



Data for FoodChain-Lab

Collection – Import – Cleaning – Transformation

Alexander Falenski, Marion Gottschald, Marco Rügen,
Christian Thöns, Matthias Filter, Annemarie Käsbohrer,
Bernd Appel, Armin A. Weiser

Topics



- FoodChain-Lab Templates
 - Start Tracing Template
 - All-in-one-Template
- Backtracing with pre-filled template
- Removing duplicates
- Annotating data for visualisation in the Tracing View

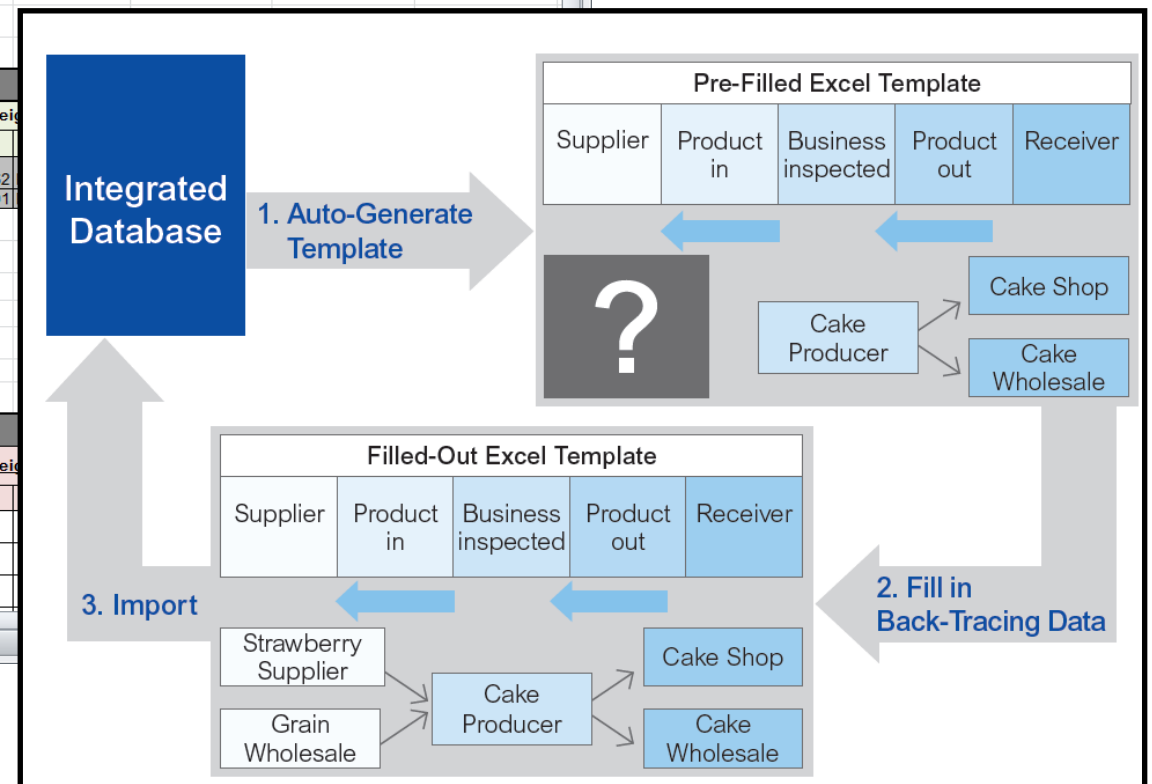
Collecting Data: Excel Template

“One step back-one step forward” principle of REGULATION (EC) No 178/2002, Article 18
Endless and complex supply chains are possible.



Backtrace_request_3001_Caterer 01.xlsx - Microsoft Excel

Reporter Information								
Reporting Officer	Reporting Date	Reporting Remark						
Station in Focus:	3001	Caterer 01						
Products Out								
Product Name	Lot Number	Delivery Date Departure			Delivery Date Arrival			Unit weight
		Day	Month	Year	Day	Month	Year	
Menü 1	M01	3	9	2014	3	9	2014	62
Menü 2	M02	3	9	2014	3	9	2014	101
Lot Information								
Lot Number of "Product Out"	Lot size							
	Quantity	Type / Unit						
M01	900	Portionen						
M02	900	Portionen						
Ingredients for Lot(s)								
Ingredient Name	Lot Number of Ingredient	Delivery Date Departure			Delivery Date Arrival (if departure unknown)			Unit weight
		Day	Month	Year	Day	Month	Year	

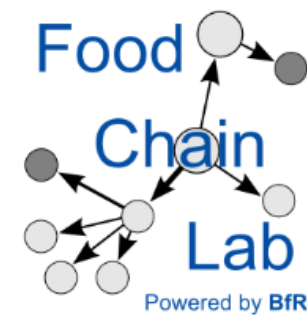


Online Validation: <https://foodrisklabs.bfr.bund.de/templatevalidator/>

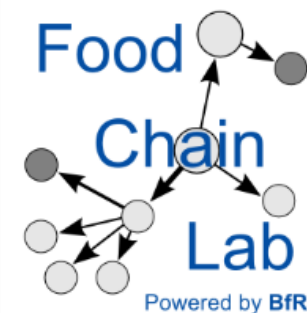
Available Templates/Data imports in FCL



- Start tracing
 - backward
 - forward
- Production tracing
 - Backward (auto generation)
 - Forward (auto generation)
- All-In-One Template
- Web-Edit
- Xml Busstop



Template Database



Tutorial 15

Data collection and import



Tracing Backward Template



Tutorial 16

Tracing Backward Template

Other approaches:

- Direct injection via
 - SQL
 - Alternative Excel sheets
 - Xml structures
 - ...



Another approach: COS – Bus stop - FoodChain-Lab



- COS collects (NRW-) Tracing Data
- COS is doing validation internally based on existing master data
- COS outputs data into newly defined xml-structure and sends it into a newly developed webservice
- FCL always looks into the webservice for new data
- FCL analyses data (semi)-automatically
- FCL prepares analysis reports and sends it back to the webservice
- COS fetches the report – management measures or further data collections will be initiated

FoodChain-Lab: Data Cleaning



SimSearch Options

Station: Name: 3 Zip: 1 Street: 3 Number: 1 City: 3 VATnumber: 0

Product: Station: 0 Denomination: 3 Item number: 0

Lot: Article: 0 LotNo: 1 BestBefore: 0 DateManufactor: 0

Delivery: Lot: 0 DeliveryDate: 0 Recipient: 0

OK Cancel

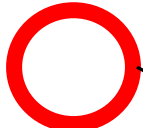
ID	Address	CompanyID	Type of busin...	VATnumber	Code	IsCase	Number Ca
1 30	Salad harvester Ltd Big Street Hometown		horticultural farm			yes	
2 31	Salad harevster Ltd Big Street Hometown		horticultural farm			yes	

Levenshtein distance

M	A	Y	O	N	N	A	I	S	E
=	=	=	=	=	-	=	=	=	=
M	A	Y	O	N		A	I	S	E

E	L	E	P	H	A	N	T
=	=	=	o	-	=	=	=
E	L	E	F		A	N	T

Works well for finding duplicates

 Bakery 1, **Sesame** Street 15, 12345 Crumbling Hills, Germany

Summer Cake

Luxury Hotel

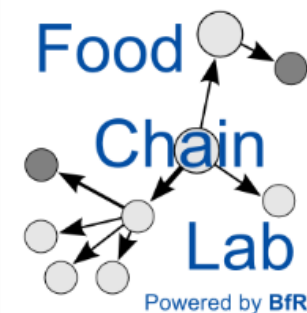
Hilltop 1

12234 Pompous Hills

Germany

Chocolate Cake

 Bakery 1, **See Same** Street 15, 12345 Crumbling Hills, Germany



Tutorial 17

Finding and removing duplicates



Thank you for your attention

Alexander Falenski

German Federal Institute for Risk Assessment
Max-Dohrn-Str. 8-10 • 10589 Berlin, GERMANY
Phone +49 30 - 184 12 - 0 • Fax +49 30 - 184 12 - 47 41
bfr@bfr.bund.de • www.bfr.bund.de/en