**Naming convention**

This document explains the naming convention used for FTIR sample files. The naming system ensures that all relevant metadata is embedded within each filename, allowing easy classification, filtering, and machine learning processing.

All files are stored in a single folder, with each filename containing information about:

* Each filename follows a structured format with \_ as a separator:
* Infrastructure Role (Road, Dikes, Sealing, Bridge, …)
* Sample Type (Binder, Mastic, Mixture, …)
* Aging Level (Fresh, Short-Term Aged, Long-Term Aged, …)
* Experimental Conditions (Temperature, Pressure, Thickness, Time, Humidity, ROS, UV Light, …)
* Additional Properties (Binder Type, Modification Type and Percentage, Biobased Binder Type and Percentage, Filler, Aggregate, Mixture Type, Layer, Location for Field Samples)
* Data Privacy Classification
* If any information is missing, unknown (**UK**) or if not applicable **None**) will be used. "Unknown" (UK) is used when the information is missing because it is not available or has not been determined. For example, if the aging temperature of a sample is not recorded or known, **"Unknown" would be used in place of the missing value. "None" is used when the information is not applicable to the specific case.** For example, if a particular condition does not involve binder modification, then "None" would be used because the concept itself is irrelevant to that case. Thus, "Unknown" indicates missing or unavailable data, whereas "None" signifies that the parameter does not apply in that context.

***Binder Samples:***

R\_B\_sample origin\_location\_binder type\_penetration grade\_modification\_modification percentage\_biobased binder type\_biobased binder percentage\_aging\_temperature\_pressure\_thickness\_time\_humidity\_ROS type\_ROS concentration\_UV exposure\_data privacy\_project code\_replicate.csv

Example: R\_B\_Lab\_L1\_Q\_PE40+60\_SBS\_5\_LGN\_2\_F\_60\_20\_1\_100\_20\_O3\_12\_40\_OP\_P1\_R1.csv

**Table 4 Naming convention**

| **No.** | **ITEM** | **Description** | **Abbreviation** |
| --- | --- | --- | --- |
| 1infra (b/m/mx) | Infra Role | Which infrastructure the sample is coming from | R = Road |
| 2type (b) | B | Sample type | B = Binder |
| 3origin | Field or Lab | Sample origin | L = Lab |
| 4location | Loc | Location for field samples, go details to cities Lab samples: university or contractor abbreviation like TUD, TUW, RWS, … Field sample: road names like NLA12, NLB13,… For test sections: location of test sections, like TS1 (Lintrack samples with exact location) | L1 = TU Delft Lab   A **logbook**[**Table 2**](https://file+.vscode-resource.vscode-cdn.net/u%3A/FTIR%20ARM/InfraDCM/data/README.md#table-2-logbook-of-location-code) is used to track updates. |
| 5bindertype (b/m/mx) | BinderType | Type of binder or UK if unknown | Q = Q8 T = Total V = Votal |
| 6grade (b/m/mx) | Penetration grade | Penetration grade or PG grade | PE70+100 PE40+60 PG60+22 |
| 7modification (b/m/mx) | Modification | Binder modification or None if not applicable In case of multiple modifications, a combination of tags should be used : SBS+RAP…… | SBS = SBS |
| 8modpct ModPct (b/m/mx) | Modification (weight) percentage(%). | For combination of modifications, multiple concentration should be provided like 5+2+6…..,else None | Numeric value (e.g., 5 for 5%) |
| 9biotype (b/m/mx) | BioType | Type of biobased binder or None if not used In case of multiple biomaterials, a combination of tags should be used : LGB+BPH | LGN = Lignin |
| 10biopct (b/m/mx) | BioPct | Biobased binder procent in weight percentage | Numeric value (e.g., 2 for 2%) |
| 11aging (b/m/mx) | Aging | Aging level | F = Fresh S = Short-Term Aged L = Long-Term Aged |
| 12temp (b/m/mx) | Temp | Test temperature in Celsius, None for fresh samples or UK if unknown, unit C | Numeric value (e.g., 60, 70) |
| 13press (b/m/mx) | Press | Pressure in bar or UK if unknown, unit bar | Numeric value (e.g., 20, 25) |
| 14thick (b/m/mx) | Thickness | Sample thickness in mm or UK if unknown, unit mm or cm. For binder and mastics, this is film thickness (mm). For Mixture, it is mixture layer thickness (cm). | Numeric value (e.g., 1, 2) |
| 15time (b/m/mx) | Time | Aging duration in hours or UK if unknown, unit hour (for lab) /year (for field) | Numeric value (e.g., 100, 200) |
| 16humi (b/m/mx) | Hum | Humidity in %RH or UK if unknown, unit % | Numeric value (e.g., 20, 40) |
| 17rostype (b/m/mx) | ROS Type | Reactive Oxygen Species, In case of multiple ROS, a combination of tags should be used : O3+NOX | O3 NOX OH |
| 18rosconc (b/m/mx) | ROS Conc | ROS concentration or None if not applicable, unit depends on the ROS type and it can be ppm, gr/m3, % for O3, Nox, OH, respectively. | Numeric value (e.g., 12, 10) |
| 19uv (b/m/mx) | UV | UV light exposure in W/m² or UK or None if not applicable | Numeric value (e.g., 40, 50) |
| 28replicate (b/m/mx) | Replicates | replicates | R1 R2 R3 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |