maximum pressure.

space: points that appear close to each other on the plot are more similar in their attributes. Dimensionality reduction is accomplished by minimizing the Kullback–Leibler divergence between a constructed Gaussian joint probability distribution of high dimensional data and a "Student-*t*" distribution of points in low dimensional space.[46] The positions of the data points in reduced dimensional space are optimized to best represent clusters of similar data in the high dimensional space. These low dimensional points can then be colorized according to a desired metric to visualize graphene properties, to look for areas of the recipe phase space that correlate to a desired metric. For example, in Figure 5c, each data point corresponding to a particular recipe has been colorized based on its maximum pressure. t-SNE enables visual inspection and search for patterns in recipes that potentially contain tens or hundreds of attributes.

A challenge to obtaining meaningful information from t-SNE is that some recipes may be missing input for selected attributes. Consequently, it can be difficult to know which sets of features have sufficient support. Gr-ResQ allows users to first determine the fraction of rows that contain valid data for a particular set of features. This fraction is referred to as support and is defined as

$$\mathrm{supp}(F) = \frac{|\{d \in D | F \subseteq d\}|}{|D|} \quad (1)$$

where d represents the set of features that are valid for a particular row, D is the set of d for all rows, and F is the chosen set of features. Using the apriori algorithm, the tool determines the support for all potential feature sets across the selected data set. Using the t-SNE widget, users can then choose a set of features according to a desired level of support or choose a set of features manually. Figure 5c shows a t-SNE plot for 12 parameters with a support of 0.8. The distances between two points reflect their relative similarities in the 12-dimensional space. A clustering pattern is observed when the data are colored according to the maximum pressure. Experiments with similar maximum pressures tend to be clustered together in a higher dimensional space, likely due to similar experimental procedures in each of the clusters.

Image Tool. The Image Tool is an image processing platform customized for analysis of SEM images of graphene, with potential applicability to other types of images as well.[47] SEM images are typically used to observe the growth results of a synthesis experiment, such as areal coverage, nucleation density, and the shape, size, and quality of graphene domains. While the visual inspection of images can sometimes be sufficient to determine the quality of graphene, it is desirable to determine quantitative metrics as well. Quantitative metrics can provide for easier comparison between experimental results and are useful as response variables when attempting to predict optimal recipes. Such metrics include the graphene coverage, which measures the fraction of the substrate that contains graphene and average domain size. The image tool provides the capability to estimate these metrics in an automated manner.

Image Segmentation. To calculate the areal coverage of graphene in an image, the pixels need to be classified as corresponding to graphene or nongraphene. The workflow for separating graphene from the substrate comprises a three-step process, as illustrated in Figure 6. A user uploads a grayscale image to the program, removes unnecessary parts of the image (frame, scale bar, etc.), and then creates a "mask" for the area containing graphene. Our tool provides cropping and erasing functionality to remove parts of the image manually but also has the ability to automatically remove scale bars from the image as well. After isolating the relevant part of the image, the last step is to distinguish pixels corresponding to graphene from those of the background substrate, which can be challenging for images with complex features. Within an SEM image of synthesized graphene, there is typically a mix of graphene (if present), substrate, and random substrate features such as contaminant particles. Depending on the microscope and the synthesis results, SEM images can look quite different. The shape and color of the graphene might vary, as can the contrast between the graphene and substrate. Consequently,

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

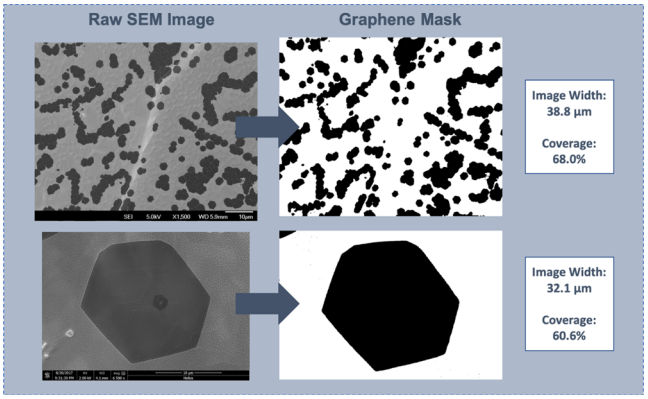


Figure 6. Workflow of a typical user of the image tool to isolate the graphene from the substrate in an SEM image and then generate a corresponding mask for further use. The user first uploads the SEM image and crops the image to remove the scale and then uses the masking tool to isolate the graphene from substrate.

simply filtering based on pixel intensity does not work in many instances. Instead, our tool uses a template matching algorithm, as implemented by OpenCV,[48] to do the bulk of the classification. Because trained scientists can simply look at images and identify graphene, in the image tool users can select one or more portions of the image containing graphene to act as "templates." These templates exhibit features on the pixel level that are consistent wherever graphene is present in the image. The tool scans the image for regions that look like the template according to a selected similarity metric. The default similarity metric is

$$R(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2 \quad (2)$$

where T is a template selected by the user from the image and I is the image. Each pixel is thus assigned a value corresponding to this measure. The user can use a sliding bar to select a cutoff threshold to determine the maximum value for determining whether a pixel matches the template (i.e., corresponds to graphene).

Because the synthesis of aligned graphene crystals (single crystal-like) is often desired, we have added the capability to measure the angular alignment of graphene domains. For instance, many experiments use relatively short synthesis times to obtain large domains of "unstitched" single-crystal graphene domains, which would appear close to hexagonal in shape. For short synthesis times yielding isolated domains of graphene that are not joined, angular alignment measurement allows the observation of distinct domain edges. The domains are often not perfectly hexagonal in shape because of dislocations or other types of defects in their structure. The image tool includes a Sobel filter, which generates a histogram of the distribution of edge angles in the image. This approach is based on convolving the grayscale image with Sobel kernels, which approximate the derivatives of the pixel intensities G_x and G_y in the x-direction and y-direction, respectively. The magnitude and direction of the gradient at each pixel is approximated as

$$G = \sqrt{G_x^{\,2} + G_y^{\,2}}$$

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

where G is the gradient magnitude and θ is the gradient direction. The distribution of the alignment of the domains can be obtained by binning the gradient directions weighted by their magnitudes in a histogram. Perfectly aligned domains will show a comb function, with a periodicity of 60° in the

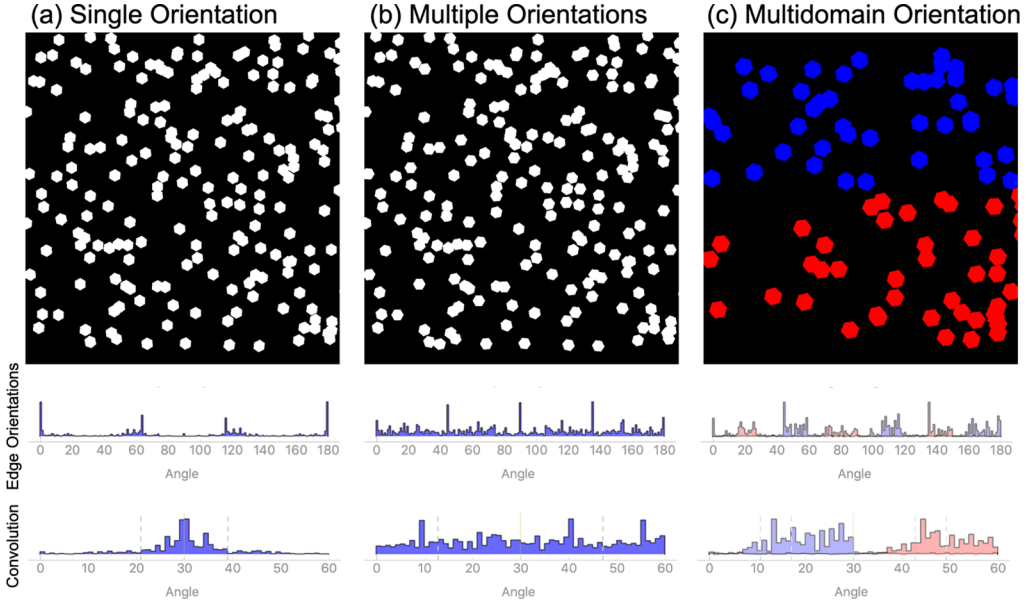


Figure 7. Edge alignment analyses of artificially constructed SEM data for (a) a single set of aligned hexagons, (b) three sets of hexagons with different orientations, and (c) two sets of hexagons with different orientations that are located in separate domains. All hexagons have noise added to their angular orientation. The top row shows the original images, whose length scale is arbitrary for this analysis. The edge orientation histogram illustrates the distribution of edges and the convolution plot is the result of the histogram plot convolved with the comb function. The periodic peaks in the orientation histogram demonstrate alignment of the domains. Histograms are normalized.

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

histogram. Figure 7 illustrates an artificially constructed example to demonstrate. In Figure 7a, the hexagons are oriented about the central angle with noise added. After applying the Sobel filter and binning the angles θ weighted by their magnitude G, a periodic pattern of peaks emerge in the resulting histogram. These peaks correspond to the average orientation of the edges of the hexagons, with the variance of the peaks arising from the distribution of orientations as well as error induced by aliasing. (By contrast, randomly oriented graphene would have no clearly discernible peaks as the edges would be uniformly distributed). We can then convolve this histogram with a comb with a periodicity of 60° and re-center the resulting convolution so that its center of mass is situated at 30°. The variance of the convolution indicates the quality of graphene, where lower variances correspond to a greater domain alignment. Figure 7b illustrates the process for hexagons with three different respective orientations scattered about the image. Correspondingly, there are three peaks in the convolution. Lastly, Figure 7c shows this procedure for an image representing growth on two different grains of a polycrystalline substrate. Again, the histogram shows the differently aligned groups and the resulting convolution displays a peak for each set of hexagons.

In summary, the SEM image tool provides the following functionality:

- Binary masking: converts all nonwhite pixels to black
- Blur: adds Gaussian blur to the image
- Canny edge detector: detects edges in the image
- Color mask: mask a portion of the grayscale spectrum
- Crop: crop the image
- Dilate: thicken edges in the image
- Domain centers: mark centers of graphene domain
- Draw scale: used to determine the number of pixels per unit length
- Erase: erases part of the image
- Erode: thins edges in the image
- Filter pattern: masks part of the image using template matching
- Remove scale: removes certain types of scales found on SEM images
- Sobel filter: applies a Sobel filter to the image

Raman Tool. The Raman tool analyzes Raman spectra of graphene but can potentially be used for other materials as well. Raman spectroscopy is commonly used to check the presence, quality, and the thickness of graphene. Typically, these spectra are analyzed manually, which can be difficult if the spectra are noisy. The Raman tool quantitatively assesses the Raman spectrum of graphene and returns relevant information. A user first uploads their raw Raman data as a .txt or .csv file. The tool applies a baseline correction by fitting the baseline to a quadratic using the method of least squares. It then uses the method of least squares to fit each of the three peaks—D, G and G'—to a Lorentzian function

$$f(x; x_0, \gamma, I) = I\left[\frac{\gamma^2}{(x - x_0)^2 + \gamma^2}\right] \quad (3)$$

where I is the peak height, x_0 is the peak location, and γ is the half-width at half-maximum. The tool outputs the fitted parameters and curves, the number of layers and the quality of graphene (ratio of the peaks in the D and G bands). Figure 8 shows two examples, one of which represents a noisy spectrum.

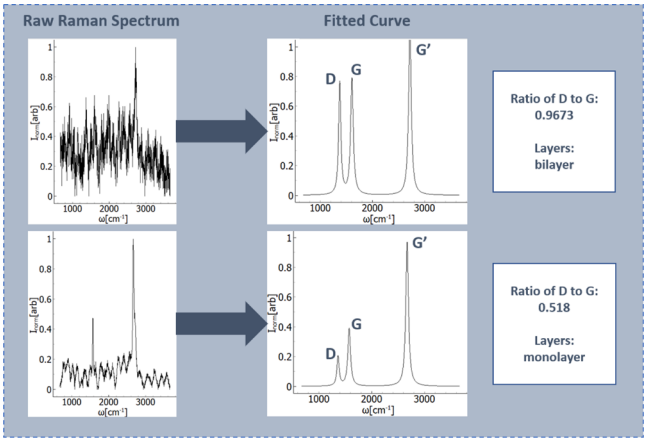


Figure 8. Examples of the use of the Raman tool.

Integration with the MDF and DOI Minting. The MDF is a set of data services designed to simplify data publication, automate metadata extraction, and discover heterogeneous materials science data sets. Its data publications service enables publication of data sets on distributed storage endpoints, access to the data sets using Globus[49,50] or standard HTTPs, invocation of custom extraction scripts to collect metadata to promote discoverability (e.g., experimental parameters, derived analysis results) from published files, and generation of citable DOIs. The MDF Discover service serves as a cloud-hosted metadata catalogue of the data set contents and the extracted metadata, enabling researchers to perform complex and granular queries against the published data set contents (e.g., match all data sets with furnace temperatures between two values or match all data sets published by a given author) and retrieve the matching metadata and data set contents. These publications and discovery capabilities are accessible through both web and programmatic (REST and Python) interfaces allowing integration with the other services needed to build Gr-ResQ. Specifically, the Gr-ResQ platform leverages the MDF services to publish data including recipes, spectrographs, electron micrographs, as well as derived analysis results and to catalogue the associated parameters and other extracted metadata. Dependency graphs are maintained in the metadata between the submitted data sets and derived analyses to make attribution of work clear and easily reportable.

In Gr-ResQ, each recipe is represented as an MDF dataset. Data sets are comprised folders that contain files including the recipe parameters in a custom JSON format, individual images, and raw Raman spectra. Upon publication, custom extractions are performed automatically on the contents of these folders, and the resulting metadata are dispatched to the MDF Discover search index. We append additional data for the Raman analysis, for which a new searchable schema was developed to support meaningful queries of the results. Platform users can use the MDF Forge Python client[51] to search the repository and find recipes based on many criteria, including these derived features of the Raman analysis. Releases of newly acquired Gr-ResQ data are periodically uploaded to MDF from the Gr-ResQ database. As an incentive to users, system administrators collect related recipes into an omnibus data set, and once published a DOI is assigned to the new version of the collection of recipes. This DOI minting mechanism enables tracking citations and other metrics related to the users' contributions to the database. To do so, Gr-ResQ

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

will periodically mint a DOI associated with all recipes provided by a particular research group or to a sub-collection of recipes considered to be a "research study" by the group. In this way, the DOIs will reference contributions on a collection basis as opposed to a unique DOI-per-recipe, limiting the cost and tedium that stems from having to mint or cite potentially thousands of DOIs. These group DOIs will be updated as contributing groups add more data to the database. For example, this can be done in preparation for a publication, where new recipes which are part of a specific research project, can have a single DOI. The current collection of Gr-ResQ datas ets in the MDF can be viewed using the service's search page[52] or using the MDF Forge Python client.

Integration with OSCM. The OSCM is a full-stack operating system to manage manufacturing hardware (machines), manufacturing data (databases), and manufacturing software (applications), in networks of cloud manufacturing.[53] We have integrated Gr-ResQ with OSCM to enable collaboration among users. Through OSCM, Gr-ResQ users can request specific experiments from their collaborators within the same research team or across research teams. This platform can enable collaborative experimentation and community-wide design of experiments. The OSCM platform is designed to make manufacturing "transactions" automated, safe, and verifiable across frictionless and scalable networks. A transaction refers to a manufacturing order that has been placed and executed by different users via their OSCM interactions. OSCM has software components to enable collaborations amongst users, including for machine owners and customers to register, administer, and access manufacturing capacity in the network through several end-user applications. OSCM integration with Gr-ResQ is one such end-user application. The OSCM client is embedded into the Gr-ResQ tool to allow users to create and interact with transactions and experimental resources via the OSCM restful API. In this case, the transaction refers to a request to run a specific synthesis experiment.

The integration between OSCM and Gr-ResQ enables users to reserve time in synthesis labs (facilities) and on equipment (specific CVD furnaces), as well as to capture all data related to the graphene synthesis experiment in a single transaction. The transaction contains experimental information such as the facility or CVD furnace, the recipe, who created the transaction (customer), who approved and executed the transaction (provider). The transaction also includes data that is not in the recipe database: live-measured process variables obtained during the experiment such as temperatures, pressures, flows, and so forth.

The interaction between Gr-ResQ users and OSCM is depicted in Figure 9. Here, the customer is a Gr-ResQ user requesting a specific synthesis experiment from their colleague or collaborator. This customer creates a recipe in the OSCM tab inside the Gr-ResQ tool. The recipe and additional information about the CVD furnace are automatically encapsulated as a transaction by the OSCM tool when the customer confirms the submission. The OSCM tool sends the transaction to the OSCM cloud and saves it in its database. The provider, who operates the CVD reactor, then downloads the recipe from the transaction in OSCM Cloud to the computer that controls the CVD furnace. The recipe is then imported as an input by the CVD controller. The provider starts the experiment, and while the process is running, an OSCM client embedded in the controller CVD furnace collects

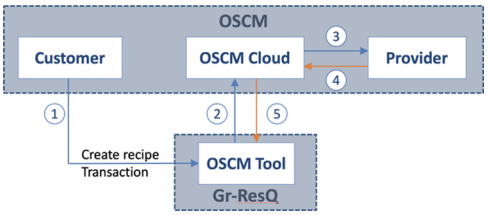


Figure 9. OSCM Tool workflow. Interaction between customer, provider, transactions, and OSCM. (1) Customer creates a recipe in the OSCM tab inside the Gr-ResQ tool. (2) OSCM tool sends the transaction to the OSCM cloud and saves it in its database. (3) Provider downloads and executes the recipe. (4) When the experiment is completed, the process data are then attached back to the transaction information located in OSCM Cloud by the provider. (5) Customer downloads the synthesis result.

and saves the process data (Figure 10). We have developed an application to run a LabVIEW controller of a graphene

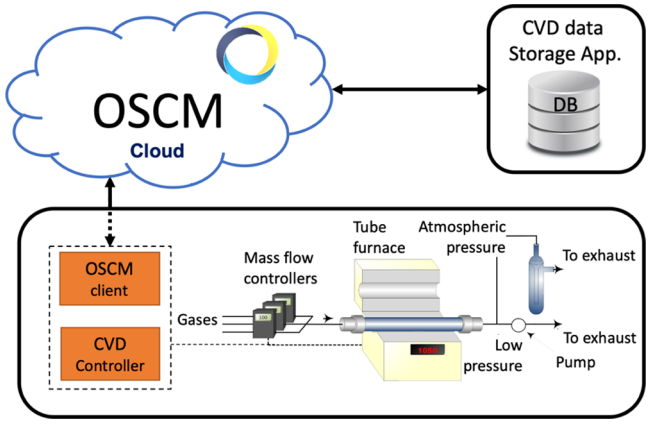


Figure 10. OSCM client is embedded in controller software. It collects and saves all process data, which is then attached to a transaction in OSCM Cloud.

synthesis CVD reactor. Such applications can greatly accelerate research even within the same laboratory. When the experiment is completed, the process data are then attached to the transaction in OSCM Cloud by the provider. Finally, the customer can pull any transaction information from the OSCM Tool for further analysis.

In short, OSCM integration closes the loop between database users and the experimental facilities generating the data. OSCM allows customers to request synthesis experiments by creating transactions using the OSCM Tool based on a requested recipe. From the provider perspective, OSCM allows an operator to download a recipe job, execute it, and upload the result. Both customers and providers are required to be registered OSCM users and can do so via the OSCM Tool or OSCM Cloud.

SUMMARY AND FUTURE WORK

We have introduced Gr-ResQ as a powerful platform which includes a synthesis database and a suite of tools to accelerate progress toward the manufacturable CVD synthesis of graphene. If adopted by the community, this platform could lead to rapid advancement in graphene synthesis, and later in other 2D materials, and enable the integration of graphene in electronic devices on the industrial scale. We view Gr-ResQ as a holistic approach to meet the objectives of graphene

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

manufacturing defined as the scalable, sustainable, cost-effective, and defect-free production of graphene on dielectric substrates for use in future electronics. While basic graphene synthesis has already reached critical milestones, such as large area single domain growth on an optimized metal catalyst, the future goals delineated here still greatly challenge the research community. The required advancements could be more effectively achieved by the community if some of the current hurdles are lifted, such as those associated with finding the most relevant recipes in the vast literature, having sufficient data to quickly replicate specific growth conditions in a different lab, and having large amounts of reliable data to enable the use of machine learning and other data-driven tools. Gr-ResQ addresses these needs by providing tools for researchers to organize and analyze their own data as well as data from the broader community, in the form of SEM and Raman spectroscopy post-processing. It also allows users to upload the results of their own recipes into a centralized and standardized table-like database, allowing researchers to learn best practices for the production of high-quality graphene. The integration of a recipe database with cyberphysical platforms such as OSCM, allows researchers to seamlessly collaborate in community-driven experimentation and machine learning-based modeling. The functional properties of graphene, including various electric, optical, and thermal properties can be added to the database in the future. However, at the current stage, we have not added these properties yet as we need to first establish a consistent protocol for taking and reporting such measurements. This includes details of transferring the graphene, post-processing, and various details of the measurements. Without these details, the reported properties will be unreliable. We are looking forward to community input regarding the inclusion of measured functional properties. The Gr-ResQ framework is flexible and expandable, but we wanted to start from a scope that can be realistically managed at this early stage. Beyond graphene, we believe that a similar approach should be adopted for the synthesis and production of all other 2D materials by CVD,[54] enabling faster realization of technologies that use these new materials.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsanm.0c02018.

Graphene synthesis parameters collected in the Gr-ResQ database, Gr-ResQ software and coding language, and links to GitHub repositories for the source code (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Elif Ertekin − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0002-7816-1803; Email: ertekin@illinois.edu

Sameh Tawfick − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0003-3645-527X; Email: tawfick@illinois.edu

Authors

Joshua A. Schiller − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Ricardo Toro − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Aagam Shah − *Department of Materials Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Mitisha Surana − *Department of Materials Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Kaihao Zhang − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Matthew Robertson − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Kristina Miller − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Kevin Cruse − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Kevin Liu − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Bomsaerah Seong − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Chae Seol − *Department of Mechanical Science & Engineering and Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Ian T. Foster − *Data Science and Learning Division, Argonne National Laboratory, Lemont, Illinois 60439, United States; Department of Computer Science and Globus, University of Chicago, Chicago, Illinois 60637, United States*

Ben J. Blaiszik − *Data Science and Learning Division, Argonne National Laboratory, Lemont, Illinois 60439, United States; Globus, University of Chicago, Chicago, Illinois 60637, United States*

Ben Galewsky − *National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Darren Adams − *National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Daniel S. Katz − *National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Placid Ferreira − *Department of Mechanical Science & Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsanm.0c02018

Notes

The authors declare no competing financial interest.

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

■ ACKNOWLEDGMENTS

J.A.S., R.T., A.S., M.S., K.Z., K.M., K.C., B.S., D.A., D.S.K., P.F., E.E., and S.T. acknowledge support from the National Science Foundation (NSF) through award #1720701. K.Z., M.S., and S.T. acknowledge partial support from the Office of Naval Research ONR award #N00014-18-1-2457 and from NSF #1825300. B.J.B., I.T.F., and B.G. acknowledge support from award #70NANB19H005 from U.S. Department of Commerce, National Institute of Standards and Technology as part of the Center for Hierarchical Materials Design (CHiMaD) and support from NSF Award #1636950.

■ REFERENCES

(1) Neto, A. H. C.; Guinea, F.; Peres, N. M.; Novoselov, K. S.; Geim, A. K. The electronic properties of graphene. *Rev. Mod. Phys.* **2009**, *81*, 109.
(2) Schwierz, F. Graphene transistors. *Nat. Nanotechnol.* **2010**, *5*, 487.
(3) Grigorenko, A. N.; Polini, M.; Novoselov, K. S. Graphene plasmonics. *Nat. Photon.* **2012**, *6*, 749.
(4) Novoselov, K. S.; Fal'ko, V. I.; Colombo, L.; Gellert, P. R.; Schwab, M. G.; Kim, K.; et al. A roadmap for graphene. *nature* **2012**, *490*, 192−200.
(5) Reiss, T.; Hjelt, K.; Ferrari, A. C. Graphene is on track to deliver on its promises. *Nat. Nanotechnol.* **2019**, *14*, 907−910.
(6) Neumaier, D.; Pindl, S.; Lemme, M. C. Integrating graphene into semiconductor fabrication lines. *Nat. Mater.* **2019**, *18*, 525.
(7) Lin, L.; Peng, H.; Liu, Z. Synthesis challenges for graphene industry. *Nat. Mater.* **2019**, *18*, 520−524.
(8) Kawamoto, M.; He, P.; Ito, Y. Green processing of carbon nanomaterials. *Adv. Mater.* **2017**, *29*, 1602423.
(9) Schiller, J. A.; Zhang, K.; Cruse, K. J.; Adams, D. K.; Ertekin, E.; Tawfick, S. H.; Surana, M.; Shah, A. R.; Toro, R. Gr-ResQ. 2019, https://nanohub.org/resources/gresq (accessed 2020-09-01).
(10) Novoselov, K. S.; Geim, A. K.; Morozov, S. V.; Jiang, D.; Zhang, Y.; Dubonos, S. V.; Grigorieva, I. V.; Firsov, A. A. Electric field effect in atomically thin carbon films. *Science* **2004**, *306*, 666−669.
(11) Wu, T.; Zhang, X.; Yuan, Q.; Xue, J.; Lu, G.; Liu, Z.; Wang, H.; Wang, H.; Ding, F.; Yu, Q.; et al. Fast growth of inch-sized single-crystalline graphene from a controlled single nucleus on Cu−Ni alloys. *Nat. Mater.* **2016**, *15*, 43−47.
(12) Vlassiouk, I. V.; Stehle, Y.; Pudasaini, P. R.; Unocic, R. R.; Rack, P. D.; Baddorf, A. P.; Ivanov, I. N.; Lavrik, N. V.; List, F.; Gupta, N.; et al. Evolutionary selection growth of two-dimensional materials on polycrystalline substrates. *Nat. Mater.* **2018**, *17*, 318−322.
(13) Hata, K.; Futaba, D. N.; Mizuno, K.; Namai, T.; Yumura, M.; Iijima, S. Water-assisted highly efficient synthesis of impurity-free single-walled carbon nanotubes. *Science* **2004**, *306*, 1362−1364.
(14) Hao, Y.; Bharathi, M. S.; Wang, L.; Liu, Y.; Chen, H.; Nie, S.; Wang, X.; Chou, H.; Tan, C.; Fallahazad, B.; et al. The role of surface oxygen in the growth of large single-crystal graphene on copper. *Science* **2013**, *342*, 720−723.
(15) Hao, Y.; Wang, L.; Liu, Y.; Chen, H.; Wang, X.; Tan, C.; Nie, S.; Suk, J. W.; Jiang, T.; Liang, T.; et al. Oxygen-activated growth and bandgap tunability of large single-crystal bilayer graphene. *Nat. Nanotechnol.* **2016**, *11*, 426.
(16) Xu, X.; Zhang, Z.; Qiu, L.; Zhuang, J.; Zhang, L.; Wang, H.; Liao, C.; Song, H.; Qiao, R.; Gao, P.; et al. Ultrafast growth of single-crystal graphene assisted by a continuous oxygen supply. *Nat. Nanotechnol.* **2016**, *11*, 930.
(17) Vlassiouk, I.; Regmi, M.; Fulvio, P.; Dai, S.; Datskos, P.; Eres, G.; Smirnov, S. Role of hydrogen in chemical vapor deposition growth of large single-crystal graphene. *ACS Nano* **2011**, *5*, 6069−6076.
(18) Wei, D.; Liu, Y.; Wang, Y.; Zhang, H.; Huang, L.; Yu, G. Synthesis of n-doped graphene by chemical vapor deposition and its electrical properties. *Nano Lett.* **2009**, *9*, 1752−1758.
(19) Dong, X.-C.; Xu, H.; Wang, X.-W.; Huang, Y.-X.; Chan-Park, M. B.; Zhang, H.; Wang, L.-H.; Huang, W.; Chen, P. 3D graphene-cobalt oxide electrode for high-performance supercapacitor and enzymeless glucose detection. *ACS Nano* **2012**, *6*, 3206−3213.
(20) Ma, D.; Liu, M.; Gao, T.; Li, C.; Sun, J.; Nie, Y.; Ji, Q.; Zhang, Y.; Song, X.; Zhang, Y.; Liu, Z. High-quality monolayer graphene synthesis on Pd foils via the suppression of multilayer growth at grain boundaries. *Small* **2014**, *10*, 4003−4011.
(21) Reina, A.; Jia, X.; Ho, J.; Nezich, D.; Son, H.; Bulovic, V.; Dresselhaus, M. S.; Kong, J. Large area, few-layer graphene films on arbitrary substrates by chemical vapor deposition. *Nano Lett.* **2009**, *9*, 30−35.
(22) Xiao, K.; Wu, H.; Lv, H.; Wu, X.; Qian, H. The study of the effects of cooling conditions on high quality graphene growth by the APCVD method. *Nanoscale* **2013**, *5*, 5524−5529.
(23) Yu, Q.; et al. Control and characterization of individual grains and grain boundaries in graphene grown by chemical vapour deposition. *Nat. Mater.* **2011**, *10*, 443−449.
(24) Babenko, V.; Murdock, A. T.; Koós, A. A.; Britton, J.; Crossley, A.; Holdway, P.; Moffat, J.; Huang, J.; Alexander-Webber, J. A.; Nicholas, R. J.; Grobert, N. Rapid epitaxy-free graphene synthesis on silicidated polycrystalline platinum. *Nat. Commun.* **2015**, *6*, 7536.
(25) Bae, S.; et al. Roll-to-roll production of 30-inch graphene films for transparent electrodes. *Nat. Nanotechnol.* **2010**, *5*, 574−578.
(26) Kim, K. S.; Zhao, Y.; Jang, H.; Lee, S. Y.; Kim, J. M.; Kim, K. S.; Ahn, J.-H.; Kim, P.; Choi, J.-Y.; Hong, B. H. Large-scale pattern growth of graphene films for stretchable transparent electrodes. *Nature* **2009**, *457*, 706−710.
(27) Choi, D. S.; Kim, K. S.; Kim, H.; Kim, Y.; Kim, T.; Rhy, S.-h.; Yang, C.-M.; Yoon, D. H.; Yang, W. S. Effect of cooling condition on chemical vapor deposition synthesis of graphene on copper catalyst. *ACS Appl. Mater. Interfaces* **2014**, *6*, 19574−19578.
(28) An, X.; Liu, F.; Jung, Y. J.; Kar, S. Large-area synthesis of graphene on palladium and their Raman spectroscopy. *J. Phys. Chem. C* **2012**, *116*, 16412−16420.
(29) Gomez De Arco, L.; Zhang, Y.; Schlenker, C. W.; Ryu, K.; Thompson, M. E.; Zhou, C. Continuous, highly flexible, and transparent graphene films by chemical vapor deposition for organic photovoltaics. *ACS Nano* **2010**, *4*, 2865−2873.
(30) Gao, L.; Ren, W.; Xu, H.; Jin, L.; Wang, Z.; Ma, T.; Ma, L. P.; Zhang, Z.; Fu, Q.; Peng, L. M.; Bao, X.; Cheng, H. M. Repeated growth and bubbling transfer of graphene with millimetre-size single-crystal grains using platinum. *Nat. Commun.* **2012**, *3*, 699.
(31) Lu, Z.-J.; Bao, S.-J.; Gou, Y.-T.; Cai, C.-J.; Ji, C.-C.; Xu, M.-W.; Song, J.; Wang, R. Nitrogen-doped reduced-graphene oxide as an efficient metal-free electrocatalyst for oxygen reduction in fuel cells. *RSC Adv.* **2013**, *3*, 3990−3995.
(32) Li, X.; Cai, W.; An, J.; Kim, S.; Nah, J.; Yang, D.; Piner, R.; Velamakanni, A.; Jung, I.; Tutuc, E.; Banerjee, S. K.; Colombo, L.; Ruoff, R. S. Large-area synthesis of high-quality and uniform graphene films on copper foils. *Science* **2009**, *324*, 1312−1314.
(33) Li, X.; Magnuson, C. W.; Venugopal, A.; An, J.; Suk, J. W.; Han, B.; Borysiak, M.; Cai, W.; Velamakanni, A.; Zhu, Y.; Fu, L.; Vogel, E. M.; Voelkl, E.; Colombo, L.; Ruoff, R. S. Graphene films with large domain size by a two-step chemical vapor deposition process. *Nano Lett.* **2010**, *10*, 4328−4334.
(34) Cao, X.; Shi, Y.; Shi, W.; Lu, G.; Huang, X.; Yan, Q.; Zhang, Q.; Zhang, H. Preparation of novel 3D graphene networks for supercapacitor applications. *Small* **2011**, *7*, 3163−3168.
(35) Leela, A.; Reddy, M.; Srivastava, A.; Gowda, S. R.; Gullapalli, H.; Dubey, M.; Ajayan, P. M. Synthesis Of Nitrogen-Doped Graphene. *ACS Nano* **2010**, *4*, 6337−6342.
(36) Chen, Z.; Ren, W.; Gao, L.; Liu, B.; Pei, S.; Cheng, H.-M. Three-dimensional flexible and conductive interconnected graphene networks grown by chemical vapour deposition. *Nat. Mater.* **2011**, *10*, 424−428.
(37) Chae, S. J.; Güneş, F.; Kim, K. K.; Kim, E. S.; Han, G. H.; Kim, S. M.; Shin, H.-J.; Yoon, S.-M.; Choi, J.-Y.; Park, M. H.; Yang, C. W.; Pribat, D.; Lee, Y. H. Synthesis of large-area graphene layers on poly-

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155

nickel substrate by chemical vapor deposition: Wrinkle formation. *Adv. Mater.* **2009**, *21*, 2328−2333.
(38) Wood, J. D.; Schmucker, S. W.; Lyons, A. S.; Pop, E.; Lyding, J. W. Effects of Polycrystalline Cu Substrate on Graphene Growth by Chemical Vapor Deposition. *Nano Lett.* **2011**, *11*, 4547−4554.
(39) Madhavan, K.; Zentner, L.; Farnsworth, V.; Shivarajapura, S.; Zentner, M.; Denny, N.; Klimeck, G. nanoHUB. org: Cloud-based services for nanoscale modeling, simulation, and education. *Nanotechnol. Rev.* **2013**, *2*, 107−117.
(40) Blaiszik, B.; Chard, K.; Pruyne, J.; Ananthakrishnan, R.; Tuecke, S.; Foster, I. The Materials Data Facility: Data services to advance materials science research. *JOM* **2016**, *68*, 2045−2052.
(41) Blaiszik, B.; Ward, L.; Schwarting, M.; Gaff, J.; Chard, R.; Pike, D.; Chard, K.; Foster, I. A data ecosystem to support machine learning in materials science. *MRS Commun.* **2019**, *9*, 1125−1133.
(42) Zhao, H.; Li, X.; Zhang, Y.; Schadler, L. S.; Chen, W.; Brinson, L. C. Perspective: NanoMine: A material genome approach for polymer nanocomposites analysis and design. *APL Mater.* **2016**, *4*, 053204.
(43) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1*, 011002.
(44) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. *npj Comput. Mater.* **2015**, *1*, 15010.
(45) Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018.
(46) van der Maaten, L.; Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579−2605.
(47) Schiller, J. A.; Robertson, M. G.; Miller, K. M.; Cruse, K. J.; Liu, K.; Adams, D. K.; Galewsky, B.; Ertekin, E.; Tawfick, S. H. SEM Image Processing Tool. 2018, https://nanohub.org/resources/gsaimage (accessed 2020-09-01).
(48) Bradski, G. *The OpenCV Library*; Dr. Dobb's Journal of Software Tools, 2000.
(49) Ananthakrishnan, R.; Chard, K.; Foster, I.; Tuecke, S. Globus platform-as-a-service for collaborative science applications. *Concurrency Comput. Pract. Ex.* **2015**, *27*, 290−305.
(50) Ananthakrishnan, R.; Blaiszik, B.; Chard, K.; Chard, R.; McCollam, B.; Pruyne, J.; Rosen, S.; Tuecke, S.; Foster, I. *Proceedings of the Practice and Experience on Advanced Research Computing*, 2018; pp 1−7.
(51) Materials Data Facility, MDF Forge Python library. n.d., https://github.com/materials-data-facility/forge (accessed 2020-09-01).
(52) Materials Data Facility, Gr-ResQ_datasets in MDF. n.d., https://petreldata.net/mdf/?filter-match-any.mdf.resource_type=dataset&filter-match-any.mdf.organizations=NanoMFG (accessed 2020-09-01).
(53) Correa, J. E.; Toro, R.; Ferreira, P. M. A new paradigm for organizing networks of computer numerical control manufacturing resources in cloud manufacturing. *Procedia Manuf.* **2018**, *26*, 1318−1329.
(54) Mounet, N.; Gibertini, M.; Schwaller, P.; Campi, D.; Merkys, A.; Marrazzo, A.; Sohier, T.; Castelli, I. E.; Cepellotti, A.; Pizzi, G.; et al. Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds. *Nat. Nanotechnol.* **2018**, *13*, 246−252.

https://dx.doi.org/10.1021/acsanm.0c02018
ACS Appl. Nano Mater. 2020, 3, 10144−10155