HTE OS: A High-Throughput Experimentation Workflow built from the Ground up

Georg Wuitschik*[,1], Vera Jost[1], Torsten Schindler[1], Michal Jakubik[2]

*georg.wuitschik@roche.com
[1] Roche Pharma Research and Early Development, Roche Innovation Center Basel
[2] MIA Solutions, Jaskový rad 209, 83101 Bratislava

Abstract

HTE OS is a free, open-source High-Throughput Experimentation workflow that supports practitioners from experiment submission all the way to results presentation. A core Google Sheet is responsible for reaction planning and execution as well communication with users and robots. All generated data is funneled into Spotfire where users can analyze it. Tools for parsing of LCMS-data and translation of chemical identifiers provide data wrangling capabilities to complete the workflow.

Introduction

High-Throughput Experimentation (HTE) is the continuation of parallel experiments with modern means. HTE's continuing rise in popularity is fueled by an ever increasing demand for complex molecules. Many of them are prepared using reactions where the outcome is strongly dependent on numerous parameters. Small variations in these often lead to significant yield changes, and the tools to predict the reaction outcome are still in their infancy.[1] Thus arises the need to reliably and efficiently run many experiments in parallel on a small scale.

[1] (a) Rinehart, N. I.; Saunthwal, R. K.; Wellauer, J.; Zahrt, A. F.; Schlemper, L.; Shved, A. S.; Bigler, R.; Fantasia, S.; Denmark, S. E. Development and Validation of a Chemoinformatic Workflow for Predicting Reaction Yield for Pd-Catalyzed C-N Couplings with Substrate Generalizability. 2023. https://doi.org/10.26434/chemrxiv-2022-hspwv-v2. (b) Fitzner, M.; Wuitschik, G.; Koller, R.; Adam, J.-M.; Schindler, T. Machine Learning C–N Couplings: Obstacles for a General-Purpose Reaction Yield Prediction. *ACS Omega* **2023**, *8* (3), 3017–3025. https://doi.org/10.1021/acsomega.2c05546. (c) Schwaller, P.; Vaucher, A. C.; Laino, T.; Reymond, J.-L. Prediction of Chemical Reaction Yields Using Deep Learning. *Mach. Learn.: Sci. Technol.* **2021**, *2* (1), 015016. https://doi.org/10.1088/2632-2153/abc81d. (d) Samha, M. H.; Karas, L. J.; Vogt, D. B.; Odogwu, E. C.; Elward, J.; Crawford, J. M.; Steves, J. E.; Sigman, M. S. Predicting Success in Cu-Catalyzed C–N Coupling Reactions Using Data Science. *Sci. Adv.* **2024**, *10* (3), https://doi.org/10.1126/sciadv.adn3478.

HTE is also growing in popularity because the means to perform it have improved.[2] This includes robust ways of dosing solids into reaction vials[3], wider adoption of glove boxes for more reproducible reaction execution, LCMS for reaction analysis[4] and an ever increasing universe of commercially available (pre-)catalysts, reagents and building-blocks. However, powerful yet user friendly open-source tools to integrate all the data and devices into a workflow are lacking.[5] Hence, we decided to build our own system orchestration and data analysis software to support our approach to HTE. We call it HTE OS and make it available for free as part of the Supporting Information.[6]

---

[2] For some examples of successful application in process development, see: (a) Selekman, J. A.; Qiu, J.; Tran, K.; Stevens, J.; Rosso, V.; Simmons, E.; Xiao, Y.; Janey, J. High-Throughput Automation in Chemical Process Development. *Annu Rev Chem Biomol* **2016**, *8* (1), 1–23. https://doi.org/10.1146/annurev-chembioeng-060816-101411. (b) Qiu, J.; Stevens, J. M. High-Throughput Classical Chiral Resolution Screening of Synthetic Intermediates: Effects of Resolving Agents, Crystallization Solvents, and Other Factors. *Org Process Res Dev* **2020**, *24* (9), 1725–1734. https://doi.org/10.1021/acs.oprd.0c00348. (c) Stevens, J. M.; Simmons, E. M.; Tan, Y.; Borovika, A.; Fan, J.; Forest, R. V.; Geng, P.; Guerrero, C. A.; Lou, S.; Skliar, D.; Steinhardt, S. E.; Strotman, N. A. Leveraging High-Throughput Experimentation to Drive Pharmaceutical Route Invention: A Four-Step Commercial Synthesis of Branebrutinib (BMS-986195). *Org Process Res Dev* **2022**, *26* (4), 1174–1183. https://doi.org/10.1021/acs.oprd.1c00443.

[3] Bahr, M. N.; Morris, M. A.; Tu, N. P.; Nandkeolyar, A. Recent Advances in High-Throughput Automated Powder Dispensing Platforms for Pharmaceutical Applications. Org Process Res Dev 2020, 24 (11), 2752–2761. https://doi.org/10.1021/acs.oprd.0c00411.

[4] Haas, C. P.; Lübbesmeyer, M.; Jin, E. H.; McDonald, M. A.; Koscher, B. A.; Guimond, N.; Rocco, L. D.; Kayser, H.; Leweke, S.; Niedenführ, S.; Nicholls, R.; Greeves, E.; Barber, D. M.; Hillenbrand, J.; Volpin, G.; Jensen, K. F. Open-Source Chromatographic Data Analysis for Reaction Optimization and Screening. *ACS Cent. Sci.* **2023**, *9* (2), 307–317. https://doi.org/10.1021/acscentsci.2c01042 .

