Conjugated Polymer Process Ontology and Experimental Data Repository for Organic Field-Effect Transistors

Aaron L. Liu,[a] Myeongyeon Lee,[b] Rahul Venkatesh,[a] Jessica A. Bonsu,[a] Ron Volkovinsky,[a] J. Carson Meredith,[a] Elsa Reichmanis,[b*] Martha A. Grover[a*]

[a]School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA, 30332

[b]Department of Chemical & Biomolecular Engineering, 124 E. Morton Street, Lehigh University, Bethlehem, PA, 18015

*Authors to whom correspondence should be addressed: elr420@lehigh.edu, martha.grover@chbe.gatech.edu

ABSTRACT

Polymer-based semiconductors and organic electronics encapsulate a significant research thrust for informatics-driven materials development. However, device measurements are described by a complex array of design and parameter choices, many of which are sparsely reported. For example, the mobility of a polymer-based organic field-effect transistor (OFET) may vary by several orders of magnitude for a given polymer, as a plethora of parameters related to solution processing, interface design/surface treatment, thin-film deposition, post-processing, and measurement settings have a profound effect on the value of the final measurement. Incomplete contextual, experimental details hamper the availability of reusable data applicable for data-driven optimization, modeling (*e.g.*, machine learning), and analysis of new organic devices. To curate organic device databases that contain reproducible and Findable, Accessible, Interoperable, and Reusable (FAIR) experimental data records, data ontologies that fully describe sample provenance and process history are required. However, standards for generating such process ontologies are not widely adopted for experimental materials domains. In this work, we design and implement an object-relational database for storing experimental records of OFETs. A data structure is generated by drawing on an international standard for batch process control (ISA-88) to facilitate the design. We then mobilize these representative data records, curated from literature and laboratory experiments, to enable data-driven learning of process-structure-property relationships. The work presented herein opens the door for the broader adoption of data management practices and design standards for both the organic electronics and the wider materials community.

1. INTRODUCTION

The domain of π-conjugated polymer semiconductors is of keen interest to both the materials informatics and organic electronics communities due to the promising opportunities this class of materials offers for large-area, printable, deformable electronic devices and energy applications.[1-3] In the years since the advent of the Materials Genome Initiative in 2011, conjugated polymer semiconductors have been associated with over 5,000 peer-reviewed articles (Web of Science, March 2023).[4] A subset of this body of literature has striven to further accelerate knowledge discovery in this materials space through applied data science, machine learning, and high-throughput experimentation techniques. For example, data-driven techniques have recently been leveraged to pursue targeted advances for thin-film device applications including organic field-effect transistors (OFETs),[2, 5-8] organic light-emitting diodes (OLEDs),[9-12] and organic photovoltaics (OPVs).[13-17] Early successes include the application of self-driven laboratory workflows to screen quaternary OPV formulations at the full device level.[13, 18] Indeed, these endeavors have positioned organic electronics as a significant research thrust within the materials informatics community.

Despite recent accomplishments, rational design of organic electronic devices, particularly those that are polymer-based, still largely materialize through one-parameter-at-a-time, hypothesis-driven studies due to the limited availability of representative experimental data. The conjugated polymer materials domain is a research area with a compelling need for experimental data management solutions. While a few examples of shared datasets or databases that target organic electronics research have been reported, such as the Harvard Clean Energy Project[19] and OCELOT,[20] they mostly include computational data on small molecules. A recent effort in Deep4Chem mined over 1,000 peer-reviewed articles to build an experimental database of chromophores,[21] but similar to prior efforts it largely targets electronic structure-property measurements and is not inclusive of process history. In extending database management effectively to polymer-based devices, providing data models that are inclusive of experimental processing information are priorities for storing accurate and reproducible records.

Experimental database design and management for polymer electronics however is nontrivial, especially when a plurality of the relevant attributes related to the provenance of the sample must be included accurately to form a robust, "reusable" data record. Data ontologies are not standardized in the organic electronics space: fully capturing all relevant experimental information is challenging, and organic device performance is highly sensitive to the many parameters associated with the active layer deposition process. For example, the charge-carrier mobility (μ), a key figure of merit for OFETs, has been shown to vary significantly for poly(3-hexylthiophene) (P3HT) ($\sim 10^{-6}$ – 10^{0} $cm^2/V{\cdot}s$)[8, 22] and poly[2,5-(2-octyldodecyl)-3,6-diketopyrrolopyrrole-alt-5,5-(2,5-di(thien-2-yl)thieno [3,2-b]thiophene)] (DPP-DTT) ($\sim 10^{-5}$ – 10^{1} $cm^2/V{\cdot}s$).[8] This performance variation is not only attributed to batch-to-batch characteristics of the polymer, but also a plethora of parameters related to the polymer's process history, starting with the solution state through the thin-film deposition process.[23, 24] Another source of variation is that mobility values are derived from device measurements via model fitting, and employing different methods/parameters (*i.e.*, models, measurement settings, voltage limits) may affect the extracted mobility value. Recording processing and measurement parameters provides indispensable contextual value to organic device data, but nonetheless recording all of them efficiently is not straightforward. Additionally, since the design space is inherently dynamic due to the evolving nature of research, data models must be designed with flexibility in mind without sacrificing consistent vocabulary. Thus, generating a representative data ontology describing the experimental device realm is a challenge that must be addressed to enable reliable database designs.[25]

Though process representations are not new problems for the sake of curating materials databases, navigating these challenges for experimental polymer domains has only been explored recently by a minority of materials data researchers. The experimental database effort in MaterialsMine promotes the inclusion of processing terms for polymer nanocomposites,[26-28] while the Community Resource for Innovation in Polymer Technology (CRIPT) proposes a framework to comprehensively describe polymer data, seeking to unify all aspects of sample provenance from synthesis, processing, characterization, properties, and instrumentation/citation metadata.[29] These active endeavors open the door for a broader adoption of polymer-based data management solutions, but it is up to various communities to enable tailored data models for their specific sub-domains.[30]

In this work, we use OFETs as a model system to propose an experimental data ontology associated with semiconducting polymer processing. We then produced a data structure that focuses on the deposition of the active semiconducting polymer layer and leveraged it to implement an experimental repository relating the semiconducting polymer process history to device performance. To guide a robust representation of that process history, this work draws upon ISA-88, an international standard for automation in batch process control, to construct generalizable relationships across process transformations within the fabrication procedure to create the semiconducting thin film.[31] Building a data repository that can handle the many nuances of this complex design space is expected to provide a platform to enhance hypothesis design, scientific decision-making, and model development within the traditionally “small data” space of the organic electronics community.

2. DATA MODEL AND KNOWLEDGE REPRESENTATION

2.1. Parameter space

Defining the required information to capture is facilitated by published reporting standards for experimental OFET device data.[32-34] An overview of the major materials and process stages involved in depositing the semiconducting polymer layer, and a non-exhaustive set of their related parameters/attributes is presented in **Figure 1**. A device recipe considers the starting materials – a polymer and a device substrate – and tracks these two inputs as they are transformed through a series of process steps and ultimately integrated into the output: an OFET on which a device measurement is made. Important parameters and nuances to evaluate device performance include materials characteristics, solution processing, substrate treatment, and the instrument parameters and models used to extract device metrics.

The primary material parameters describe the components of the active semiconducting layer as well as any other materials (*i.e.*, solvents, chemical treatments) included in the processing procedure. The polymer and any other components are dissolved into a solution that is ultimately deposited onto the device substrate. As the behavior of the polymer in the solution state is a key determinant of its thin-film behavior, all information associated with the solution makeup and its processing is especially crucial to capture for data provenance purposes.[35] Choices in the material characteristics of the polymer (*i.e.*, molecular weight distributions, regioregularity, tacticity, *etc.*) yield a range of structural and morphological motifs that in turn significantly impact device performance.[36] The identity of the solvent(s) used affects not only the polymer-solvent interactions in the solution-state, but also the dynamics of the thin-film deposition process, thereby influencing the structure and performance of the final active layer.[37]

The OFET device substrate is a layered device structure, comprising substrate, gate electrode, dielectric, and source/drain electrodes. The most common electrode configurations used in the organic electronics community include (a) bottom-gate bottom-contact (BGBC), (b) bottom-gate top-contact (BGTC), (c) top-gate bottom-contact (TGBC), and (d) top-gate top-contact (TGTC). The electrode configuration is a necessary contextual detail for reporting a device measurement, as substrate designs often take advantage of charge carrier behavior at different interfaces. Additionally, as substrate design parameters (*i.e.*, channel dimensions such as width and length)[7] and material choices (*i.e.*, electrode material, dielectric, *etc.*) may influence device measurements,[38] this design information is important to include in a data entry to promote experimental reproducibility.[32]

An emphasis of the work herein is that processing information is indispensable for the purpose of storing reproducible device data. Seemingly minor differences in processing can lead to significant changes in the recorded charge-carrier mobility of an OFET sample. Omitting information related to this process history will therefore lead to errors or inaccuracies when comparing device data across experiments. Even prior to deposition, the solution and device substrate undergo process transformations that can affect the deposited thin-film and device characteristics. Solution-based processes may include operations such as sonication, aging, poor solvent addition, cooling, etc.[39] in a prescribed sequence to promote solution-state aggregation,[40-42] while surface pretreatment procedure may include, for example, a cleaning process (e.g., UV-ozone or plasma treatment) followed by a surface modification step *via* self-assembled monolayer (SAM).[43] As an example, differences in solution aging times (*i.e.,* 3 hr, 6 hr, 24 hr prior to coating) can lead to noticeable structural changes that affect the final value of the device measurement in P3HT.[44, 45]

The coating process could be performed through a plethora of solution casting methods including drop casting, spin coating, blade coating, inkjet printing, slot die coating, *etc*. that all have different physical impacts on thin-film morphology and therefore device performance, especially when coupled with solution pre-treatment.[46, 47] Meniscus-guided coating techniques, for example, yield a set of deposition regimes governed by a complex parameter space that includes flow conditions, coating speeds, stage temperatures, drying times, contact angles, *etc*.[48] In combination with solution properties and surface interactions, coating conditions are often chosen carefully to tune the morphology of the deposited thin film. Post-processing operations may also be performed after the coating stage, such as annealing[49] or selective etching,[50] to further control the thin-film morphology. Throughout all processing steps, the ambient environment (humidity, air *vs.* inert atmosphere, temperature, *etc.*) may also play a role in the final properties.[51]

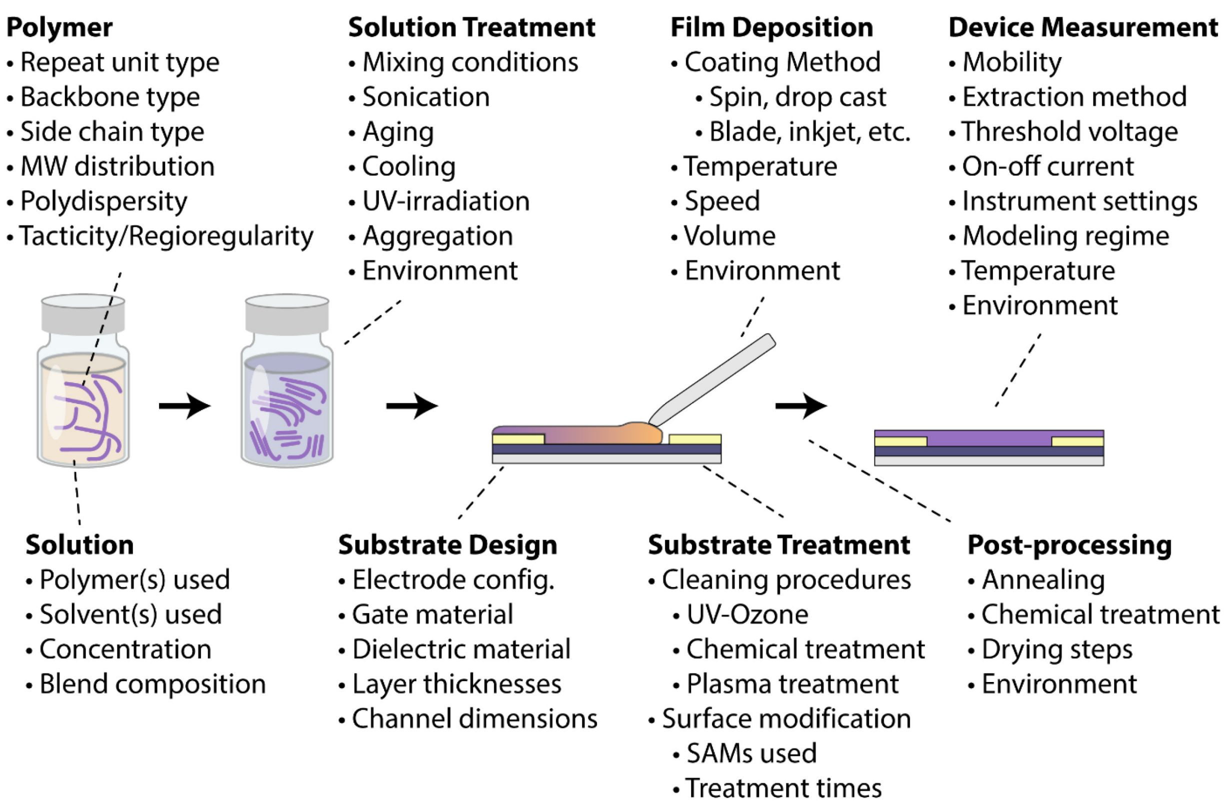

Figure 1. Overview of the process associated with transforming a semiconducting polymer into the active layer of an OFET sample, including the common parameters associated with a sample recipe.

Instrument settings are classified as attributes of the measurement rather than the fabrication process, but this metadata is important as it contextualizes the reported value. Particularly, OFET performance metrics can be nontrivial to represent because a charge-carrier mobility value is not a measurement per se; it is a parameter value derived from curve fitting of the actual measurement, a transfer curve sweep. Therefore, the measurement and fitting protocol used to extract the charge-carrier mobility benchmark from the actual transfer curve is an important consideration. Unreported details about measurement regimes, voltage sweep range/direction, the measurement environment, etc., can lead to misinterpretations about the provenance of device metrics such as the mobility. In some cases, mobilities extracted from the same device data can differ significantly when different extraction methods are used (*e.g.*, space-charge limited current-voltage (SCLC), time-of-flight (ToF), *etc.*)[52] or when different voltage ranges are chosen to obtain the fitted mobility value.[53] Guidelines on robust mobility extraction protocols and measurement metadata reporting are relevant here, and are elaborated on in the literature.[52-54]

2.2. Sample representation

The translation of the real-world parameter space and its relationships to a robust data model requires definition and elaboration of an ontology. We direct the reader to introductory SQL and database literature to facilitate conceptual understanding of the database model enumerated below.[30, 55] At a high level, the entity-relationship diagram in **Figure 2** shows how an organic device sample with its associated reported measurement (*i.e.*, a charge-carrier mobility value) may be conceptually encapsulated as an experimental data record. This diagram provides an important visualization of how various parameters, data, and information in the experimental real world are captured as attributes of related objects, to facilitate the organization of data in constructing a database. Rectangles represent entities or objects, diamonds represent relationships between entities, and labeled ovals represent attributes containing the data or information associated with the various entities, where underlined labels denote a unique identifier for that object.

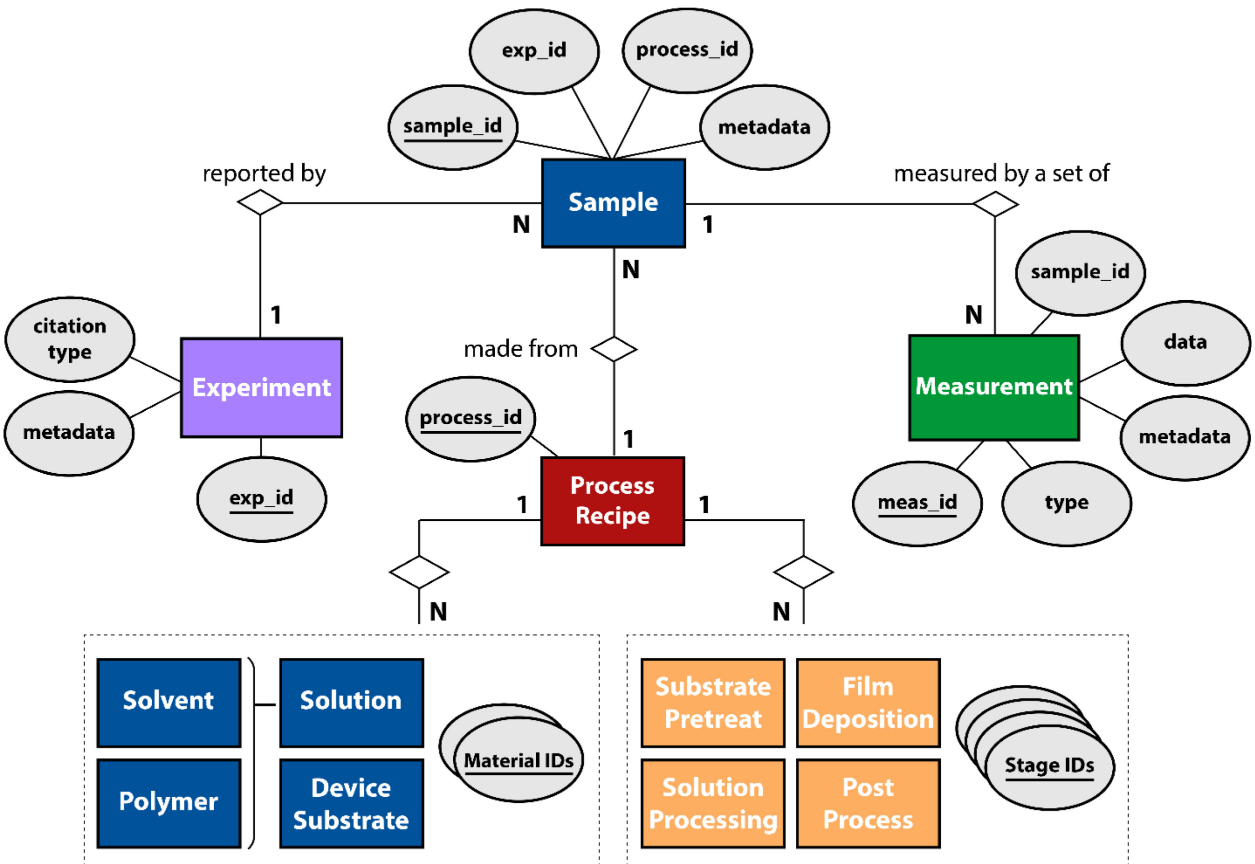

Figure 2. High level entity-relationship diagram depicting the reporting of an experimental Sample, including relationships with Experiment information, the Process recipe, and the associated Measurement data. Grey ovals are attributes associated with the labeled entities. Inputs to the process recipe are represented by material nodes and process stages (see Figure 3).

A **sample** refers to a single organic device in the form in which all its characterization data (a set of **measurement** objects) were collected and contains explicit information associating it with its reporting origin (**experiment**) and its physical origin/process history (**process**) (**Figure 2**). **Sample** has a one-to-many (1-N) relationship with **measurement**, since a single OFET could be measured a number of times by a variety of characterization methods not limited to device performance. A **measurement** has a type (*e.g.*, transfer curve, spectroscopy, scattering, thickness, *etc.*) and the heterogeneous data (*e.g.*, value type, value, error, *etc.*) and metadata (*e.g.*, instrument information, date measured, *etc.*) associated with it. An **experiment** refers to the source information associated with the reporting of the sample, and therefore has a citation type (*e.g.*, laboratory, journal article, dissertation) and metadata associated with that source (*e.g.*, digital object identifier (DOI), date published, author information, *etc.*). **Experiment** and **sample** have a one-to-many relationship, as a single **experiment** may report multiple samples. A **process recipe** refers to the unique material ingredients and process sequence through which the sample was generated. The process recipe contains foreign keys (*i.e.*, references to material nodes) that link information about the **device substrate** and the **solution** (the latter contains **polymer** and **solvent** information), and to process stage nodes that subdivide the process history (*vide infra*). Metadata fields linked to the material entities contain information such as polymer batch or lot information, supplier information, *etc*. These material nodes also serve as placeholders to expand the data model to include more details on synthetic routes (for **polymer**) and/or device fabrication routes (for **device substrate**) in future database development. **Sample** and **process recipe** have a many-to-one (N-1) relationship, as a given sample device can only be associated with one process recipe but the same process recipe could be used for multiple samples.

2.3. Process representation

Comparing device data reported from multiple sources requires that the various nuances of the experimental design space are accurately represented in a robust data format. Particularly, understanding the sensitivity of the process space is non-negotiable for the sake of reproducibility and accurate data representation. However, in contrast to the other entities in **Figure 2**, it is not straightforward to manifest a data structure that broadly represents the process history for the conjugated polymer layer. This is not only because the real-world process history is extremely complex, but also because the various events in a process history have an explicit order, and the events may not occur consistently across samples. For example, as discussed above, the solution processing procedure may include sequenced pre-treatment techniques to induce polymer aggregation in the solution state. This procedure may include multiple steps, and the ordering of those steps may affect the final film and properties (*i.e.*, sonication then aging, or aging then sonication).[44] The example above exemplifies a broader challenge in robustly handling information in both dynamic and nuanced ways in sample recipes.

One avenue to formulate a data structure is to sub-divide the sample generation process into a series of sub-processes that appear in a consistent ordering for any given polymer active layer in an OFET, model the relationships among a standard domain of entities within those sub-processes, and use the resulting graph to sort data. Recently, Walsh *et al.* proposed a generalized polymer data structure in CRIPT, introducing a data format for process entities that can be sequentially arranged to represent successive material transformations.[29] However, there is no established standard for defining the boundaries of a "process" for the sake of knowledge representation in materials data structures. We propose that incorporating a universal standard to help compose and arrange individual process stages would foster the adoption of generalized data models that can be used to model a broad set of application domains that are sensitive to complex processing histories.

Therefore, herein we adopt an international automation standard in ISA-88 to facilitate the conceptual modeling of the conjugated polymer deposition in a logical way (**Figure 3a**).[31] ISA-88 is a standard that is used in batch process control to organize the various pieces of data associated with a complex network of instrumentation and process stages, wherein a batch process input material is fed to a defined order of processing actions (*e.g.*, pieces of equipment) in series or in parallel to obtain some output material. Section 4.1 of ISA-88 defines a series of hierarchical subdivisions that are increasingly descriptive of an overall batch process. If the first level of the hierarchy is the overall process, the process is subdivided at the second level as an ordered set of process stages which operate independently from each other, usually in a planned sequence of physical changes in the material being processed. Process stages can be broken up into individual process operations, which are defined as major activities that result in a chemical or physical change in the material inputs. At the lowest level of the process model, process actions represent the minor activities that make up a process operation. Within each level of complexity, entities are a directed set of process sub-nodes organized in serial, parallel, or both.[31]

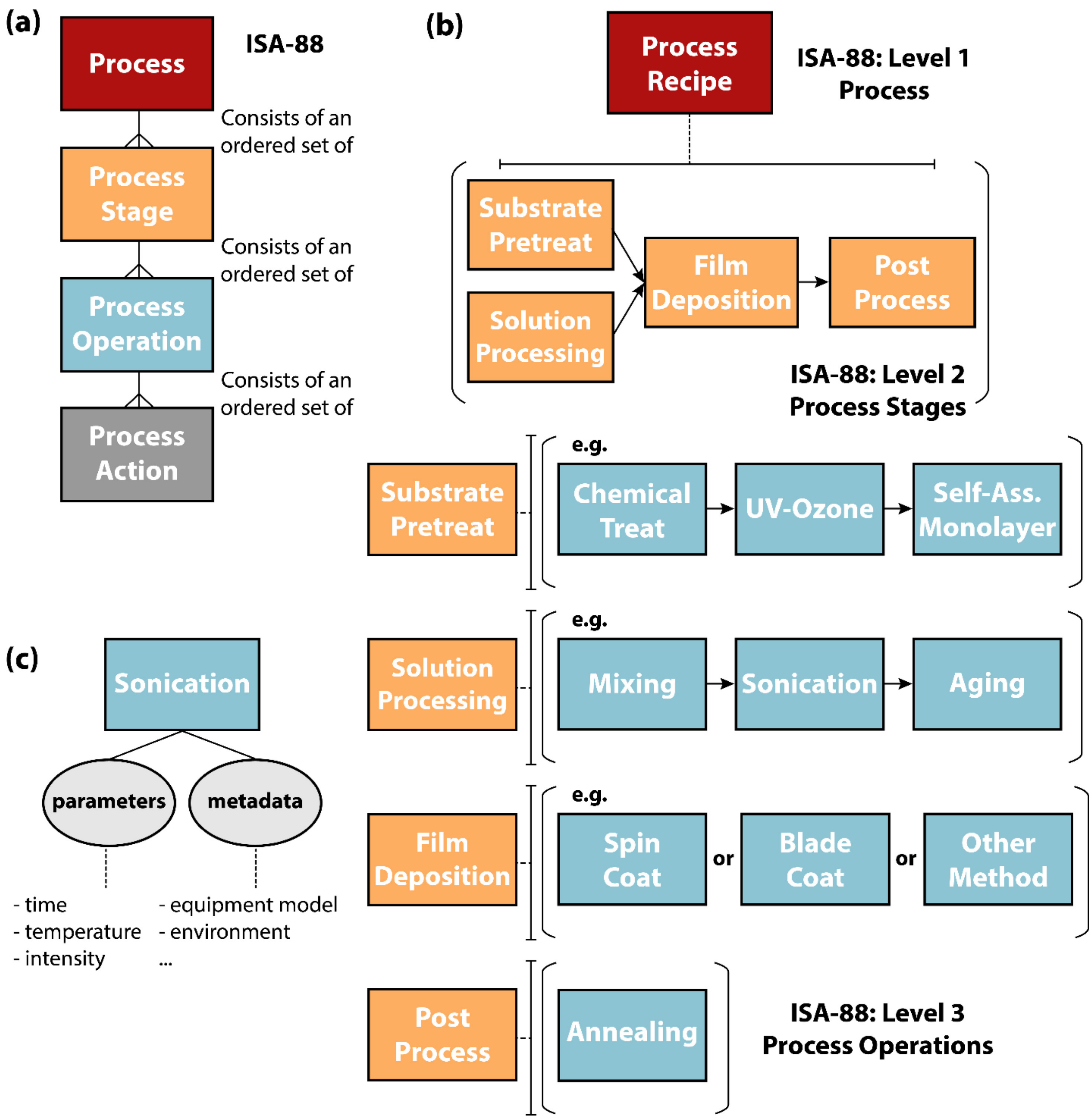

Figure 3. (a). Entity-relationship diagram of ISA-88, an international standard in batch process control. Material nodes and attributes are not explicitly included in the process model. Here, the former participate as inputs and outputs to the process, while attributes may describe a process node at the lowest level of complexity (See Figure 2). (b) Expanded graphical representation of an OFET fabrication recipe into a set of process stages, using the second level of ISA-88. (c). Process stage expansion to directed relationships between process operations at the third level of ISA-88. Ovals represent attributes at the model's lowest level.

The conjugated polymer process recipe is expanded to process stages of **substrate pretreatment, solution processing, film deposition,** and **post processing** at the second level of ISA-88 (**Figure 3b**). These entities are *fixed stages* and represent logical subdivisions that broadly describe the active layer transformation of the input materials. **Figure 3c** represents a third level of process description, showing that the sequence of process operations that falls within the boundaries of each polymer process stage may vary from sample to sample. **Substrate pretreatment** refers to all the processing activities related to preparing the patterned device substrate for the deposition process. This sequence may or may not include cleaning treatments (*e.g.*, chemical washing, UV-ozone, plasma treatment, *etc.*) proceeded by one or more surface modification steps, such as the use of a self-assembled monolayer. **Solution processing** includes the various ordered steps (*e.g.*, mixing, sonication, cooling, aging, *etc.*) that transform the polymer/solvent components into the final solution or ink formulation that ultimately gets coated onto the treated substrate during the **film deposition** stage. The film deposition stage includes only one node that contains information about the parameters and metadata of the coating method (*e.g.*, blade coating, spin coating, drop casting, inkjet printing, *etc.*). Finally, post-deposition operations such as further chemical treatment or annealing appear under the **post process** stage node.

Further expanding the process operations into individual process actions is possible but provides a level of complexity that may not be required for sufficient database design, as the main purpose of the process model is to identify the most appropriate object classification to store attributes. For example, a "sonication" node might be expanded to process actions such as "set sonication time" or "set sonication intensity", but this information can be just as easily represented by storing the attributes "sonication time" and "sonication intensity" in the parent process operation node (**Figure 3c**). However, it should be noted that this next level of process expansion may be useful for relating data commands or readings to computer-integrated or fully automated instrumentation. In either case, all the raw parameters, data, and metadata information are stored in the nodes at the lowest process level.

3. OFET-db: A database implementation and demonstration

The prior section proposes a general data structure and ontology for storing information related to the formation of the conjugated polymer active layer in a device such as an OFET. ISA-88's sequenced process model also allows for process history to be effectively captured in a data model, as the conjugated polymer is transformed into the final active layer of the measured device through a batch process. The process representation provides a high-level fixed structure (the main process stages of solution processing, substrate treatment, *etc.*) to promote aspects of a consistent relational schema, while providing flexibility for storing dynamic information within each of the stages. Using the data model described above, an experimental repository of OFET device measurements was curated from a set of published, peer-reviewed literature data (**Supporting Information**) and unpublished laboratory data. The following section discusses the initial construction and continued curation of experimental device records into the database and provides a brief demonstration of utilizing the database for meaningful searching and data visualization.

3.1. Data sourcing

The database was seeded using a set of experimental datasets for a set of three model polymer systems for electronic devices: P3HT, DPP-DTT, and poly([N,N'-bis(2-octyldodecyl)naphthalene-1,4,5,8-bis(dicarboximide)-2,6-diyl]-alt-5,5'-(2,2'-bithiophene)) (N2200). The P3HT, DPP-DTT, and N2200 data were curated from a body of literature combining over 50 peer-reviewed journal articles reporting OFET device performance, containing a heterogeneous set of information including process-related parameters and some structural characterization data. A subset of the database was also curated using unpublished records from laboratory experiments.

3.2. Database management system (DBMS)

The database was constructed using PostgreSQL, an open-source, relational database management system primarily based on the structured query language (SQL) that has strong support for NoSQL features. This allowed the database to have the preferred functionalities of the relational model (e.g., data normalization, data redundancy, error checking, etc.), while allowing for storage flexibility where attributes may be dynamic.[30] Preserving relationships is also a key factor in representing sample provenance in a robust manner, which makes certain aspects of the relational model attractive for the sake of interoperability with other community databases. At the same time, PostgreSQL can handle storage and queries on a variety of complex data types, including JSON, XML, and binary objects, which is not a feature that is always available for SQL databases. The mix of SQL and NoSQL features allows the implementation of a data model that provides more convenient and robust organization for structured aspects (process stages, *e.g.,* film deposition) and flexible storage for unstructured information or data fields that may evolve with research thrusts over time (process operations, *e.g.,* coating parameters, and metadata fields). A complete description of the DBMS table schema, based on the data model described earlier, is available in the **Supporting Information**.

3.3. Vocabulary for data curation

The diversity of categorical or text descriptors in the OFET parameter space requires the use of a consistent vocabulary of keywords to guide the naming of attributes and free-text fields.

Naming errors can in part be mitigated through built-in DBMS functionalities but may persist, especially in JSON formats, due to flexible key-value naming. At the time of writing, the implementation of OFET-db uses some keywords borrowed from other large-scale materials database efforts (*e.g.*, where applicable and available, citation/source keywords from MaterialsMine,[26] process and material keywords from CRIPT,[56] *etc.*), but largely uses custom keywords that provide more specificity to descriptors relevant to the organic device fabrication domain (*e.g.*, *blade*, *spin*, *inkjet*, *dip*, *etc*. to specify different classes of solution coating/deposition techniques). A full list of terms is available in the **Supporting Information** for the database implementation version described herein. Future design efforts will employ updated terminologies from shared community resources when available, as shared vocabularies promote consistent descriptions and interoperability. An experimental data entry template that incorporates this controlled vocabulary has also been adapted from a similar template shared by MaterialsMine.[27] This template is intended to not only reduce the time and inconvenience that is inevitable for an experimentalist or domain expert to fill out a data record for database entry but provides a tool to reduce the error checking and validation workload on the back end. Future template versions could be implemented as user-friendly webforms, a web application, or directly coupled to electronic laboratory notebooks or integrated laboratory instrumentation to facilitate the process of database inserts for newly curated experimental records.

3.4. Data visualization

Data reads from PostgreSQL are facilitated through built-in Python libraries, such as *psycopg2* and *pandas*. The following section demonstrates basic data analyses generated from such read queries to highlight the usefulness of enabling databases for experimental research purposes.

In the future, it is envisioned that a larger population of data can facilitate data-driven knowledge discovery activities through the utilization of data science or machine learning techniques.

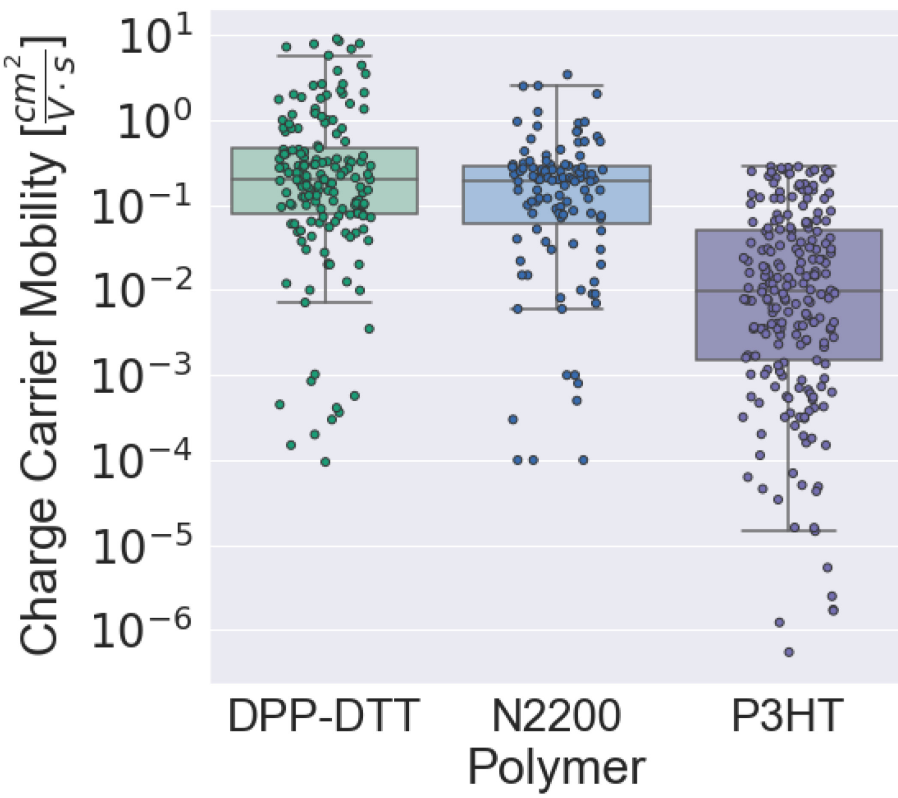

Figure 4. Single-axis scatterplots superimposed on boxplots, visualizing the mobility distributions for three representative polymers in OFET-db. The interquartile range (IQR) is defined between 25^{th} and 75^{th} percentiles, where the whisker endpoints are defined by 1.5*IQR. Data is shown only for pure-component active layers; blends are omitted.

Figure 4 shows the distribution of charge-carrier mobilities generated using data queried from OFET-db, showing the total spread and statistics of performance values for three relatively well-represented polymers in OFET-db: DPP-DTT, N2200, and P3HT. OFET samples fabricated from all individual polymer types show performance variations that span several orders of magnitude. The higher average and maximum charge-carrier mobility achieved for the DPP-DTT and N2200 data also reflects the general performance advantage of donor-acceptor copolymers versus the model homopolymer P3HT, even despite this large variation. Polymer material characteristics, such as molecular weight and polydispersity, are important factors in performance differences, as demonstrated by **Figure 5**. Molecular weight is a well-studied parameter for conjugated polymers in OFETs, and it is generally understood that longer conjugated backbones promote entanglements and aggregates in solution, and thereby enhance long-range molecular order and charge transport pathways in the thin film.[36, 41, 57] The positive correlation between molecular weight and mobility is generally visible for P3HT and DPP-DTT. Variation in mobility for constant molecular weight is visible for all three polymers when, for example, the same study uses the same polymer batch to explore the effect of a chosen processing motif on the device performance, highlighting the importance of including such process details. A similar positive trend in mobility is visible for the polydispersity index (PDI), where devices made from a higher PDI polymer are more likely to have mobilities in the upper range of the dataset.

Compared to molecular weight, however, the effect of PDI on the charge-carrier mobility is not as well-understood by the community. While this behavior may at first glance be due to a correlation between molecular weight and PDI, a Spearman Correlation analysis shows that these two variables are positively correlated only for P3HT ($r = 0.79$) but remain uncorrelated for DPP-DTT (0.03) and N2200 (0.23) (Table 1). PDI may be less frequently studied as a tunable experimental parameter since batch characteristics are sensitive to the synthetic procedure, hindering the design of controlled experiments for OFET performance comparisons. Recently, McBride *et al.* blended different M_w batches of P3HT and found that a wider molecular weight distribution exhibited beneficial effects due to a synergistic behavior between shorter tie chains connecting aggregated domains of larger chains in aged solutions.[58] However, the data analysis above shows that a relationship between dispersity and mobility may be a common trend for copolymer systems, which potentially motivates a broader study into the structural mechanisms behind the mobility dependence on molecular weight distributions.

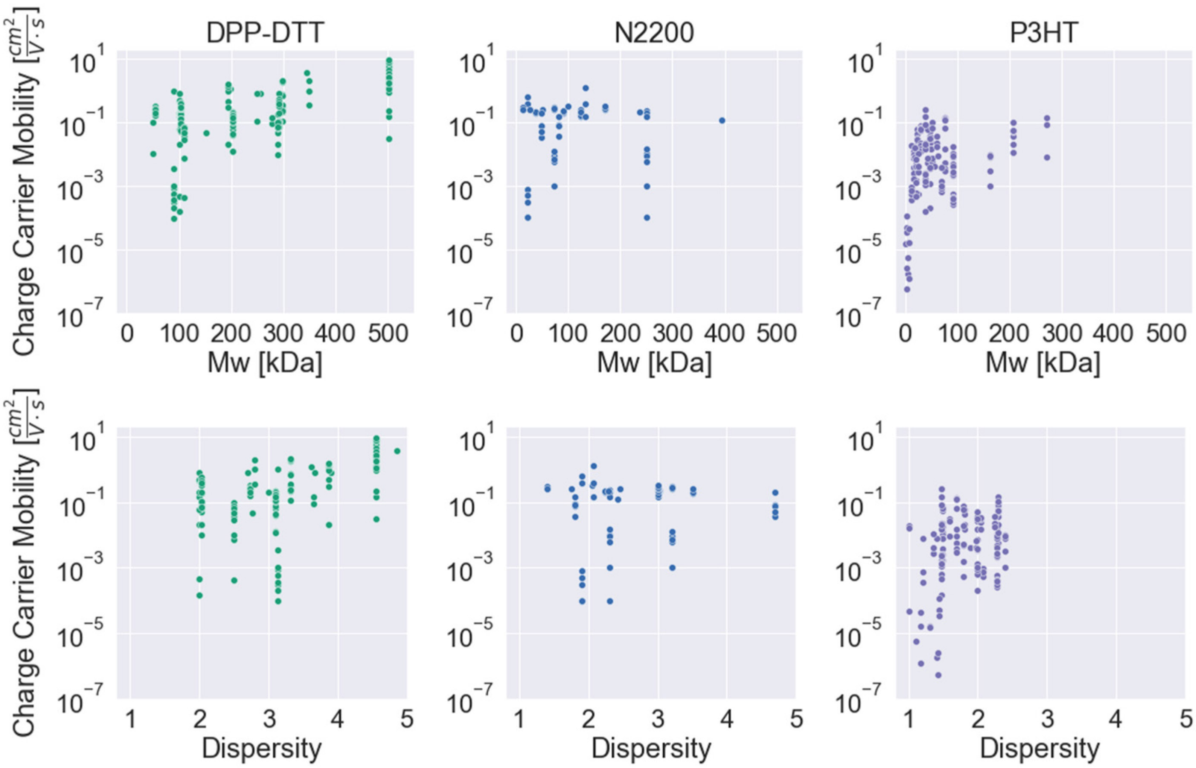

Figure 5. Charge-carrier mobility trends plotted against polymer material characteristics (a) molecular weight, (b) dispersity

Table 1. Spearman correlation coefficients calculated between molecular weight (M_w) and PDI for datasets classified by polymer

Polymer	**Spearman Correlation Coefficient: M_w vs PDI**
DPP-DTT	0.03
N2200	0.23
P3HT	0.79

The data model implemented herein also provides the flexibility to index information from structural measurements (*e.g.*, spectroscopic signals, microscopic images, *etc.*), which enhances the ability to interrogate process-structure-property relations. The population of structural measurements is much smaller than the number of device measurements, but here we show that structural information is available and representative.

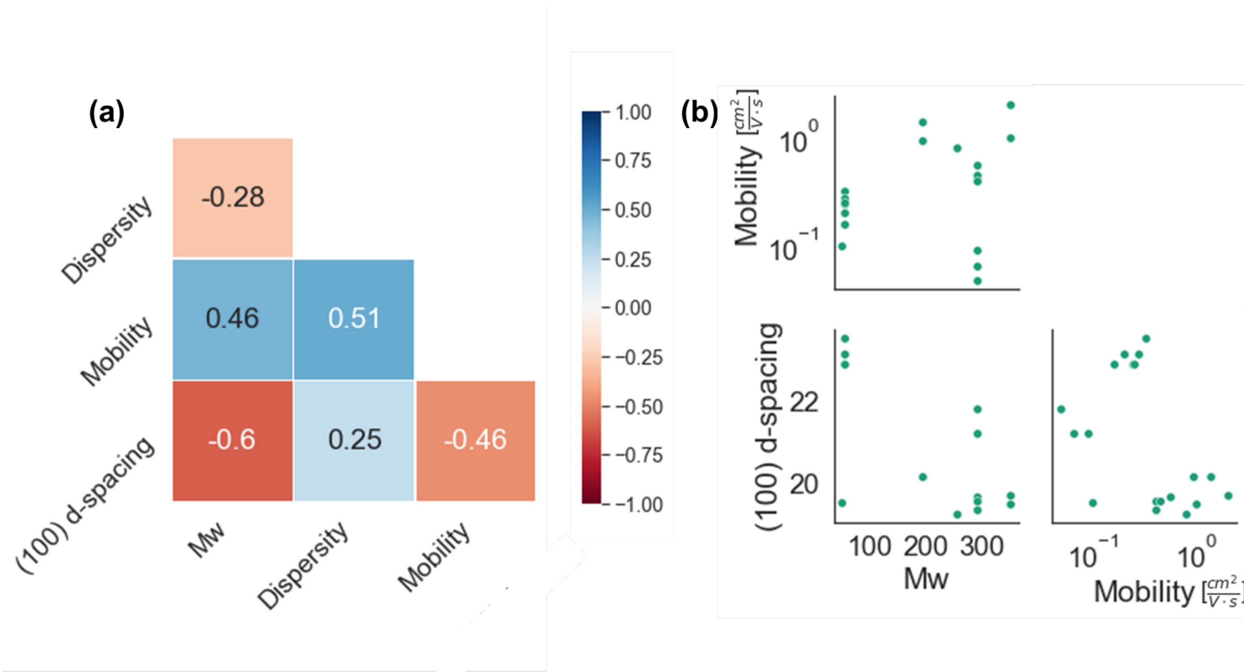

Figure 6. (a) Spearman correlation matrix on a subset of material information with thin film structural information (100) d-spacing) in GIWAXS. (b). Pair plots showing a 3-dimensional relationship between molecular weight, hole mobility and available (100) d-spacing from GIWAXS measurements of pure-component DPP-DTT devices.

For example, **Figure 6** analyzes the polymer characteristics of DPP-DTT with respect to (100) d-spacing extracted from available GIWAXS data. Most notably, Figure 6 shows a strong negative Spearman correlation value between molecular weight and (100) d-spacing, indicating that the lamellar spacing tends to decrease with an increasing conjugated backbone length for DPP-DTT. A corresponding negative correlation between d-spacing and mobility suggests that this change in lamellar spacing is a potential indicator for improved charge transport characteristics in the thin film, as decreased lamellar spacing could facilitate charge hopping.[59] Though this distance is merely one factor characterizing the crystalline domain, this observation draws interest in considering further parameters (*e.g.*, full width at half maximum, degree of crystallinity, *etc.*) to study the impact on device performance. As the data reported here is relatively sparse, with only 31 of the DPP-DTT device samples registering associated GIWAXS data, more meaningful structure-property observations could potentially be extracted when a richer set of structural data is recorded. Nonetheless, we demonstrate the utility of using our preliminary body of populated data to inform future work toward a richer experimental repository and greater physical understanding. Such observations can suggest new hypotheses that can then be tested with additional experiments.

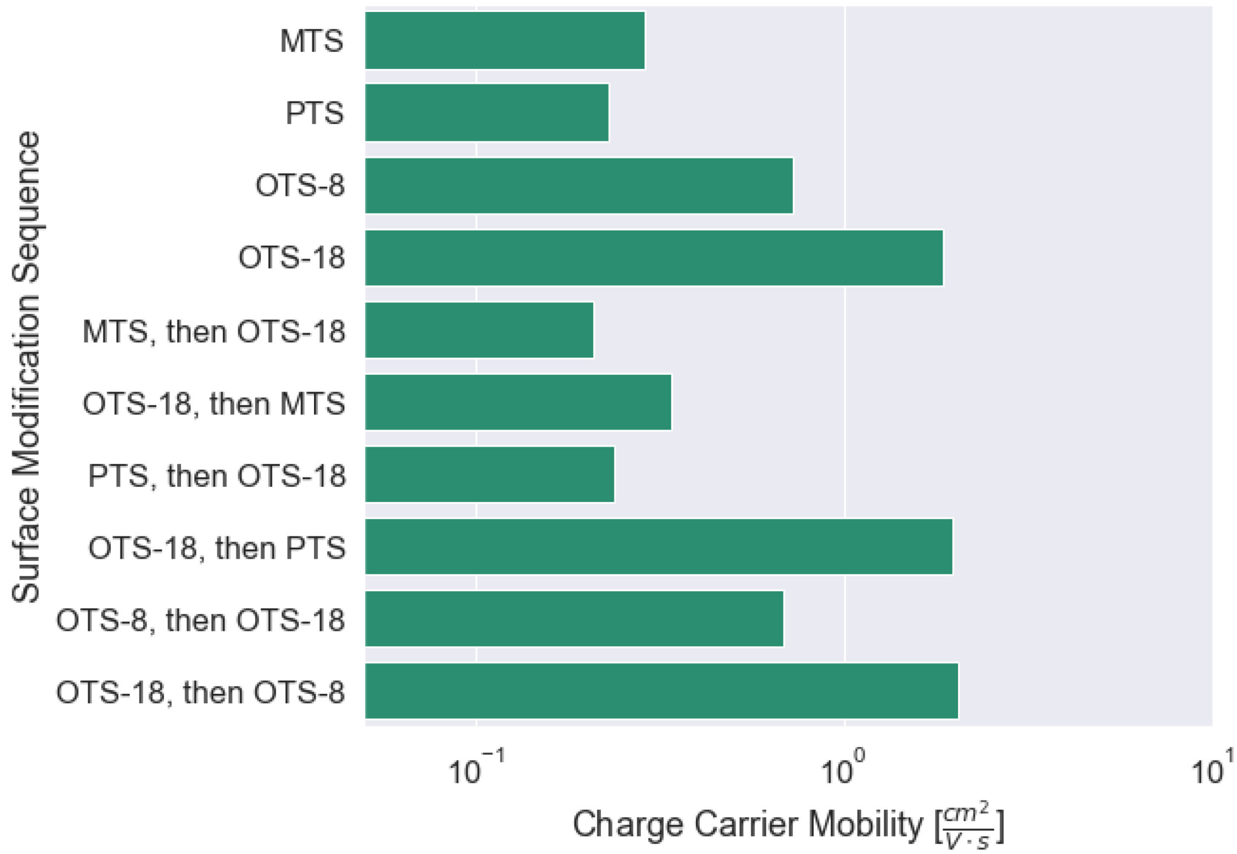

Figure 7. Impact of sequential substrate surface modification on charge carrier mobility for a DPP-DTT study. Two different surface modifiers, including methyltrichlorosilane (MTS), octyltrichlorosilane (OTS-8), octadecyltrichlorosilane (OTS-18), and phenyltrichlorosilane (PTS), are used sequentially to treat the substrate as detailed in ref [38].

We expect that future data-driven studies that use OFET-db as a resource will benefit from representative storage of process history, as materials characteristics alone often provide insufficient information for fully understanding the experimental sensitivity of device performance. To that end, with our proposed process ontology we aim to enable the curation of experimental data with the associated process history of samples, with the future intention of providing advanced analyses based on process recipes. A preliminary demonstration of the ability to curate, select, and plot process sequenced data is presented briefly herein. We used a query that searched for devices deposited on substrates treated by more than one surface modification agent

(**Supporting Information**) to show the subset of data extracted from the database in **Figure 7**. This subset highlights results from a single study[38] that compares the OFET mobility of devices deposited onto substrates sequentially pretreated with either of three pairs of silanes: methyltrichlorosilane (MTS) and octadecyltrichlorosilane (OTS-18), OTS-18 and phenyltrichlorosilane (PTS), and octyltrichlorosilane (OTS-8) and OTS-18. For all three of the pairings, noticeable differences in performance were observed depending on whether OTS-18 treatment was performed before or after the other silane agent.

This comparison highlights the necessity to capture process history (rather than simply process parameters) in fully describing device performance, as such processes will affect the relationship between solution-state and thin-film assembly, and therefore the final device properties. A major challenge, however, is that drawing device comparisons across different authors based on a standard recipe (*i.e.*, isolating a large set of devices made with a standard set of process conditions that appear frequently) would require more data since the process space is very large.[60] Currently, "low" availability of curated data precludes a more comprehensive meta-analysis of the process-structure-property relationships that govern the device performance of conjugated polymers. While the process-structure-property analyses discussed above for OFET-db demonstrate the challenge in driving data-driven studies in a "small data" environment, they also highlight the potential in mobilizing a database that can capture the various experimental nuances that could be indispensable toward greater physicochemical understanding of conjugated polymer-based organic devices. Therefore, the application of process ontologies is necessary for a broader adoption of representative materials databases, which the work herein addresses for conjugated polymer processing in OFETs.

4. CONCLUSIONS

Herein, we demonstrate the design and implementation of a data model for the experimental domain of OFETs as a foundation for broadly enabling database curation and management for organic thin-film electronics. Specifically, capturing process history in a manner that conforms to standard data protocols was a key challenge that was addressed by employing ISA-88, a batch process data model. Then, a database was constructed based on the model using PostgreSQL, enabling storage capabilities for both SQL (structured data) and NoSQL (document-based data) to provide flexibility without sacrificing the advantages of data redundancy/normalization in the relational model. While the work and discussion presented provides an experimental database that applies to the OFET active layer, it also serves as a model for adaptation to other aspects of organic device experiments by designing the data structure around an accepted process standard. Moving forward, enhancing materials ontologies to comprehensively capture classifications of process steps would facilitate the future design of data models in other domains that accurately manifest the real-world experimental processes. This is a necessary pursuit in elucidating the format in which a sample's provenance is recorded within a database in FAIR data structures. Additionally, future work will build upon the preliminary body of curated experimental data to mobilize data-driven experimentation for polymer-based organic electronics.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge online.

Detailed relational database table schema as implemented in PostgreSQL; vocabulary list for free-text attributes; list of publications from which initial seed data was extracted. Supporting code and a local database implementation for OFET-db is available on GitHub at https://github.com/aaronliu64/ofetdb_public.

AUTHOR INFORMATION

Corresponding Authors

Elsa Reichmanis – Department of Chemical & Biomolecular Engineering, Lehigh University, Bethlehem, PA 18015, USA

Martha A. Grover – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Other Authors

Aaron L. Liu – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Rahul Venkatesh – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Myeongyeon Lee – Department of Chemical & Biomolecular Engineering, Lehigh University, Bethlehem, PA 18015, USA

Jessica A. Bonsu – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Ron Volkovinsky – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Carson Meredith – School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332

ORCID:

Aaron L. Liu: 0000-0001-7347-5347
Elsa Reichmanis: 0000-0002-8205-8016
Martha Grover: 0000-0002-7036-776X
J. Carson Meredith: 0000-0003-2519-5003
Rahul Venkatesh: 0000-0003-1008-6568
Myeongyeon Lee: 0000-0002-3035-9909
Jessica A. Bonsu: 0009-0005-8233-7129

Notes

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Financial support from NSF DMREF (Grant No. 1922111, DMREF: Collaborative Research: Achieving Multicomponent Active Materials through Synergistic Combinatorial, Informatics-Enabled Materials Discovery) is gratefully acknowledged. Dr. Shamkant Navathe is gratefully acknowledged for advising on aspects of database design and implementation. Dr. Ying Diao and Azzaya Khasbaatar are acknowledged for helpful discussions on the organic electronics parameter space and piloting data entry. Dr. Chad Risko, Vinayak Bhat, Rebekah Duke, and Shubham More are also acknowledged for discussions on data management. Yulong Zheng is acknowledged for valuable discussions related to DPP-DTT structural characterization.

REFERENCES

(1) Root, S. E.; Savagatrup, S.; Printz, A. D.; Rodriquez, D.; Lipomi, D. J. Mechanical Properties of Organic Semiconductors for Stretchable, Highly Flexible, and Mechanically Robust Electronics. *Chem Rev* **2017**, *117* (9), 6467-6499. DOI: 10.1021/acs.chemrev.7b00003.
(2) Saeki, A.; Kranthiraja, K. A high throughput molecular screening for organic electronics via machine learning: present status and perspective. *Japanese Journal of Applied Physics* **2020**, *59* (SD), SD0801. DOI: 10.7567/1347-4065/ab4f39.
(3) Zhang, Q.; Hu, W.; Sirringhaus, H.; Müllen, K. Recent Progress in Emerging Organic Semiconductors. *Advanced Materials* **2022**, *34* (22), 2108701, https://doi.org/10.1002/adma.202108701. DOI: https://doi.org/10.1002/adma.202108701 (acccessed 2023/03/09).
(4) *Materials Genome Initiative*. https://www.mgi.gov/ (accessed September 25, 2023).
(5) Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. Machine learning for the structure–energy–property landscapes of molecular crystals. *Chemical Science* **2018**, *9* (5), 1289-1300, 10.1039/C7SC04665K. DOI: 10.1039/C7SC04665K.
(6) Yang, J.; De, S.; Campbell, J. E.; Li, S.; Ceriotti, M.; Day, G. M. Large-Scale Computational Screening of Molecular Organic Semiconductors Using Crystal Structure Prediction. *Chemistry of Materials* **2018**, *30* (13), 4361-4371. DOI: 10.1021/acs.chemmater.8b01621.
(7) McBride, M.; Persson, N.; Reichmanis, E.; Grover, M. Solving Materials' Small Data Problem with Dynamic Experimental Databases. *Processes* **2018**, *6* (7), 1-17. DOI: 10.3390/pr6070079.
(8) Venkatesh, R.; Zheng, Y.; Viersen, C.; Liu, A.; Silva, C.; Grover, M.; Reichmanis, E. Data Science Guided Experiments Identify Conjugated Polymer Solution Concentration as a Key Parameter in Device Performance. *ACS Materials Letters* **2021**, *3* (9), 1321-1327. DOI: 10.1021/acsmaterialslett.1c00320.
(9) Zhuo, Y.; Brgoch, J. Opportunities for Next-Generation Luminescent Materials through Artificial Intelligence. *The Journal of Physical Chemistry Letters* **2021**, *12* (2), 764-772. DOI: 10.1021/acs.jpclett.0c03203.
(10) Gómez-Bombarelli, R.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Duvenaud, D.; Maclaurin, D.; Blood-Forsythe, M. A.; Chae, H. S.; Einzinger, M.; Ha, D.-G.; Wu, T.; et al. Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. *Nature Materials* **2016**, *15* (10), 1120-1127. DOI: 10.1038/nmat4717.
(11) Janai, M. A. B.; Woon, K. L.; Chan, C. S. Design of efficient blue phosphorescent bottom emitting light emitting diodes by machine learning approach. *Organic Electronics* **2018**, *63*, 257-266. DOI: https://doi.org/10.1016/j.orgel.2018.09.029.
(12) Shu, Y.; Levine, B. G. Simulated evolution of fluorophores for light emitting diodes. *The Journal of Chemical Physics* **2015**, *142* (10), 104104. DOI: 10.1063/1.4914294 (acccessed 2023/03/07).
(13) Langner, S.; Häse, F.; Perea, J. D.; Stubhan, T.; Hauch, J.; Roch, L. M.; Heumueller, T.; Aspuru-Guzik, A.; Brabec, C. J. Beyond Ternary OPV: High-Throughput Experimentation and Self-Driving Laboratories Optimize Multicomponent Systems. *Advanced Materials* **2020**, *32* (14), 1907801, https://doi.org/10.1002/adma.201907801. DOI: https://doi.org/10.1002/adma.201907801 (acccessed 2021/05/31).
(14) Rodríguez-Martínez, X.; Pascual-San-José, E.; Campoy-Quiles, M. Accelerating organic solar cell material's discovery: high-throughput screening and big data. *Energy & Environmental Science* **2021**, *14* (6), 3301-3322, 10.1039/D1EE00559F. DOI: 10.1039/D1EE00559F.
(15) An, N. G.; Kim, J. Y.; Vak, D. Machine learning-assisted development of organic photovoltaics via high-throughput in situ formulation. *Energy & Environmental Science* **2021**, *14* (6), 3438-3446, 10.1039/D1EE00641J. DOI: 10.1039/D1EE00641J.
(16) Abbasi Jannat Abadi, E.; Sahu, H.; Javadpour, S. M.; Goharimanesh, M. Interpretable machine learning for developing high-performance organic solar cells. *Materials Today Energy* **2022**, *25*, 100969. DOI: https://doi.org/10.1016/j.mtener.2022.100969.
(17) Sahu, H.; Rao, W.; Troisi, A.; Ma, H. Toward Predicting Efficiency of Organic Solar Cells via Machine Learning and Improved Descriptors. *Advanced Energy Materials* **2018**, *8* (24), 1801032, https://doi.org/10.1002/aenm.201801032. DOI: https://doi.org/10.1002/aenm.201801032 (acccessed 2023/03/07).
(18) Du, X.; Lüer, L.; Heumueller, T.; Wagner, J.; Berger, C.; Osterrieder, T.; Wortmann, J.; Langner, S.; Vongsaysy, U.; Bertrand, M.; et al. Elucidating the Full Potential of OPV Materials Utilizing a High-Throughput Robot-Based Platform and Machine Learning. *Joule* **2021**, *5* (2), 495-506. DOI: 10.1016/j.joule.2020.12.013 (acccessed 2021/06/29).

(19) Hachmann, J.; Olivares-Amaya, R.; Atahan-Evrenk, S.; Amador-Bedolla, C.; Sánchez-Carrera, R. S.; Gold-Parker, A.; Vogt, L.; Brockway, A. M.; Aspuru-Guzik, A. The Harvard Clean Energy Project: Large-Scale Computational Screening and Design of Organic Photovoltaics on the World Community Grid. *The Journal of Physical Chemistry Letters* **2011**, *2* (17), 2241-2251. DOI: 10.1021/jz200866s.
(20) Ai, Q.; Bhat, V.; Ryno, S. M.; Jarolimek, K.; Sornberger, P.; Smith, A.; Haley, M. M.; Anthony, J. E.; Risko, C. OCELOT: An infrastructure for data-driven research to discover and design crystalline organic semiconductors. *The Journal of Chemical Physics* **2021**, *154* (17), 174705. DOI: 10.1063/5.0048714 (acccessed 2023/03/09).
(21) Joung, J. F.; Han, M.; Jeong, M.; Park, S. Experimental database of optical properties of organic compounds. *Scientific Data* **2020**, *7* (1), 295. DOI: 10.1038/s41597-020-00634-8.
(22) Holliday, S.; Donaghey, J. E.; McCulloch, I. Advances in Charge Carrier Mobilities of Semiconducting Polymers Used in Organic Transistors. *Chemistry of Materials* **2013**, *26* (1), 647-663. DOI: 10.1021/cm402421p.
(23) Khasbaatar, A.; Xu, Z.; Lee, J. H.; Campillo-Alvarado, G.; Hwang, C.; Onusaitis, B. N.; Diao, Y. From Solution to Thin Film: Molecular Assembly of pi-Conjugated Systems and Impact on (Opto)electronic Properties. *Chem Rev* **2023**. DOI: 10.1021/acs.chemrev.2c00905 From NLM Publisher.
(24) Deng, J.; Zheng, L.; Ding, C.; Guo, Y.; Xie, Y.; Wang, J.; Ke, Y.; Li, M.; Li, L.; Janssen, R. A. J. Determinant Role of Solution-State Supramolecular Assembly in Molecular Orientation of Conjugated Polymer Films. *Advanced Functional Materials* **2023**, *33* (1), 2209195, https://doi.org/10.1002/adfm.202209195. DOI: https://doi.org/10.1002/adfm.202209195 (acccessed 2023/07/05).
(25) Zhao, S.; Qian, Q. Ontology based heterogeneous materials database integration and semantic query. *AIP Advances* **2017**, *7* (10). DOI: 10.1063/1.4999209.
(26) Zhao, H.; Wang, Y.; Lin, A.; Hu, B.; Yan, R.; McCusker, J.; Chen, W.; McGuinness, D. L.; Schadler, L.; Brinson, L. C. NanoMine schema: An extensible data representation for polymer nanocomposites. *APL Materials* **2018**, *6* (11), 111108. DOI: 10.1063/1.5046839 (acccessed 2023/03/09).
(27) 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 Materials* **2016**, *4* (5), 053204. DOI: 10.1063/1.4943679 (acccessed 2023/02/09).
(28) Brinson, L. C.; Deagen, M.; Chen, W.; McCusker, J.; McGuinness, D. L.; Schadler, L. S.; Palmeri, M.; Ghumman, U.; Lin, A.; Hu, B. Polymer Nanocomposite Data: Curation, Frameworks, Access, and Potential for Discovery and Design. *ACS Macro Letters* **2020**, *9* (8), 1086-1094. DOI: 10.1021/acsmacrolett.0c00264.
(29) Walsh, D. J.; Zou, W.; Schneider, L.; Mello, R.; Deagen, M. E.; Mysona, J.; Lin, T.-S.; de Pablo, J. J.; Jensen, K. F.; Audus, D. J.; Olsen, B. D. Community Resource for Innovation in Polymer Technology (CRIPT): A Scalable Polymer Material Data Structure. *ACS Central Science* **2023**. DOI: 10.1021/acscentsci.3c00011.
(30) Duke, R.; Bhat, V.; Risko, C. Data storage architectures to accelerate chemical discovery: data accessibility for individual laboratories and the community. *Chemical Science* **2022**, *13* (46), 13646-13656, 10.1039/D2SC05142G. DOI: 10.1039/D2SC05142G.
(31) *International Society of Automation (ISA) 88 : Batch Control* https://www.isa.org/standards-and-publications/isa-standards/isa-standards-committees/isa88 (accessed March 9th 2023).
(32) Choi, D.; Chu, P.-H.; McBride, M.; Reichmanis, E. Best Practices for Reporting Organic Field Effect Transistor Device Performance. *Chemistry of Materials* **2015**, *27* (12), 4167-4168. DOI: 10.1021/acs.chemmater.5b01982.
(33) Paterson, A. F.; Singh, S.; Fallon, K. J.; Hodsden, T.; Han, Y.; Schroeder, B. C.; Bronstein, H.; Heeney, M.; McCulloch, I.; Anthopoulos, T. D. Recent Progress in High-Mobility Organic Transistors: A Reality Check. *Adv Mater* **2018**, e1801079. DOI: 10.1002/adma.201801079 From NLM Publisher.
(34) Gamota, D.; Brazis, P. IEEE standard test methods for the characterization of organic transistors and materials. *IEEE Std* **2004**, 1620-2008.
(35) Callaway, C. P.; Liu, A. L.; Venkatesh, R.; Zheng, Y.; Lee, M.; Meredith, J. C.; Grover, M.; Risko, C.; Reichmanis, E. The Solution is the Solution: Data-Driven Elucidation of Solution-to-Device Feature Transfer for π-Conjugated Polymer Semiconductors. *ACS Applied Materials & Interfaces* **2022**, *14* (3), 3613-3620. DOI: 10.1021/acsami.1c20994.
(36) Noriega, R.; Rivnay, J.; Vandewal, K.; Koch, F. P.; Stingelin, N.; Smith, P.; Toney, M. F.; Salleo, A. A general relationship between disorder, aggregation and charge transport in conjugated polymers. *Nat Mater* **2013**, *12* (11), 1038-1044. DOI: 10.1038/nmat3722.
(37) Kim, Y.-J.; Lee, S.; Niazi, M. R.; Hwang, K.; Tang, M.-C.; Lim, D.-H.; Kang, J.-S.; Smilgies, D.-M.; Amassian, A.; Kim, D.-Y. Systematic Study on the Morphological Development of Blade-Coated Conjugated Polymer Thin Films via In Situ Measurements. *ACS Applied Materials & Interfaces* **2020**, *12* (32), 36417-36427. DOI: 10.1021/acsami.0c07385.

(38) Lei, Y.; Wu, B.; Chan, W.-K. E.; Zhu, F.; Ong, B. S. Engineering gate dielectric surface properties for enhanced polymer field-effect transistor performance. *Journal of Materials Chemistry C* **2015**, *3* (47), 12267-12272. DOI: 10.1039/c5tc02579f.
(39) Persson, N. E.; Chu, P. H.; McBride, M.; Grover, M.; Reichmanis, E. Nucleation, Growth, and Alignment of Poly(3-hexylthiophene) Nanofibers for High-Performance OFETs. *Acc Chem Res* **2017**, *50* (4), 932-942. DOI: 10.1021/acs.accounts.6b00639 From NLM PubMed-not-MEDLINE.
(40) Xu, Z.; Park, K. S.; Kwok, J. J.; Lin, O.; Patel, B. B.; Kafle, P.; Davies, D. W.; Chen, Q.; Diao, Y. Not All Aggregates Are Made the Same: Distinct Structures of Solution Aggregates Drastically Modulate Assembly Pathways, Morphology, and Electronic Properties of Conjugated Polymers. *Advanced Materials* **2022**, *34* (32), 2203055, https://doi.org/10.1002/adma.202203055. DOI: https://doi.org/10.1002/adma.202203055 (acccessed 2023/03/10).
(41) Hu, H.; Zhao, K.; Fernandes, N.; Boufflet, P.; Bannock, J. H.; Yu, L.; de Mello, J. C.; Stingelin, N.; Heeney, M.; Giannelis, E. P.; Amassian, A. Entanglements in marginal solutions: a means of tuning pre-aggregation of conjugated polymers with positive implications for charge transport. *Journal of Materials Chemistry C* **2015**, *3* (28), 7394-7404. DOI: 10.1039/c5tc01425e.
(42) Zhao, K.; Khan, H. U.; Li, R.; Su, Y.; Amassian, A. Entanglement of Conjugated Polymer Chains Influences Molecular Self-Assembly and Carrier Transport. *Advanced Functional Materials* **2013**, *23* (48), 6024-6035. DOI: 10.1002/adfm.201301007.
(43) Boudinet, D.; Benwadih, M.; Altazin, S.; Verilhac, J.-M.; De Vito, E.; Serbutoviez, C.; Horowitz, G.; Facchetti, A. Influence of Substrate Surface Chemistry on the Performance of Top-Gate Organic Thin-Film Transistors. *Journal of the American Chemical Society* **2011**, *133* (26), 9968-9971. DOI: 10.1021/ja2010576.
(44) Kleinhenz, N.; Persson, N.; Xue, Z.; Chu, P. H.; Wang, G.; Yuan, Z.; McBride, M. A.; Choi, D.; Grover, M. A.; Reichmanis, E. Ordering of Poly(3-hexylthiophene) in Solutions and Films: Effects of Fiber Length and Grain Boundaries on Anisotropy and Mobility. *Chemistry of Materials* **2016**, *28* (11), 3905-3913. DOI: 10.1021/acs.chemmater.6b01163.
(45) McBride, M.; Bacardi, G.; Morales, C.; Risteen, B.; Keane, D.; Reichmanis, E.; Grover, M. A. Control of Nucleation Density in Conjugated Polymers via Seed Nucleation. *ACS Appl Mater Interfaces* **2019**, *11* (41), 37955-37965. DOI: 10.1021/acsami.9b10967 From NLM PubMed-not-MEDLINE.
(46) Park, K. S.; Kwok, J. J.; Dilmurat, R.; Qu, G.; Kafle, P.; Luo, X.; Jung, S.-H.; Olivier, Y.; Lee, J.-K.; Mei, J.; et al. Tuning conformation, assembly, and charge transport properties of conjugated polymers by printing flow. *Science Advances* **2019**, *5* (8), eaaw7757. DOI: 10.1126/sciadv.aaw7757 (acccessed 2023/03/09).
(47) Chang, M.; Choi, D.; Egap, E. Macroscopic Alignment of One-Dimensional Conjugated Polymer Nanocrystallites for High-Mobility Organic Field-Effect Transistors. *ACS Appl Mater Interfaces* **2016**, *8* (21), 13484-13491. DOI: 10.1021/acsami.6b02216 From NLM PubMed-not-MEDLINE.
(48) Gu, X.; Shaw, L.; Gu, K.; Toney, M. F.; Bao, Z. The meniscus-guided deposition of semiconducting polymers. *Nature Communications* **2018**, *9* (1), 534. DOI: 10.1038/s41467-018-02833-9.
(49) Afzal, T.; Iqbal, M. J.; Iqbal, M. Z.; Sajjad, A.; Raza, M. A.; Riaz, S.; Kamran, M. A.; Numan, A.; Naseem, S. Effect of post-deposition annealing temperature on the charge carrier mobility and morphology of DPPDTT based organic field effect transistors. *Chemical Physics Letters* **2020**, *750*, 137507. DOI: https://doi.org/10.1016/j.cplett.2020.137507.
(50) Tran, V. V.; Jeong, G.; Wi, E.; Lee, D.; Chang, M. Design and Fabrication of Ultrathin Nanoporous Donor–Acceptor Copolymer-Based Organic Field-Effect Transistors for Enhanced VOC Sensing Performance. *ACS Applied Materials & Interfaces* **2023**, *15* (17), 21270-21283. DOI: 10.1021/acsami.3c00105.
(51) Li, D.; Borkent, E.-J.; Nortrup, R.; Moon, H.; Katz, H.; Bao, Z. Humidity effect on electrical performance of organic thin-film transistors. *Applied Physics Letters* **2005**, *86* (4). DOI: 10.1063/1.1852708 (acccessed 9/16/2023).
(52) Blakesley, J. C.; Castro, F. A.; Kylberg, W.; Dibb, G. F. A.; Arantes, C.; Valaski, R.; Cremona, M.; Kim, J. S.; Kim, J.-S. Towards reliable charge-mobility benchmark measurements for organic semiconductors. *Organic Electronics* **2014**, *15* (6), 1263-1272. DOI: https://doi.org/10.1016/j.orgel.2014.02.008.
(53) McCulloch, I.; Salleo, A.; Chabinyc, M. Avoid the kinks when measuring mobility. *Science* **2016**, *352* (6293), 1521-1522. DOI: 10.1126/science.aaf9062 From NLM PubMed-not-MEDLINE.
(54) Horowitz, G. Organic Field-Effect Transistors. *Advanced Materials* **1998**, *10* (5), 365-377, https://doi.org/10.1002/(SICI)1521-4095(199803)10:5<365::AID-ADMA365>3.0.CO;2-U. DOI: https://doi.org/10.1002/(SICI)1521-4095(199803)10:5<365::AID-ADMA365>3.0.CO;2-U (acccessed 2023/03/10).
(55) Elmasri, R.; Navathe, S. B. *Fundamentals of Database Systems*; Pearson, 2015.
(56) *CRIPT Controlled Vocabulary*. 2023. https://criptapp.org/keys/ (accessed 03/22/2023).
(57) Koch, F. P. V.; Rivnay, J.; Foster, S.; Müller, C.; Downing, J. M.; Buchaca-Domingo, E.; Westacott, P.; Yu, L.; Yuan, M.; Baklar, M.; et al. The impact of molecular weight on microstructure and charge transport in semicrystalline

polymer semiconductors–poly(3-hexylthiophene), a model study. *Progress in Polymer Science* **2013**, *38* (12), 1978-1989. DOI: https://doi.org/10.1016/j.progpolymsci.2013.07.009.
(58) McBride, M.; Persson, N.; Keane, D.; Bacardi, G.; Reichmanis, E.; Grover, M. A. A Polymer Blend Approach for Creation of Effective Conjugated Polymer Charge Transport Pathways. *ACS Applied Materials & Interfaces* **2018**, *10* (42), 36464-36474. DOI: 10.1021/acsami.8b13255.
(59) Liu, R.; Yang, W.; Xu, W.; Deng, J.; Ding, C.; Guo, Y.; Zheng, L.; Sun, J.; Li, M. Impact of Chemical Design on the Molecular Orientation of Conjugated Donor–Acceptor Polymers for Field-Effect Transistors. *ACS Applied Polymer Materials* **2022**, *4* (4), 2233-2250. DOI: 10.1021/acsapm.1c01676.
(60) Persson, N.; McBride, M.; Grover, M.; Reichmanis, E. Silicon Valley meets the ivory tower: Searchable data repositories for experimental nanomaterials research. *Current Opinion in Solid State and Materials Science* **2016**, *20* (6), 338-343. DOI: 10.1016/j.cossms.2016.06.002.

For Table of Contents Only: