# D 1.7 Data Management Plan

## Date: 28-02-2019

Smart in-line metrology and control for boosting the yield and quality of high-volume manufacturing of Organic Electronics
(SmartLine)
Grand Agreement: 768707

| Project co-funded by the European Commission within Horizon 2020 Research and Innovation Programme | | |
|---|---|---|
| **Dissemination Level** | | |
| PU | Public | X |
| PP | Restricted to other programme participants (including the Commission Service | |
| RE | Restricted to a group specified by the consortium (including the Commission Services) | |
| CO | Confidential, only for members of the consortium (excluding the Commission Services) | |

Horizon 2020
European Union funding
for Research & Innovation

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Call H2020 FOF-08-2017

# Copyright



| | | |
|---|---|---|
| Consisting of Coordinator: | Aristotle University of Thessaloniki (AUTh) | Greece |
| Partners: | Centro Ricerche FIAT SCPA (CRF) | Italy |
| | Aixtron SE (AIXTRON) | Germany |
| | Suragus GmbH (SUR) | Germany |
| | IBS Precision Engineering BV (IBS) | The Netherlands |
| | Organic Electronic Technologies Private Company IKE (OET) | Greece |
| | Laytec Aktiengesellschaft (LayTec) | Germany |




This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Call H2020 FOF-08-2017

# Contents

# 1. Introduction

The SmartLine Project is set out to create intelligent and zero-defect manufacturing processes by developing robust and non-destructive in-line metrology tools (optical, electrical, structural) and a process control platform to achieve the reliable and closed-loop manufacturing of Organic Electronic devices (OPVs and OLEDs for lighting) by R2R printing and OVPD pilot lines.

The SmartLine tools, methodologies and platform will boost the manufacturing yield for OPV and OLED devices by achieving unprecedented homogeneity in nanolayer thickness, optical, electrical and structural properties, optimum resource utilization, minimization of process waste, and predictive/preventive corrections on the entire process.

This project will have a huge impact and will transform the manufacturing processes for Organic Electronics Industry and for other Industries as Thin Films (e.g. functional films, antimicrobial coatings, barrier thin films), Electronics, Wearables, Energy, Automotive, Transport, Space, Health, etc., and incorporate them towards the Factory of the Future.

This deliverable constitutes the Data Management Plan (DMP) of the SmartLine project. This is a short, general outline of the project policy for data management (i.e. collecting and sharing procedures), that includes the following issues:

1) Outline the types of data already generated and/or foreseen to be generated at the project, including the context and procedures of this generation, as well as the degree of privacy and confidentiality of the data.
2) Outline the procedures that will be followed to assess the generated/collected data with respect to their sensitiveness.
3) Outline the data acquisition plan for the duration of the project.
4) Outline the measures that are foreseen for the adequate management of the data from the ethical and security points of view.

Based on the above, the DMP is a document outlining how the research data that will be generated by SmartLine will be handled during the project, and after it is completed. The overall purpose of the DMP is to support the data management life cycle for all data that will be collected, processed or generated by the project.

The described policy herein reflects the current state of consortium agreements regarding data management and is consistent with those referring to exploitation and protection of results. This deliverable can also be considered as a living document that will evolve during the project and will be updated accordingly by the partners.

# 2. SmartLine Experimental Data

This paragraph reports on the organization of the SmartLine datasets of the individual metrology tools, that will be conducted in the context of SmartLine and their characteristics, along with the details on the date to be generated. In more details the SmartLine experimental data include datasets for the following tools:

- Spectroscopic Ellipometry
- Raman Spectroscopy
- Reflectometry
- Wavelength Scanning Interferometry
- Eddy Current Measurements

These datasets include characterization data and information from specific materials and their combinations in multilayer architectures to form Organic Electronic (OE) devices, such as Organic Photovoltaics – OPVs and Organic Light Emitting Diodes – OLEDs.

Within the course of the project, the partners will develop the methodologies for the optimum measurement and analysis of the optical properties of OE nanomaterials and devices by SE and RS, as well as novel tools (Reflectometer, WSI, EC) and the associated acquisition and modelling methodologies for characterization of OE nanomaterials and devices. Therefore, the partners will generate experimental data during the testing of the developed tools ex-situ, as well as in-line when these tools will be integrated at the novel Roll-to-Roll (R2R) printing and OVPD pilot lines.

These data will be used for the understanding of the structure-property relationships of the studied nanomaterials and for the optimization of the nano-manufacturing processes of OE devices.

The Key attributes, characteristics, and other information related to SmartLine datasets, are described in the following sections.

## 2.1. DataSet 1: SmartLine Spectroscopic Ellipsometry Data

**Data Summary**

| | |
|---|---|
| Purpose of the Data | Information of the measured optical properties of samples (single, multilayers), either ex-situ, or in-situ, or in-line |
| Type and Format of data | Proprietary software format (.spe, .clc). Also, these can be converted to other types of format for processing by standard editors, such as Tab delimited text (.txt) |
| Reused-Data | Data can be re-used as reference for the modelling of SE measurements from complex material structures (multilayers) |
| Data origin | Proprietary software of the SE tool manufacturer (Horiba, DeltaPsi2©) |
| Data size | Variable, starting from kilobytes (for ex-situ recorded spectra) up to several megabytes (for in-line recorded spectra that contain several measurements in the same file) |
| Data Security and Storage | Data are not encrypted and they are locally stored on the SE tool. These can be also copied to external storage media for additional processing, modelling and analysis |
| Data value<br>(Long Term) | Long term reference purposes, archiving |

**FAIR DATA- Making Data Findable**

| | |
|---|---|
| Discoverability of data (metadata provision) | No metadata provided in the data filename. The body of the data file includes specific metadata in the first 60 lines of the file, with information for the experimental conditions used for the measurement, which include the spectral range, azimuth angles of optical elements, acquisition speed, acquisition resolution, number of data, date. |
| Identifiability of data (refer to standard id mechanisms) | Identifiable by file name fields (i.e. run, material, date). Also, the file extension and format of the data refers to the functionality of the data; spe=measured data, clc=calculated data based on simulation, or modelling routines |
| Naming conventions used | User defined (i.e. run_material_date.ext) |
| Search keywords approach | Search by user defined fields |
| Clear versioning approach | Only user defined name of the file |
| Standards or procedures for metadata creation applied | The metadata included in the file are automatically generated by the SE tool based on the experimental conditions used for the measurement, which include the spectral range, azimuth angles of optical elements, acquisition speed, acquisition resolution, number of data, date. |

**FAIR DATA- Making Data Openly Accessible**

| | |
|---|---|
| Data openly available or kept close | Measurement data are kept close. Modelling data (e.g. calculated bulk refractive index or dielectric function) can be made available after release from AUTh |
| How data will be made available | Specific reference data for specific materials (calculated bulk refractive index or dielectric function) can be made available after release from AUTh. These can be opened with any text Editor software |

| | |
|---|---|
| Methods or SW tools for data access | LAN access via local Windows server and applicable remote control SW |
| SW documentation and other information needed | Credentials for access over closed network |
| Repository for deposit of data, metadata, documentation and code | All data are stored locally at the control systems (PC) of the SE tool |
| Access restrictions | Credential and password protection |
| Data interoperability assessment | The .txt data (simulated spectra) can be opened by any conventional text Editor program, or analysis software |

**FAIR DATA- Making data interoperable**

| | |
|---|---|
| Standard vocabulary or mapping to commonly used ontologies | ASCII |
| Data licensing for wide reuse | Not applicable |

**FAIR DATA- Increase data re-use (through clarifying licenses)**

| | |
|---|---|
| Timing of data availability for re-use (incl. indications on embargo) | Confidential availability upon request/agreement/licensing |
| Data usability by Third Parties (after the end of the project) | Confidential usability upon request/agreement/licensing |
| Restrictions to data re-use | Confidentiality and license induced restrictions |
| Quality assurance process | Restricted access backups |
| Length of time of data re-usability | No time restriction |

**Allocation of resources**

| | |
|---|---|
| Costs estimates for making data FAIR | Upon request/agreement/licensing from manufacturer/proprietary software provider |
| Data Management Responsibilities | Periodically ensured storage quality and protection |

## 2.2. DataSet 2: SmartLine Raman Spectroscopy Data

**Data Summary**

| | |
|---|---|
| Purpose of the Data | Material identification |
| Type and Format of data | • Proprietary software format (.l6s)<br>• Tab delimited text (.txt) |
| Reused-Data | License attained data prompt to confidentiality |
| Data origin | Proprietary software (LabSpec 6) |
| Data size | Varies from kBs up to MB order of magnitude |
| Data Security and Storage | Unencrypted data locally stored |
| Data value<br>(Long Term) | Long term reference purposes, archiving |

**FAIR DATA- Making Data Findable**

| | |
|---|---|
| Discoverability of data (metadata provision) | No metadata provided |
| Identifiability of data (refer to standard id mechanisms) | Identifiable by file name fields (i.e. run, material, date) |
| Naming conventions used | User defined (i.e. run_material_date.ext) |
| Search keywords approach | Search by user defined fields |
| Clear versioning approach | No versioning of data |
| Standards or procedures for metadata creation applied | No metadata provided |

**FAIR DATA- Making Data Openly Accessible**

| | |
|---|---|
| Data openly available or kept close | Kept close |
| How data will be made available | Available over local network |
| Methods or SW tools for data access | LAN access via local Windows server and applicable remote control SW |
| SW documentation and other information needed | Credentials for access over closed network |
| Repository for deposit of data, metadata, documentation and code | Hardware-paired PC's local storage |

| | |
|---|---|
| Access restrictions | Credential and password protection |
| Data interoperability assessment | • *.l6b data only operable by proprietary software<br>• *.txt data interoperable by txt supporting software |

**FAIR DATA- Making data interoperable**

| | |
|---|---|
| Standard vocabulary or mapping to commonly used ontologies | ASCII |
| Data licensing for wide reuse | Not applicable |

**FAIR DATA- Increase data re-use (through clarifying licenses)**

| | |
|---|---|
| Timing of data availability for re-use (incl. indications on embargo) | Confidential availability upon request/agreement/licensing |
| Data usability by Third Parties (after the end of the project) | Confidential usability upon request/agreement/licensing |
| Restrictions to data re-use | Confidentiality and license induced restrictions |
| Quality assurance process | Restricted access backups |
| Length of time of data re-usability | No time restriction |

**Allocation of resources**

| | |
|---|---|
| Costs estimates for making data FAIR | Upon request/agreement/licensing from manufacturer/proprietary software provider |
| Data Management Responsibilities | Periodically ensured storage quality and protection |

## 2.3. DataSet 3: SmartLine Reflectometry Spectroscopy Data

**Data Summary**

| | |
|---|---|
| Purpose of the Data | Provision of layer thickness |
| Type and Format of data | The statements made in this document by LayTec refer to data that are the results of the evaluation of raw data. Raw data are only available for LayTec-internal use.<br>Layer thickness: entry in SQL database<br>Meta data:<br>• Number of optical head<br>• Position of measurement along web<br>• Timestamp<br>• Data type<br>• Process name (given by operator, e.g. layer type) |
| Reused-Data | Thickness values of bottom layers are used when calculating thickness of top layer |
| Data origin | Sensors of metrology system<br>Output of Software |
| Data size | 1 KB |
| Data Security and Storage | All data is stored on local hard drive |
| Data value<br>(Long Term) | Layer thickness |

**FAIR DATA- Making Data Findable**

| | |
|---|---|
| Discoverability of data (metadata provision) | Data can be searched an identified according to<br>• Number of optical head<br>• Position of measurement along web<br>• Timestamp<br>• Data type<br>• Process name (given by operator) |
| Identifiability of data (refer to standard id mechanisms) | Data can be searched an identified according to<br>• Number of optical head<br>• Position of measurement along web<br>• Timestamp<br>• Data type<br>Process name (given by operator) |
| Naming conventions used | - |
| Search keywords approach | - |
| Clear versioning approach | - |

| | |
|---|---|
| Standards or procedures for metadata creation applied | - |

**FAIR DATA- Making Data Openly Accessible**

| | |
|---|---|
| Data openly available or kept close | openly available |
| How data will be made available | Data is stored in openly accessible database |
| Methods or SW tools for data access | SQL database management software |
| SW documentation and other information needed | Knowledge of how use SQL database management software |
| Repository for deposit of data, metadata, documentation and code | Local database |
| Access restrictions | none |
| Data interoperability assessment | - |

**FAIR DATA- Making data interoperable**

| | |
|---|---|
| Standard vocabulary or mapping to commonly used ontologies | - |
| Data licensing for wide reuse | - |

**FAIR DATA- Increase data re-use (through clarifying licenses)**

| | |
|---|---|
| Timing of data availability for re-use (incl. indications on embargo) | - |
| Data usability by Third Parties (after the end of the project) | yes |
| Restrictions to data re-use | none |
| Quality assurance process | - |
| Length of time of data re-usability | infinite |

**Allocation of resources**

| | |
|---|---|
| Costs estimates for making data FAIR | 2 person months |
| Data Management Responsibilities | Product owner software development |

## 2.4. DataSet 4: SmartLine Wavelength Scanning Interferometry Data

**Data Summary**

| | |
|---|---|
| Purpose of the Data | Topology measurement |
| Type and Format of data | 3D profile measurement in various windows 7 formats. |
| Reused-Data | None |
| Data origin | WSI |
| Data size | MBytes |
| Data Security and Storage | Local secure network |
| Data value<br>(Long Term) | NA |

**FAIR DATA- Making Data Findable**

| | |
|---|---|
| Discoverability of data (metadata provision) | Basic metadata; structured storage system |
| Identifiability of data (refer to standard id mechanisms) | NA |
| Naming conventions used | IBS in-house |
| Search keywords approach | No |
| Clear versioning approach | NA |
| Standards or procedures for metadata creation applied | No |

**FAIR DATA- Making Data Openly Accessible**

| | |
|---|---|
| Data openly available or kept close | Closed |
| How data will be made available | Not public |
| Methods or SW tools for data access | NA |
| SW documentation and other information needed | NA |
| Repository for deposit of data, metadata, documentation and code | NA |

| | |
|---|---|
| Access restrictions | Company in-house |
| Data interoperability assessment | NA |

**FAIR DATA- Making data interoperable**

| | |
|---|---|
| Standard vocabulary or mapping to commonly used ontologies | NA |
| Data licensing for wide reuse | NA |

**FAIR DATA- Increase data re-use (through clarifying licenses)**

| | |
|---|---|
| Timing of data availability for re-use (incl. indications on embargo) | NA |
| Data usability by Third Parties (after the end of the project) | NA |
| Restrictions to data re-use | Company in-house |
| Quality assurance process | ISO9001 |
| Length of time of data re-usability | NA |

**Allocation of resources**

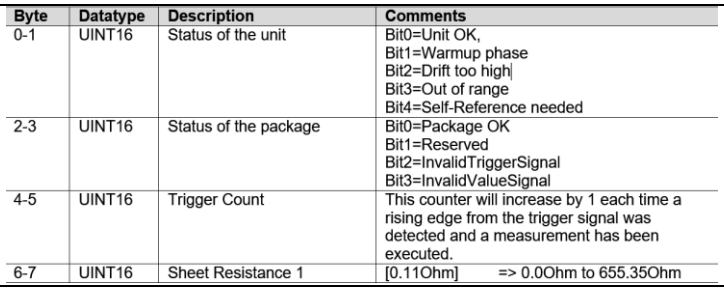
| | |
|---|---|
| Costs estimates for making data FAIR | NA |
| Data Management Responsibilities | NA |

## 2.5. DataSet 5: SmartLine Eddy Current Data

**Data Summary**

| | |
|---|---|
| Purpose of the Data | Resistance values (integrity of printed structure, printing and dying quality) |
| Type and Format of data | User defined protocol (UDP) with customized 16 bit data fields |
| Reused-Data | N.A. |
| Data origin | SURAGUS software creates UDP packages |
| Data size | < 100 bytes (typically 7 bytes) |
| Data Security and Storage | PC is equipped with fire wall |
| Data value<br>(Long Term) | N.A. |

**FAIR DATA- Making Data Findable**

| | |
|---|---|
| Discoverability of data (metadata provision) | N.A. |
| Identifiability of data (refer to standard id mechanisms) | N.A. |
| Naming conventions used |  |
| Search keywords approach | N.A. |
| Clear versioning approach | N.A. |
| Standards or procedures for metadata creation applied | N.A. |

| Byte | Datatype | Description | Comments |
|---|---|---|---|
| 0-1 | UINT16 | Status of the unit | Bit0=Unit OK,<br>Bit1=Warmup phase<br>Bit2=Drift too high<br>Bit3=Out of range<br>Bit4=Self-Reference needed |
| 2-3 | UINT16 | Status of the package | Bit0=Package OK<br>Bit1=Reserved<br>Bit2=InvalidTriggerSignal<br>Bit3=InvalidValueSignal |
| 4-5 | UINT16 | Trigger Count | This counter will increase by 1 each time a rising edge from the trigger signal was detected and a measurement has been executed. |
| 6-7 | UINT16 | Sheet Resistance 1 | [0.11Ohm] => 0.0Ohm to 655.35Ohm |

**FAIR DATA- Making Data Openly Accessible**

| | |
|---|---|
| Data openly available or kept close | N.A. |
| How data will be made available | N.A. |
| Methods or SW tools for data access | N.A. |

| | |
|---|---|
| SW documentation and other information needed | English protocol description and English software manual |
| Repository for deposit of data, metadata, documentation and code | N.A. |
| Access restrictions | PC password and user and expert mode in software (required for recalibration) |
| Data interoperability assessment | N.A. |

**FAIR DATA- Making data interoperable**

| | |
|---|---|
| Standard vocabulary or mapping to commonly used ontologies | N.A. |
| Data licensing for wide reuse | N.A |

**FAIR DATA- Increase data re-use (through clarifying licenses)**

| | |
|---|---|
| Timing of data availability for re-use (incl. indications on embargo) | N.A. |
| Data usability by Third Parties (after the end of the project) | N.A. |
| Restrictions to data re-use | N.A. |
| Quality assurance process | N.A. |
| Length of time of data re-usability | N.A. |

**Allocation of resources**

| | |
|---|---|
| Costs estimates for making data FAIR | N.A. |
| Data Management Responsibilities | N.A. |

# 3. Administrative data and reports

The dataset includes all administrative files related to the reporting and dissemination activities of the project. After the finalization of the project, the data collected from the project (including documents, reports, photographs, deliverables, minutes, presentations) will be maintained by each project partner for a period determined by each partner. The project coordinator (AUTh) will be responsible for the archiving and maintaining the project data files after the completion of the project, for a period covering at least five years. These data include the following:

- Progress reports
- Deliverables
- Meeting documents (photos, agendas, minutes)
- Scientific publications
- Conference material (presentations, posters, photos, conference agenda)
- Patents
- Press releases
- Brochures, flyers

This data will be shared between the partners in readable formats with standard software tools. These formats include PDF, PPTX, DOCX, JPG. Concerning the scientific publications, these will be submitted to Open Access journals. In case that any partner will submit to closed access journals or conferences, the consortium will attempt to make publicly available any versions of them eligible for open access.

After the completion of the project, the data collected from the project will be maintained by each project partner for a period determined by each partner. The project coordinator (AUTh) will be responsible for the archiving and maintaining the project data files after the completion of the project, for a period covering at least five years. All this information will be stored and will be made available to the project partners in the secure area of the SmartLine project.

# 4. Archiving and Preservation of generated SmartLine Data

The general aspects related to the storage, archiving and preservation of the SmartLine data are described below. Each partner involved in the generation of a specific dataset is the owner of this dataset and it will be responsible for the storage and preservation of these data.

After the completion of the project, the data collected from the project (including documents, reports, photographs, deliverables, minutes, presentations) will be maintained by each project partner for a period determined by each partner. The project coordinator (AUTh) will be responsible for the archiving and maintaining the project data files after the completion of the project, for a period covering at least five years. These data include the information exchanged between the participants for the preparation of the reports, deliverables, presentations and minutes of the meetings.

The SmartLine website will be maintained for at least five years after the completion of the project and all the information hosted in the web - site will be maintained by AUTh. The estimated data volume is not expected to necessitate the use of special storage systems. The repository provided by AUTh is hosted on its private infrastructure and therefore the availability is guaranteed for the project duration. It will be put offline after project completion for maintenance and bandwidth cost reduction reasons, however it will still be available for internal use, conforming to each partner's archiving practices.

The generated datasets of the metrology tools will remain at the control systems connected to the metrology tools. Selected datasets from representative materials may be extracted by the responsible partner for distribution, either at on-line media (e.g. in the project website) or to other partner for dissemination purposes.

In the case of the SURAGUS software, this contains a database which has a maximum size of 10 GB. Data can be exported in the csv format. In the SmartLine project, SURAGUS is only storing data for testing purposes. During the project the measurement data is transferred to a global data base that covers required functions for all applied metrology.

# 5. Ethics and confidentiality of generated Data

Any ethical issue that might arise will be dealt by SmartLine consortium with the appropriate care, in a professional way, following very closely established EU regulations and corresponding national laws about privacy, confidentiality, and consent.

All personal and professional data will be addressed by the consortium to ensure that they do not breach national laws on the protection of data. All systems and metrology tools developed in SmartLine will conform to both current legislation and that anticipated in the future. The systems that will be developed in SmartLine will consider "Trust and confidence" as a key attribute due to the sensitivity of personal and business confidential information in terms of exploitation. Therefore, extra care will be taken into preserving this data.

The measurement data obtained on products of OET and AUTh in Ohm/sq provided by SURAGUS can be used by OET, AUTh and SUR.

# 6. Conclusions

In this deliverable, we have presented the data that will be generated in the context of SmartLine including data from the metrology tools and the administrative data related to the project implementation.

The Data Management Plan is designed to comply with the confidentiality level and sharing strategy between the partners and with external entities. Therefore, the project consortium will not share or open data that are considered confidential or rated as not shareable according to ethical principles and rules.

This DMP document will be further updated during the course of the project and it will be reported in month 36. During this period, the partners will define better the specific requirements that will ensure that the innovations of the project and the advances on the development of the Factory of the Future in the manufacturing of high quality flexible Organic Electronic devices, will be shared to the public. In this way, the impact of the project will be extended to much more industrial sectors and applications.