[5] A number of publications exist that deal with the practical execution of HTE: (a) Shevlin, M. Practical High-Throughput Experimentation for Chemists. *Acs Med Chem Lett* **2017**, *8* (6), 601–607. https://doi.org/10.1021/acsmedchemlett.7b00165. (b) Cook, A.; Clément, R.; Newman, S. G. Reaction Screening in Multiwell Plates: High-Throughput Optimization of a Buchwald–Hartwig Amination. *Nat Protoc* **2021**, *16* (2), 1152–1169. https://doi.org/10.1038/s41596-020-00452-7. (c) Impastato, A. C.; Brown, J. T. C.; Wang, Y.; Tu, N. P. Readily Accessible High-Throughput Experimentation: A General Protocol for the Preparation of ChemBeads and EnzyBeads. *ACS Med. Chem. Lett.* **2023**, *14* (4), 514–520. https://doi.org/10.1021/acsmedchemlett.2c00491. (d) Fordham, J. M.; Kollmus, P.; Cavegn, M.; Schneider, R.; Santagostino, M. A "Pool and Split" Approach to the Optimization of Challenging Pd-Catalyzed C–N Cross-Coupling Reactions. J Org Chem **2022**, *87* (6), 4400–4414. https://doi.org/10.1021/acs.joc.2c00104. (e) Caldentey, X.; Romero, E. High-Throughput Experimentation as an Accessible Technology for Academic Organic Chemists in Europe and Beyond**. Chem Methods 2023. https://doi.org/10.1002/cmtd.**2022**00059. (f) Biyani, S. A.; Moriuchi, Y. W.; Thompson, D. H. Advancement in Organic Synthesis Through High Throughput Experimentation. *Chem Methods* **2021**, 1 (7), 323–339. https://doi.org/10.1002/cmtd.202100023.

[6] A live-demonstration of HTE OS can be found here: https://youtu.be/7QsfGkBi_xo?si=LDD9Ivr0QbGhjQM2&t=1010

Lab Workflow Integration through HTE OS

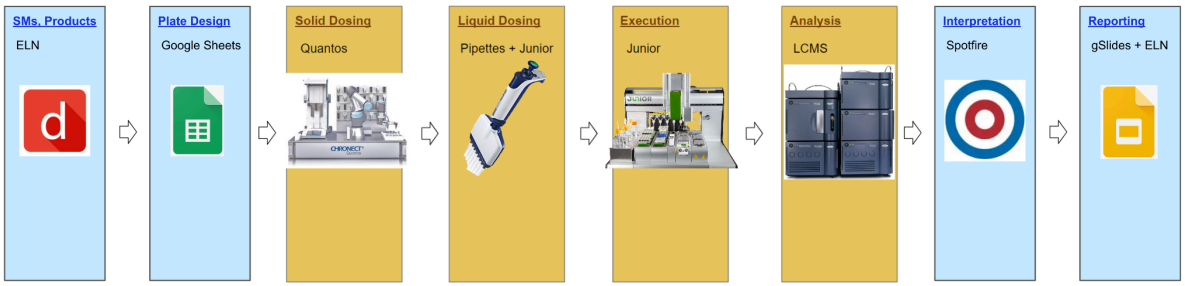


Figure 1: The HTE OS workflow streamlines lab tasks in brown boxes using a highly automated data workflow in blue boxes, guiding the operator from request initiation to result reporting.

After four years of building our workflow, we feel confident sharing it and the underlying principles. We support chemists from discovery all the way to late-stage process development. Our customers reach us by email, chat or stop by. We put this information into our Dotmatics Electronic Lab Notebook (ELN, Version: 6.1-814-s, Figure 1). Then we design plates and generate all files needed to run them in the lab in a self-built Google Sheet application. Every plate is sampled at least twice and the analytical data is automatically pushed into Spotfire[7] for data analysis. Key results are documented in a Google Slides presentation the scaffold of which is generated automatically. On that basis, our customers decide whether to continue screening or proceed with scale-up. We refined and built the workflow over the course of more than 1,000 plates and we continue adding new features. Since we didn't know at the outset how the end product should look like, we chose software platforms that allowed us to retain control, build ourselves and iterate quickly. In the following, we will walk you through the workflow and the design rationales behind it.[8]

[7] Tibco Spotfire 12.0 is a data analysis software that allows users to connect to a variety of data sources and to process/visualize data. For another example of using Spotfire in chemical process development, see: Kaiser, D.; Yang, J.; Wuitschik, G. Using Data Analysis To Evaluate and Compare Chemical Syntheses. *Org Process Res Dev* **2018**, *22* (9), 1222–1235. https://doi.org/10.1021/acs.oprd.8b00199.

[8] The supporting information contains more technical detail as well as the source code and instructions on how to implement everything presented herein. We are aware that our workflow relies on proprietary software such as Google Sheets and Spotfire, which might limit flexibility and customization. This dependence could pose challenges in terms of data security, software updates, and integration with other platforms.

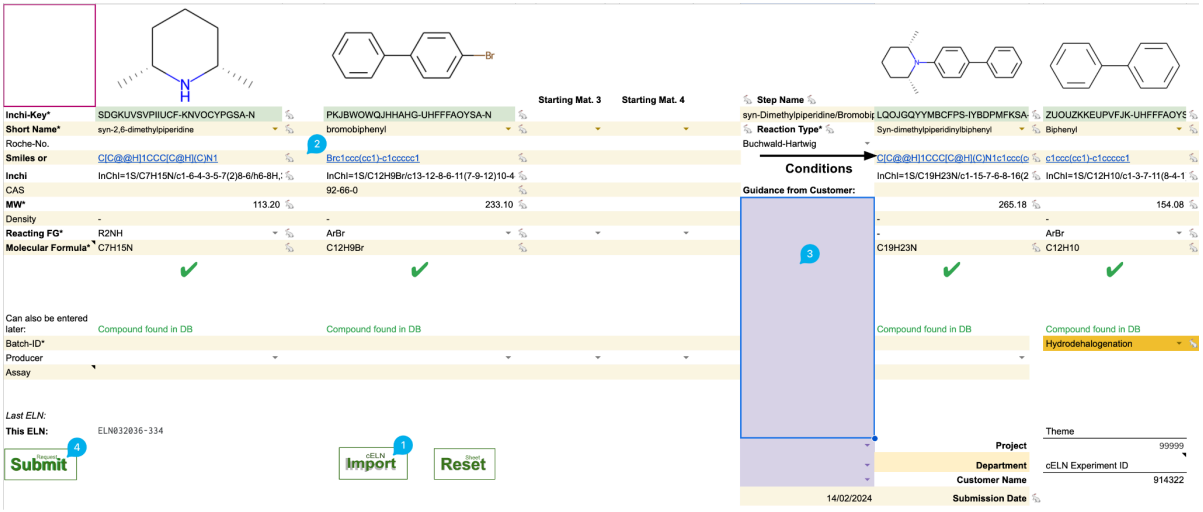


Figure 2: Registering a new reaction in the sheet *Submit Request* can be achieved either through importing the reactants (1) or pasting their SMILES strings (2). After entering metadata such as project name, customer and reacting functional group (3), the request can be submitted (4).

Our workflow is based on merging customer requests with plate layouts to arrive at a final plate design. SMILES[9] strings of reactants and products are exported from our ELN[10] as an Excel file and imported into our Google Sheet application (Figure 2). In there, other data points such as Inchi Key, molecular weight and formula are automatically generated using a FASTAPI wrapper[11] for RDKit[12] and isospec++[13]. After specifying additional metadata such as customer guidance and type of (side-) reaction, the request is processed and a Google Slides presentation generated.

[9] Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28* (1), 31–36. https://doi.org/10.1021/ci00057a005.

[10] We use an ELN application from Dotmatics, but the tool can be adapted to any other file formats as long as they contain the SMILES-strings of the starting materials and products in a defined manner.

[11] FastAPI is a high-performance Python web framework for building APIs, and a FastAPI wrapper is used to simplify and expose complex functionalities, such as cheminformatics tools, as easy-to-consume web services: FastAPI
Ramírez, S. repository-code: https://github.com/tiangolo/fastapi ,url: https://fastapi.tiangolo.com, license: MIT

[12] RDKit: Open-source cheminformatics. https://www.rdkit.org

[13] (a) Łącki, M. K.; Startek, M.; Valkenborg, D.; Gambin, A. IsoSpec: Hyperfast Fine Structure Calculator. *Anal. Chem.* **2017**, 89 (6), 3272–3277. https://doi.org/10.1021/acs.analchem.6b01459. (b) Łącki, M. K.; Valkenborg, D.; Startek, M. P. IsoSpec2: Ultrafast Fine Structure Calculator. *Anal. Chem.* **2020**, 92 (14), 9472–9475. https://doi.org/10.1021/acs.analchem.0c00959.

Figure 3: New plates can be created by modifying a standardized plate (1) or starting from scratch with different substance categories and substances (2). The sheet automatically computes all possible permutations for each entered substance, allowing the operator to select which combinations to include on the plate (3).

As a next step, users can build and design screening plates (Figure 3). They can choose up to three substance categories to be placed in rows or columns and start either from a blank design or a standard plate.[14] For each substance category, individual compounds can be selected as well as the number of variations of the same compound. The latter can be used to compare the performance of different batches, molar equivalents or ways of dosing. The system provides all permutations in rows and columns for the user to include in the plate design. There is also the possibility to screen more than twelve members of a substance category and the tool will alert users if there are no dosing heads or bottles of any selected substance available in the gloveboxes or on the solid dosing robot.

[14] We include a number of plate designs for common reactions that are based on our own experience and literature-based cheatsheets: Fitzner, M.; Wuitschik, G.; Koller, R. J.; Adam, J.-M.; Schindler, T.; Reymond, J.-L. What Can Reaction Databases Teach Us about Buchwald–Hartwig Cross-Couplings? *Chem Sci* **2020**, *11* (48), 13085–13093. https://doi.org/10.1039/d0sc04074f.

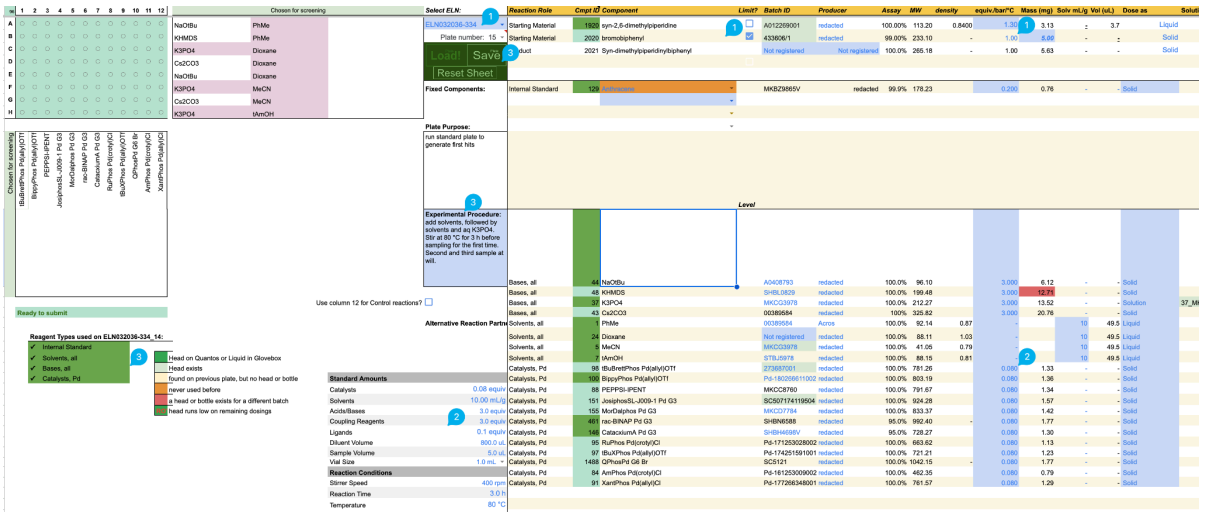


Figure 4: Saving the plate in the sheet *FileGenerator*: Select a reaction (1), choose the limiting starting material and its amount. Choose which batch and how much (2) of each component should be dosed. This can be done individually or by substance category. Note the color coding which indicates availability of compounds for dosing (green) and if there's not enough (red). Save the plate (3) after making sure all types of components are present and everything is filled in correctly. This will trigger the generation of all necessary input files and support documents.

In a third sheet called *FileGenerator*, we combine the plate with a reaction and choose for each component how it should be dosed and how much (Figure 4). While the field is automatically pre-filled based on the presence of density information, the operator can also choose from solutions registered for this compound. The sheet also provides a choice of different batches and indicates for each batch whether it's available for dosing. It will automatically flag intended amounts, if the total weight needed to run the plate exceeds what is available in the glovebox. Finally, the user has to define reaction temperature, time, goal and procedure before saving the plate. During that process, all data is saved to different data tables, the corresponding presentation updated and the files needed for solid-dosing, liquid-handling, LC(MS) and plate setup are written to Google Drive. Previously saved plates can be reused for other reactions or overwritten.

Next we put solid dosing heads on the robot and dosing commences. A script running in the background of the Google Sheet continuously monitors the state of the solid dosing robot and alerts the users when the dosing is finished or in case of issues. After that, the user loads log files into the system and decides whether additional dosings are required to reach the target weight. If that's the case, the system will generate a correction file.

When all solids are dosed, the plate is manually transferred to a second glovebox[15] where we dose the liquids with an electronic pipette following instructions generated by the Google Sheet. After stirring for the prescribed time and at the prescribed temperature, the plate is allowed to cool down.

[15] We consciously separate gloveboxes for solid and liquid handling to eliminate the risk of liquid vapor-mediated solid decomposition.

We take a sample and re-heat the plate. For the final sample, we quench the reaction mixtures with aqueous acetonitrile and draw an aliquot. LCMS sample submission relies on sequence files which are also generated while saving the plate. These contain molecular formulas of all the reaction's starting materials and expected products. When the LCMS-sequence is finished, the resulting text file[16] is automatically uploaded to the Google Cloud Platform, imported into the database and a notification sent using Google Chat. Other analytical data, for example originating from SFC, GC or HPLC, can be imported using the Google Sheet by copying and pasting a peak list.

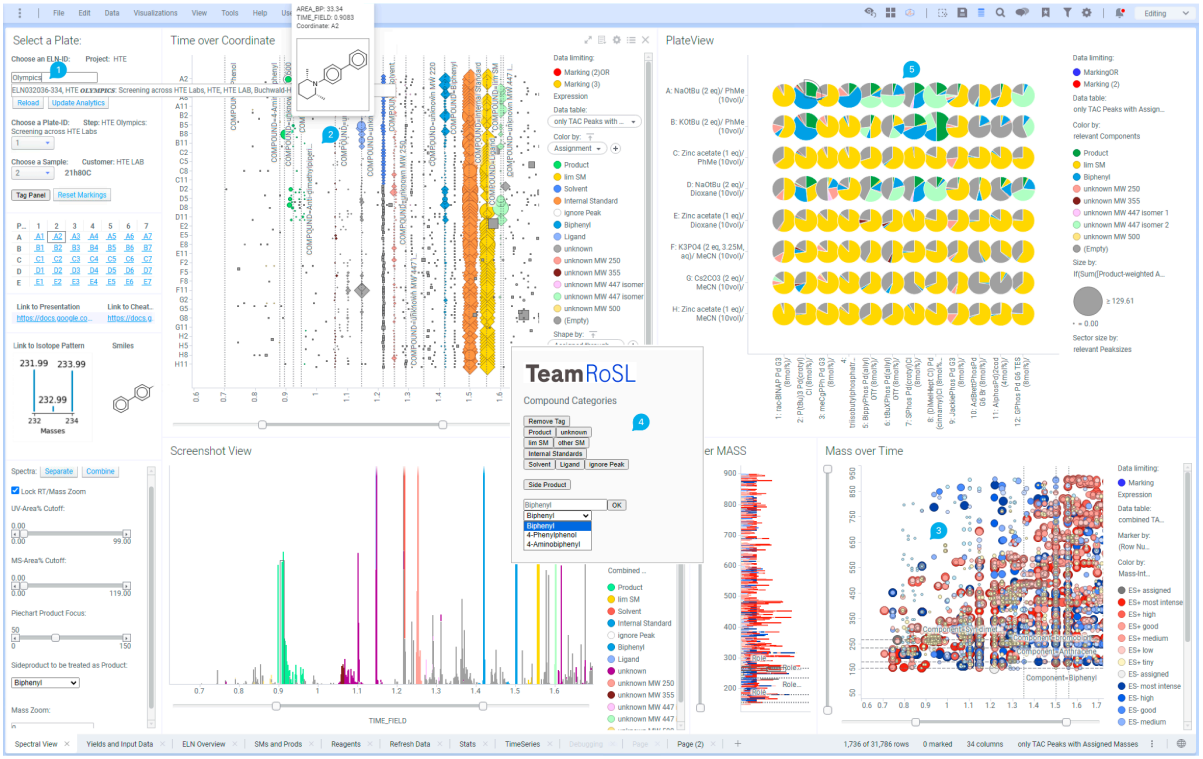


Figure 5: *Spectral View* in HTE OS - Data Analysis provides tools to browse large amounts of analytical data and tag peaks with chemical information. Full-text search (1) to load all analytical data of one customer request with the ability to focus on individual plates and samples. Peaks are colored by the assigned structure (2) which is displayed on mouse-over. Retention times of compounds that were tagged in previous samples are indicated with dotted lines. MS-peaks are positioned on a 2D-plane (3) as a function of retention time and mass. Groups of UV-peaks selected based on their retention time or mass can be tagged (4) with chemical information or flagged to be ignored. The composition of each vial is displayed as a pie chart (5) and all pie charts are put into the chemical context of the plate.

Once the import is finished, users can analyze the data in Spotfire. All peaks in which the LCMS detected one of the specified compounds are automatically tagged with the corresponding structure (Figure 5). Since not all assignments are correct and not all compounds will ionize all the

[16] The data comes as rpt-files which are generated by Waters MassLynx V4.2 SCN989. They contain all peaks of all samples present on the plate as well as aggregated mass data found in each peak.

time, the users can mark peaks and tag them manually based on retention time or the masses present. In the visualization *Mass over Time* (Figure 5, bottom right), the five most intense mass signals of all peaks are displayed as a function of retention time. Therefore, the combined dataset can be analyzed by retention time, mass or both. The automatic drill down makes it easy to recognize isotope patterns, identify all peaks that contain certain masses and thus assign structures.[17] Based on the peak assignments the Spotfire tool automatically generates a pie chart visualization that displays the composition of all vials on the plate.

Figure 6: *Yields and Input Data* is visualizing Yields and Robot Data allows moving beyond peaks. Navigation to select which ELN-ID, plate or sample to display (1). Visualization of product yield (2) and other reaction components of interest. Simple kinetics (3) displaying how the chosen component yield changes over time. Suite of visualizations (4) highlighting potential analytical outliers and actual compound amounts dosed.

If an internal standard is present, the tool will also generate relative yields for each component in each vial from the actual weights dosed of limiting starting material and internal standard as well as the respective peak areas (Figure 6). These yields will be absolute, if a reference sample of the component in question together with the internal standard was measured. It also generates simple kinetics by displaying the change in yield over time. This is useful to track reaction progress and detect decomposition. The Spotfire tool also visualizes the relationship between calculated yields and peak areas. Outliers will show up as data points that don't obey the commonly observed linear relationship. These often result from issues in peak integration or sampling errors. If we detect outliers, we tag them and the system attempts to correct them. It will also display the actual weights

[17] We recorded a live demonstration of the platform which can be found here:

dosed as well as the equivalents relative to limiting starting material for each vial. This feature enables the explanation of anomalies in reaction performance caused by dosing errors.

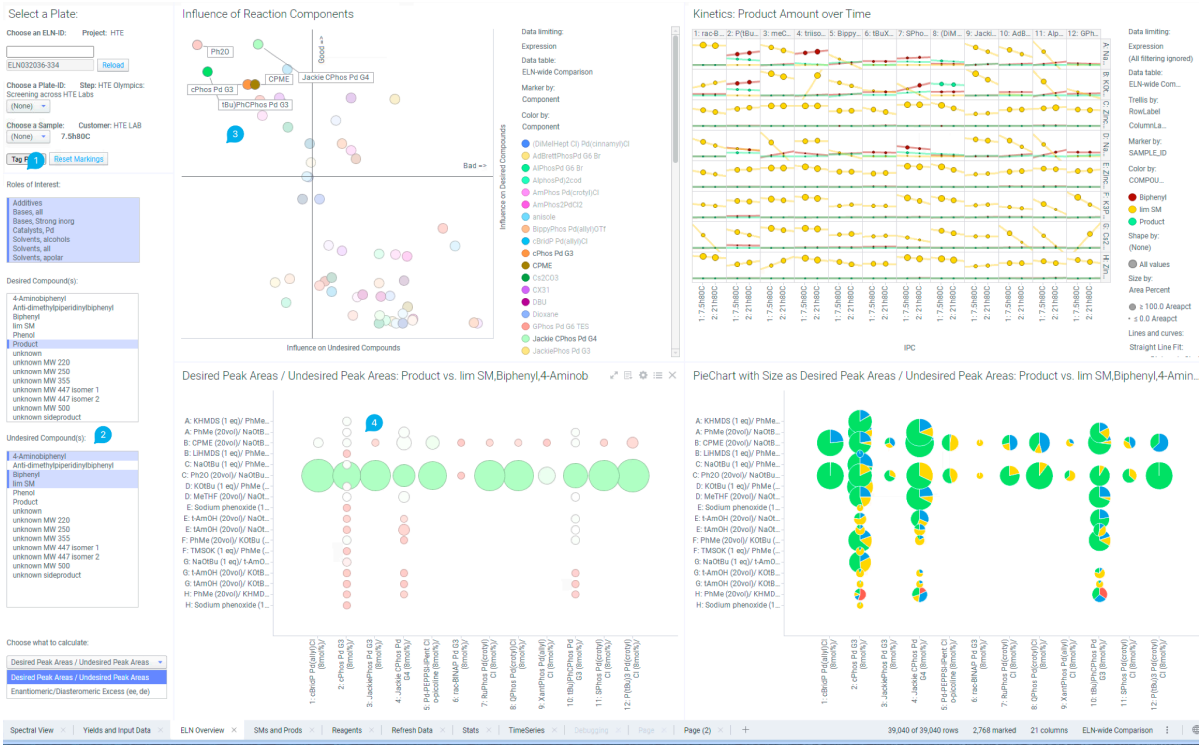
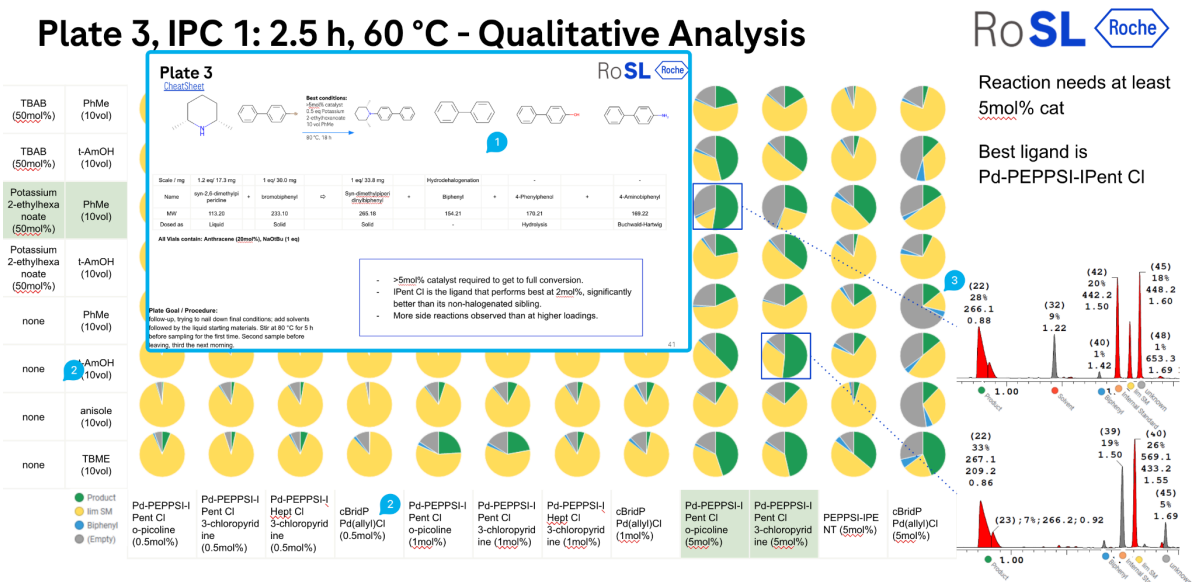


Figure 7: *ELN Overview* allows users to gauge the influence of individual components on the outcome of a reaction. Users can choose which reaction, plate(s), sample(s) and types of reaction variables to consider (1). Next, users can pick desired and undesirable compounds (2). This selection determines which peaks are considered for the calculation of the two axes Good/Bad in the *Influence of Reaction Components* visualization (3). In it, sums of peak areas for vials that contain the component in question are compared with vials that don't. This visualization (4) displays the ratio of desired compounds over undesirable ones for the reaction components selected in *Influence of Reaction Components*.

Knowing what are the most beneficial reaction inputs can be difficult when several plates are run for a chemical transformation. For that purpose, our Spotfire tool offers a set of visualizations in which the user can evaluate all plates at once (Figure 7). The user has the option to select one or more desired and undesired compounds, along with one or more types of reaction inputs such as catalysts, solvents, or bases. Spotfire will then show how much any of these components influence the formation of the desired and undesired components. In a separate visualization, the tool will calculate a simple ratio or enantiomeric/diastereomeric excess of desired to undesired compounds.

Figure 8: Search through all your HTE-data using *SMs and Prods*. Filter your data using different characteristics (1) including substructure. Browse the results and jump to individual reactions (2).

When going beyond a single transformation, the tool offers full substructure search and drilldown into all previous screening campaigns (Figure 8). This is helpful for plate design, but also for deriving trends and consulting with customers on next steps. This tool provides backlinks to detailed results as well as a peek into the best conditions identified for each transformation and a rating on how successful the screening was deemed to be. A similar tool exists that provides sub-structure search and drilldown for the compound inventory including available batches, amounts and solutions.



https://doi.org/10.26434/chemrxiv-2024-c071p **ORCID:** https://orcid.org/0000-0001-6682-1213 Content not peer-reviewed by ChemRxiv. **License:** CC BY 4.0

Figure 9: Google Slides generated when saving a plate come prefilled with an overview that includes reaction equation and input table (1) as well as the plate design (2). The pie charts are pasted in from Spotfire and chromatograms inserted for vials of interest (3).

We present the results using Google Slides by amending the scaffold generated by the Google Sheet (Figure 9). Customers are provided access to the same Spotfire tool via its webclient and how the plate was executed. They receive a recommendation on how to proceed, but are always in control over the next steps. When the optimization is finished, we conduct a scale-up or hand back the starting materials and help in transferring the results to the lab environment. We also follow up with customers to make sure they could isolate the product in sufficient quantity and confirm its structure[18]. This adds to our knowledge of conditions that don't work well outside of the glovebox or result in problems upon scale-up.

Our Setup and Workflow

We believe that manual data entry and copying should be avoided, and that all relevant data should be easily accessible.We want users to enjoy doing HTE and relieve them from tedious tasks. Our goal is to facilitate easy error detection and prevention, making it challenging to commit errors initially.We want to make collaboration and data exploration enjoyable. HTE is a means to solve our customers' problems.

In order to reach these goals, we need to create results that stay relevant beyond the task at hand. Good quality data is a gift that keeps on giving, bad data will continue to mislead and is difficult to identify. We arrived at a standardized workflow which we continuously optimize, also to maximize throughput. We recognize that standardization comes with compromises and we accept them.[19]

As part of the standardization, we aim to use solid-dosing and a standard LCMS-method as much as possible. Precision will suffer for the former when aiming below one milligram target amount and, for the latter, peak shape or separation may not always be ideal. Reactions are run at a scale between one to five milligrams of limiting starting material in the presence of an internal standard and twenty to one hundred microliters of solvent. They are mostly conducted in one milliliter vials as part of ninety six or twenty four well plates. We perform liquid dosing using electronic pipettes and each reaction mixture is sampled manually at several time/temperature points.

[18] If the structure is not confirmed, the tagging will be amended.

[19] Christensen, M.; Yunker, L. P. E.; Shiri, P.; Zepel, T.; Prieto, P. L.; Grunert, S.; Bork, F.; Hein, J. E. Automation Isn't Automatic. *Chem Sci* **2021**, *12* (47), 15473–15490. https://doi.org/10.1039/d1sc04588a.

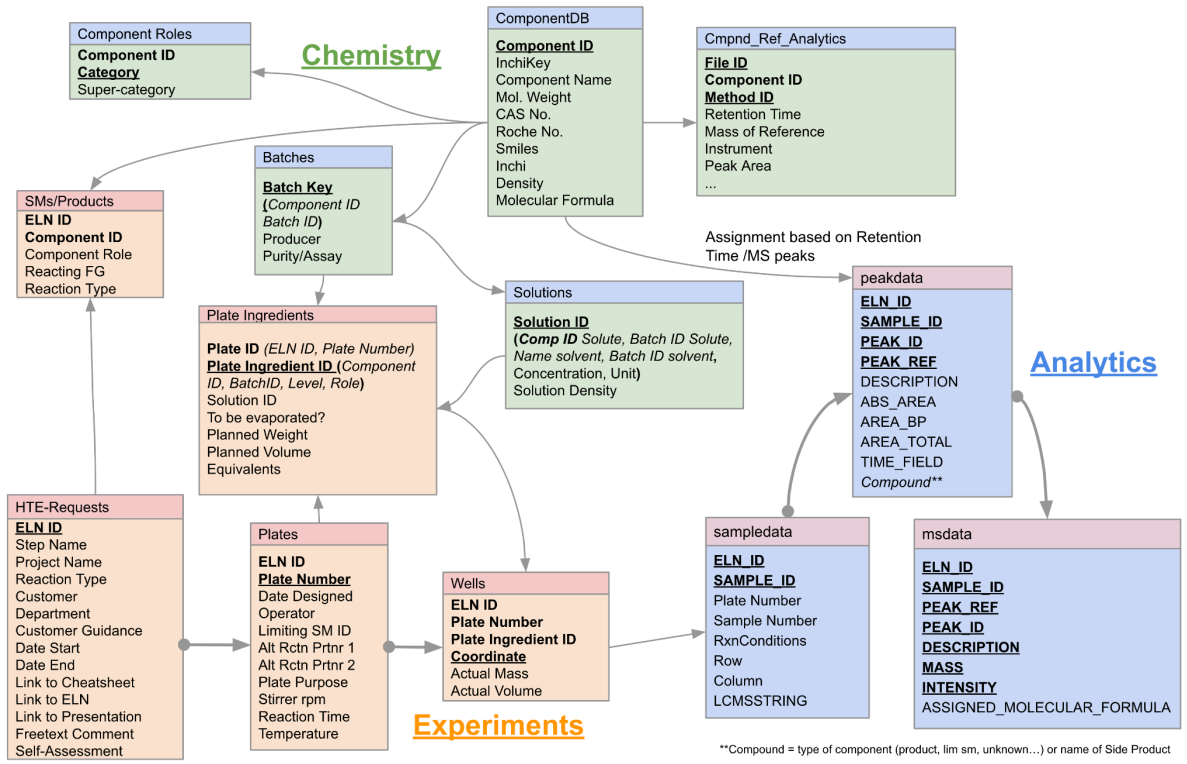


Figure 10: Our way of structuring HTE-data: Individual data fields are combined in tables that are linked via common identifiers. Fields that uniquely identify a row in the table are shown in bold, identifiers that emanate from this table are additionally underlined.

In order to support the workflow, we developed a hierarchical schema to organize our data (Figure 10). We clustered the data fields around the topics of Chemistry, Analytics and Experiments. Within these, related data is combined in data tables in which every line is guaranteed to be unique with respect to one or a combination of several identifiers. A data structure results from linking individual tables containing the same identifiers. All data is stored in a MS SQL Server database unless needed for constant lookup by our Google Sheet application. The way the data is structured and what data points are captured, determines which questions can be answered. In our case, the data structure allows for example to figure out the amount and type of material in each vial and know which compounds were assigned to which peaks in LCMS afterwards. Data is written and consumed through user interfaces or automatically through a Google Cloud Function parsing Waters rpt-files.[20] The latter extracts data from these and writes them to the three data tables labeled Analytics in the bottom right of Figure 10.

[20] Another great parser for rpt-files was recently published: Mason, J.; Wilders, H.; Fallon, D. J.; Thomas, R. P.; Bush, J. T.; Tomkinson, N. C. O.; Rianjongdee, F. Automated LC-MS Analysis and Data Extraction for High-Throughput Chemistry. *Digit. Discov.* **2023**, *2* (6), 1894–1899. https://doi.org/10.1039/d3dd00167a.

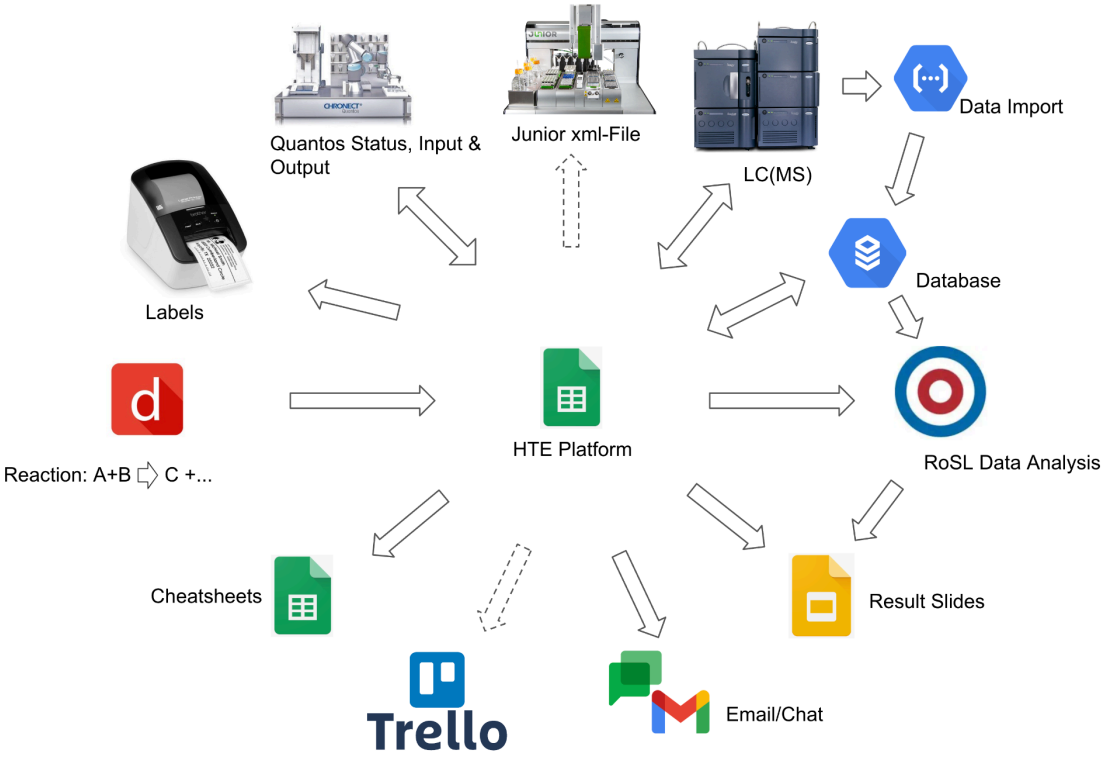


Figure 11: The HTE Platform serves as the workflow backbone, importing reaction schemes from the ELN, generating plate designs and input files. Analytical data is automatically imported into the database. A Spotfire application combines LCMS data with chemical and solid dosing data for interpretation.

Shown in Figure 11, the Google Sheet *HTE Platform* functions as the primary workflow hub. It tracks available compounds for dosing and keeps an inventory of available solids and liquids. Thus, the user knows during plate design whether a compound is present in the gloveboxes, which batches and how much. The HTE Platform reads and writes files to Google Drive using Google Apps Script and sends notifications by email or chat. The files written include, among others, input for Quantos Chronect, Waters Masslynx, and Google Slide presentations. The latter contain reaction equations including chemical structures and plate designs to minimize the time spent on reporting results.

Analysis of the data takes place in Spotfire which is connected to all experimental, chemical and analytical data. We take screenshots of these results, put them into the Google Slides presentation combined with conclusions and screenshots of selected chromatograms. The presentation is archived in the ELN and serves as the backbone for customer communication, although customers also have access to all raw data using the Spotfire web client. We also use the Spotfire application to inform plate designs, be it follow-up plates or initial screens for newly submitted transformations.

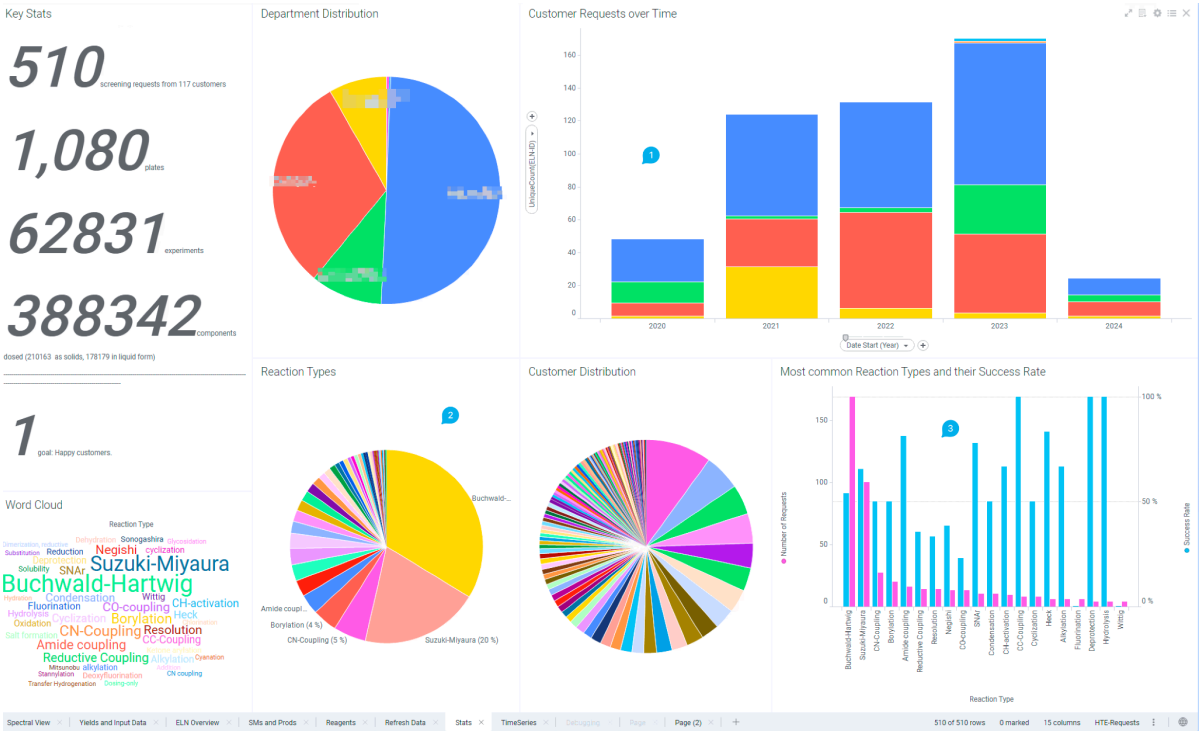


Figure 12: Statistics visualizing changes over time (1), composition with respect to chemical transformations (2) and success rates (3).

We see that our customers like HTE and that they have a clear preference for coupling reactions (Figure 12). We also observe that our success rates vary greatly with reaction type. Seeing that we find an acceptable solution for little over fifty per cent of Suzuki-Miyaura couplings is unexpected, but may be a result of survivorship bias. Moreover, the success rate only considers reactions that were submitted for screening. It doesn't take into account reactions where our cheat sheets[21] already yielded satisfactory results. Customer preferences change over time as do their targets. Our goal is to make chemists happy and learn as much as possible on the way.

Conclusion

The popularity of HTE is still increasing and it's approached from many different angles. We built our dream HTE lab from scratch and this is what we learned: HTE should be the first line of defense, not the last hope. Thus, we encourage our customers to contact us early. HTE is great for quickly exploring chemical space and generating leads. Yet it is no panacea. Reaction miniaturization and execution in the glovebox improve robustness and reduce material consumption. But that also means that not everything will work on a larger scale in the lab. It's our job to anticipate problems and design plates with the customer's intention in mind.

[21] Based on the cheat sheets (Ref. 14) and our own data, we assembled a set of recommendations for Suzuki-Miyaura and Buchwald-Hartwig coupling. These are part of the supporting information.

We believe in picking the best hardware for the job and in replacing humans only when it makes sense. We haven't seen for example a flexible and efficient solution for dosing a large variety of different liquids on a small scale and none for sampling from often heterogeneous reaction mixtures at less than 50 uL total volume. Therefore, we perform these steps manually using multi-channel and electronic pipettes respectively. This results in a partially automated workflow that is simple, robust and highly productive.

Software is a core competency in HTE. If you don't want to write your own tools, insist on having a developer within your team. Prioritize flexibility over robustness until you have a clearer understanding of your specific requirements. When choosing platforms to build your software on, go for ones that can easily be modified, are widely used and familiar.

Data is a gift that keeps on giving, but only if it is well structured, centrally stored, comparable and easily accessible. There is much to learn about chemical reactivity through HTE, but only if one can efficiently draw conclusions from all of the data generated and flexibly design plates. We hope that by making our tools available, a first step has been made to a true open-source platform for HTE.

Acknowledgements

We'd like to thank Andreas Marx, Christian Bartelmus, Monique Ucar, Maxim Popov, Ryan Burwood, Marco Nestola, Rick Sidler, Jean-Michel Adam, Irene Marzuoli, Raphael Bigler, Raffael Josef Koller and Andreas Schuster for their input in building the laboratory and feedback.