



**Exploiting the multifunctional potential of  
belowground biodiversity in horticultural farming**

## **Delivarable 1.4**


**Database on soil mapping and physico-  
chemical characteristics of experimental fields**

**Grant Agreement: 817946**

**Horizon 2020 Programme**

**Authors: Antonio Gerardo Pepe, Loredana Canfora, Stefano  
Mocali, Corrado Costa (CREA)**

**Version: 28-01-2021 (v1.4)**


	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	1/12
		Rev	4
		Date	28-01-2021

**Title:** Exploiting the multifunctional potential of belowground biodiversity in horticultural farming

**Project acronym:** EXCALIBUR

<b>Work Package</b>	WP1
<b>Deliverable Number</b>	1.4
<b>Due date</b>	30-11-2020
<b>Submission date</b>	31-01-2021
<b>Dissemination level</b>	Public
<b>Dealiverable lead</b>	CREA
<b>Authors</b>	Antonio Gerardo Pepe, Corrado Costa, Loredana Canfora, Stefano Mocali
<b>Abstract</b>	The aim of this document is to describe the database structure. A database including all the project data related to soil features and biodiversity will be provided and shared among all the participants. To ensure interoperability of data across the project, the FAIR principles will be applied, supporting data reuse ( <a href="http://www.dtls.nl/fair-data/">http://www.dtls.nl/fair-data/</a> ). The aim of this deliverable is to describe the database structure paying particular attention to metadata provided be soil mapping and physico-chemical characteristics of experimental field.
<b>Keywords</b>	Database




	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	2/12
		Rev	4
		Date	28-01-2021

## List of participants

Partner	Acronym	Country
CONSIGLIO PER LA RICERCA IN AGRICOLTURA E L'ANALISI DELL'ECONOMIA AGRARIA	CREA	Italy
INSTYTUT OGRODNICTWA	INHORT	Poland
C.R.P.V. CENTRO RICERCHE PRODUZIONI VEGETALI SOCIETA COOPERATIVA A RESPONSABILITA LIMITATA	C.R.P.V	Italy
NATURAL HISTORY MUSEUM	NHM	United Kingdom
NIAB EMR	NIAB EMR	United Kingdom
KMETIJSKI INSTITUT SLOVENIJE - AGRICULTURAL INSTITUTE OF SLOVENIA	KIS	Slovenia
UNIVERSITA DEGLI STUDI DI TORINO	UNITO	Italy
KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN - KNAW	KNAW-NIOO	Netherlands
KOBENHAVNS UNIVERSITET	UCPH	Denmark
TECHNISCHE UNIVERSITAET GRAZ	TU GRAZ	Austria
INOCULUMPLUS SAS	INplus	France
UNIVERSIDAD DE GRANADA	UGR	Spain
INTERMAG SPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA	Intermag	Poland
NSF EURO CONSULTANTS	NSF EC	Belgium
KOMPETENZZENTRUM OBSTBAU BODENSEE	KOB	Germany
FORDERGEMEINSCHAFT OKOLOGISCHER OBSTBAU EV (FOKO)	FÖKO	Germany



	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	3/12
		Rev	4
		Date	28-01-2021

## Version history

Version	Date	Author(s)	Description
1.0	24-11-2020	Antonio G. Pepe	Initial draft version, tables relation schema e and tables structure and new proposal of key coding
1.1	26-11-2020	Corrado Costa, Antonio G. Pepe	Second draft version
1.2	27-11-2020	Loredana Canfora. Antonio G. Pepe	Soil biological parameters review
1.3	27-01-2021	Loredana Canfora. Antonio G. Pepe, Stefano Mocali, Corrado Costa	Trial and Sample Codes review
1.4	28-01-2021	Loredana Canfora. Antonio G. Pepe	Soil metadata review






Page	4/12
Rev	4
Date	28-01-2021

## TABLE OF CONTENTS

List of participants .....	2
Version history.....	3
TABLE OF CONTENTS .....	4
EXECUTIVE SUMMARY.....	5
1. DATABASE LOGIC.....	5
1.1 Trial Code.....	6
1.2 Sample Code.....	6
2. Database structure .....	7
2.1 Trials .....	7
2.2. Secondary Tables.....	9



	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	5/12
		Rev	4
		Date	28-01-2021

## EXECUTIVE SUMMARY

This document describes the database collecting in a standard mode all the project data related to soil features and biodiversity provided, in .csv or .xlsx format, by partners.

Such a vast amount of data will be collected into a project database managed by CREA. Excalibur will generate data of different nature, including high-throughput sequencing data of bacteria, fungi, oomycetes, nematodes and microbial eukaryotes, transcriptomic (WP1, WP4, WP5), data of mesofauna, arthropods and earthworms biodiversity (WP1, WP4), metabolic profile of microbial bioinoculants (WP2), soil mineralogical and chemical-physical parameters (WP1, WP4, WP5), bioassay data of various formulations of bioinoculants and bioeffector's performance (WP2), field measurements of bioindicators of the soil and plant health status from experimental and on-farm trials (WP3, WP4, WP5), biomarkers and molecules driving plant-soil-microorganism interactions (WP5), detailed description of the current and innovative strategy designed for each case study including agronomical (crop yield and quality), economical (input and product prices, average farm size, etc.), and environmental (soil quality, functions and biodiversity, tracking inocula persistence in soil) data (WP3, WP4), personal data collected from stakeholder expectations (WP1, WP2 and WP6), georeferenced data as input of the DSSs (WP3, WP5), pedo-climatic and weather data (all WPs).

All the provided data will be checked and uploaded through Python scripts in an Azure SQL RDBMS on the Microsoft cloud managed by CREA, and it will be possible to connect to the resource with different technologies through dedicated connectors (Python, R, C#, etc.), useful for a model or DSS developers. Also, periodic reports will be provided for all users in the project intranet.

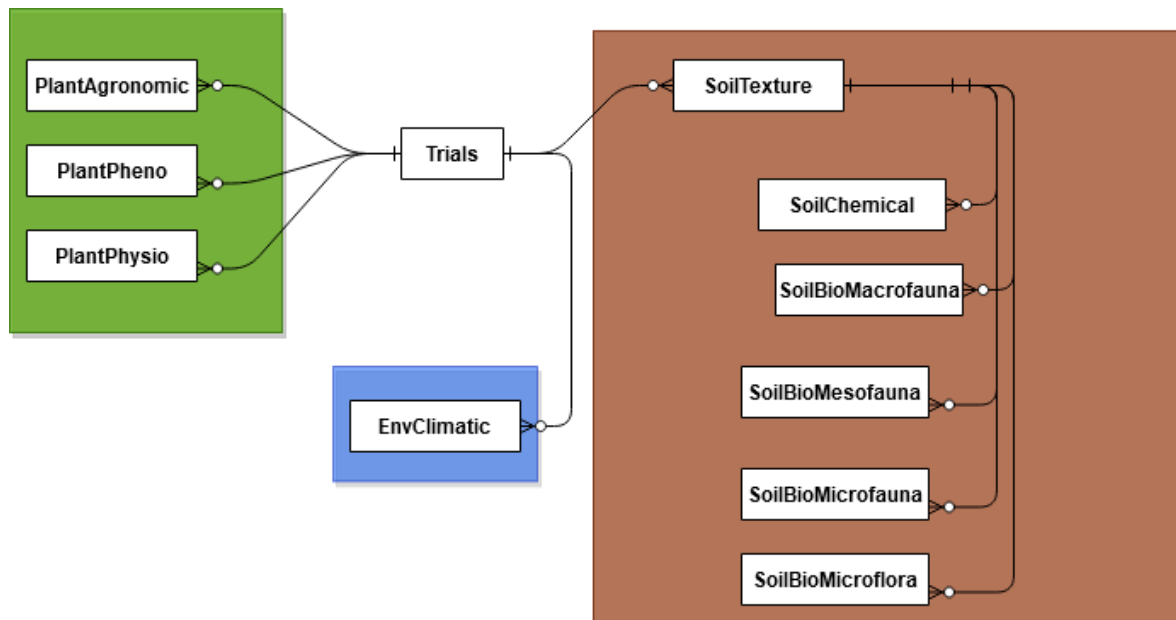
## 1. DATABASE LOGIC

As described in 'D7.4 Data Management Plan' the data each partner provides, may refer to three categories: soil, plant and environment. The database structure has been designed to collect all relevant data from WPs 1-5. This database is made up of the main table (Trials) and secondary tables, one for each measured and/or modelled variable (e.g. SOC, N, biological parameters etc.) collected for all trials over time. This database allows for cross-checking the consistency of the same type of data coming from different trials. Each trial is identifiable through a unique key code (see par. 1.1 Trial Code) used as a reference for all categories of data that will be provided by each partner (see fig. 1). All samples will have a **SampleCode** based on the **TrialCode**. We have introduced a code system for the codification of each measured parameters/variables for tracking the samples delivered to other partners and for monitoring the samples dedicated to the further analysis (soil, plant).

Relational and critical integrity is managed at the database level.



**Fig. 1. This figure displays the tables content and the relationships among the different tables.**



### 1.1 Trial Code

The Trial Code (TrialCode) is of pivotal importance and discriminately identify a trial. The TrialCode is also useful to define the *sample code* referring to the different samples collected in a trial (SampleCode) for the downstream analyses (chemical, physical biological, physiological ect.).

The structure of the TrialCode proposed by CREA and agreed by the Executive Committee is **CMC-PPP**, where:


- **CMC** = *Crop* (A, apple; S, strawberry; T, tomato) and *Management Method* (O, organic; I, integrated), *Condition* (L, lab; G, greenhouse; F,)
- **PPP** = *partner 3 letter code* (see PartnerCode in Tab. 3 Partners details)

Each partner will take care to provide the TrialCode following the example showed in tab. 1.1.

### 1.2 Sample Code

The samples collected by each trial will be distinguished by a unique code 'SampleNumber' (001 – 999) for tracking the samples delivered to other partners. The core referred to parameters will allow distinguishing the analyses to be carried out (ParamsCode). This is a code made of eight numerical characters.

Finally, we have a SampleCode made by TrialCode+SampleNumber+ParamsCode that must be used by each partner. Following, the SampleCode:

	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	7/12
		Rev	4
		Date	28-01-2021

**CMC-PPP-NNN-(1-8)** where

- **CMC- PPP = TrialCode**
- **NNN = unique Sample Number, each partner is responsible for keeping track of the unique numbering per partner.**
- **1-8 = Parameters to be monitored 1 letter code or 0 (zero) if not monitored, in order**
  - 1 = Physical (Soil)
  - 2 = Chemical (Soil)
  - 3 = Biological (Soil)
  - 4 = Biological, QBS-ar (Soil)
  - 5 = Biological, earthworms (Soil)
  - 6 = Agronomic (Plant)
  - 7 = Phenological (Plant)
  - 8 = Physiological (Plant)

Then, each sample should include the following information in SampleInfo:

**D-S-R-(GPS), where:**

- **D = Date** (YYYYMMDD)
- **S = Soil type** (P (profile); B (bulk); R, rhizosphere)
- **R = Replicate** (1-5)
- **GPS = GPS coordinates** from latitude and longitude of the sample

## 2. Database structure

The database is made of secondary table, one for each measured variable.

### 2.1 Trials

The Table 'Trials' is the main and first element describing the trial and the TCode (Crop Code, Management method Partner and Environmental Geographic information's like Location and coordinates 'Latitude and Longitude').





**Tab. 1. Trial data Metadata**


Property	Abbreviation used in DATA	Type	Unit used in DATA and in Auxiliary DATA	Type of sample used	Method
Crop code	CropCode	text	[adimentional]	A,S,T = Apple, Strawberry, Tomatoes	[-]
Management Method	ManMethod	text	[adimentional]	O,I,P,U = Organic, Integrated, Protected, Untreated	[-]
Condition	Condition	text	[adimentional]	L, G, F = Lab, Greenhouse; Field	[-]
Partner code	PartnerCode	text	[adimentional]	[-]	[-]
Trial Latitude	Lat	numeric	[decimal degrees]	Coordinate system	WGS 84
Trial Longitude	Lon	numeric	[decimal degrees]	Coordinate system	WGS 84
Trial code	TrialCode	text	[adimentional]	CMC-PPP	CMC = CropCode+ManMethod+Condition; PPP = partner 3 letter code

Each partner will use the TrialCode as showed in the example below:

**Tab.2 – Trials details example**

CropCode	ManMethod	Condition	PartnerCode	Location	Latitude	Longitude	TrialCode
A	O	F	INH	Skierniewice -Podlesna (PL)	51.9547	20.1583	<b>AOF-INH</b>
A	I	F	INH	Skierniewice -Podlesna (PL)	51.9547	20.1583	<b>AIF-INH</b>
S	I	G	INH	Skierniewice -Podlesna (PL)	51.9547	20.1583	<b>SIG-INH</b>
A	O	F	CRP	Forli' (FC)	44.22123	12.04269	<b>AOF-CRP</b>
A	I	F	CRP	Cesena (FC)	44.13876	12.23912	<b>AIF-CRP</b>
S	O	G	CRP	Cesena (FC)	44.13876	12.23912	<b>SOG-CRP</b>
S	I	G	CRP	Cesena (FC)	44.13876	12.23912	<b>SIG-CRP</b>
T	O	F	CRP	Cesena (FC)	44.13876	12.23912	<b>TOF-CRP</b>
T	I	F	CRP	Cesena (FC)	44.13876	12.23912	<b>TIF-CRP</b>



	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	9/12
		Rev	4
		Date	28-01-2021

**Tab. 3 – Partner details**

PartnerCode	PartnerAcronym	PartnerName
CRE	CREA	CONSIGLIO PER LA RICERCA IN AGRICOLTURA E L'ANALISI DELL'ECONOMIA AGRARIA
INH	INHORT	INSTYTUT OGRODNICTWA
CRP	CRPV	CENTRO RICERCHE PRODUZIONI VEGETALI SOC. COOP. A RESPONSABILITA LIMITATA
NHM	NHM	NATURAL HISTORY MUSEUM
NIA	NIAB EMR	NIAB EMR
KIS	KIS	KMETIJSKI INSTITUT SLOVENIJE - AGRICULTURAL INSTITUTE OF SLOVENIA
UNI	UNITO	UNIVERSITA DEGLI STUDI DI TORINO
KNA	KNAW-NIOO	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN - KNAW
UCP	UCPH	KOBENHAVNS UNIVERSITET
TUG	TU GRAZ	TECHNISCHE UNIVERSITAET GRAZ
INP	INplus	INOCULUMPLUS SAS
UGR	UGR	UNIVERSIDAD DE GRANADA
INT	Intermag	INTERMAG SPOLKA Z OGRANICZONA ODPOWIEDZIALNOSCIA
NSF	NSF EC	NSF EURO CONSULTANTS
KOB	KOB	KOMPETENZENTRUM OBSTBAU BODENSEE
FOK	FOKO	FORDERGEMEINSCHAFT OKOLOGISCHER OBSTBAU EV


## 2.2. Secondary Tables

Two examples are displayed to explain how the monitored variables will be organized in the secondary table, providing the metadata. The secondary tables will send to each partner to implement the different tables referred to each metadata.

**Tab.4 Soil texture metadata**

Property	Abbreviation used in DATA	Type	Unit used in DATA and in Auxiliary DATA	Type of sample used	Method
Sampling Date Code	Date	date	[YYYY-MM-DD]	[-]	[-]
Soil Type	SoilType	text	[adimentional]	Soil type (P (profile); B (bulk); R, rhizosphere)	[-]
Sample Replicate code	Replicate	numeric	[adimentional]	[-]	[-]
Layer	Layer	text	[adimentional]	[-]	[-]



	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	10/12
		Rev	4
		Date	28-01-2021


Property	Abbreviation used in DATA	Type	Unit used in DATA and in Auxiliary DATA	Type of sample used	Method
Silt	Silt	numeric	[%]	[-]	[-]
Clay	Clay	numeric	[%]	[-]	[-]
Sand	Sand	numeric	[%]	[-]	[-]
Trial code	TrialCode	text	[adimentional]	CMC-PPP	CMC = CropCode+ManMethod+Condition; PPP = partner 3 letter code
Unique Sample Number	SampleNumber	text	[adimentional]	NNN	each partner is responsible for keeping track of the unique numbering per partner
Sample Latitude	Lat	numeric	[decimal degrees]	Coordinate system	WGS 84
Sample Longitude	Lon	numeric	[decimal degrees]	Coordinate system	WGS 84
Parameters to be monitored	ParamsCode	text	[adimentional]	12345678	1 letter code or 0 (zero) if not monitored
Sample Code	SampleCode	text	[adimentional]	CMC-PPP-NNN-(1-8)	TrialCode+SampleNumber+ParamsCode
Sample Information Code	SampleInfo	text	[adimentional]	D-S-R-(GPS)	Date+SoilType+Replicate+(Lat Lon)

Following a 'Soil texture data example' including preliminary data uploaded on Excalibur intranet:

**Tab.5 Soil texture data example**

Date	Soil Type	Replicate	layer	Silt	Clay	Sand	Trialcode	Sample Number	Latitude	Longitude	Params Code	SampleCode	SampleInfo
20191112	B	1	A	14	0	86	AOF-INH	6	51.954799	20.158310	12000000	AOF-INH-006-12000000	20191112-B-1-(51.954799,20.15831)
20191112	B	1	Bg1	18	4	78	AOF-INH	5	51.954799	20.158311	12000000	AOF-INH-005-12000000	20191112-B-1-(51.954799,20.15831)
20191112	B	1	Bg2	33	7	60	AOF-INH	4	51.954798	20.158311	12000000	AOF-INH-004-12000000	20191112-B-1-(51.954798,20.15831)
20191112	B	1	A	11	1	88	AIF-INH	1	51.954798	20.158312	12000000	AIF-INH-001-12000000	20191112-B-1-(51.954798,20.15831)
20191112	B	1	Bg1	32	2	66	AIF-INH	3	51.954797	20.158312	12000000	AIF-INH-003-12000000	20191112-B-1-(51.954797,20.15831)
20191112	B	1	Bg2	20	3	77	AIF-INH	2	51.954797	20.158313	12000000	AIF-INH-002-12000000	20191112-B-1-(51.954797,20.15831)



	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields		Page	11/12
			Rev	4
			Date	28-01-2021


20191112	B	1	Ap	20	1	79	SIG-INH	9	51.954797	20.158314	12000000	SIG-INH-009-12000000	20191112-B-1-(51.954797,20.158314)
20191112	B	1	3	20	2	78	SIG-INH	7	51.954796	20.158315	12000000	SIG-INH-007-12000000	20191112-B-1-(51.954796,20.158315)
20191112	B	1	Bg	19	2	79	SIG-INH	8	51.954796	20.158316	12000000	SIG-INH-008-12000000	20191112-B-1-(51.954796,20.158316)

**Tab.6 Soil chemical metadata**

Property	Abbreviation used in DATA	Type	Unit used in DATA and in Auxiliary DATA	Type of sample used	Method
Sampling Date Code	Date	date	[YYYY-MM-DD]	[-]	[-]
Soil Type	SoilType	text	[adimensional]	Soil type (P (profile); B (bulk); R, rhizosphere)	[-]
Sample Replicate code	Replicate	numeric	[adimensional]	[-]	[-]
Layer	Layer	text	[adimensional]	[-]	[-]
pH	pH	numeric	[adimensional]	[-]	Extraction with 0,01 M CaCl <sub>2</sub> (VDLUF 1991); Extraction with 0,1 N KCl or H <sub>2</sub> O
NaCl/l	NaCl/l	numeric	[adimensional]	[-]	[-]
N-NO <sub>3</sub>	N-NO <sub>3</sub>	numeric	[adimensional]	[-]	[-]
P	P	numeric	[adimensional]	[-]	Aqua regia, measurement with ICP-OES
K	K	numeric	[adimensional]	[-]	Aqua regia, measurement with ICP-OES
Ca	Ca	numeric	[adimensional]	[-]	Aqua regia, measurement with ICP-OES
Mg	Mg	numeric	[adimensional]	[-]	Aqua regia, measurement with ICP-OES
Trial code	TrialCode	text	[adimensional]	CMC-PPP	CMC = CropCode+ManMethod+Condition; PPP = partner 3 letter code
Unique Sample Number	SampleNumber	text	[adimensional]	NNN	each partner is responsible for keeping track of the unique numbering per partner
Sample Latitude	Lat	numeric	[decimal degrees]	Coordinate system	WGS 84
Sample Longitude	Lon	numeric	[decimal degrees]	Coordinate system	WGS 84
Parameters to be monitored	ParamsCode	text	[adimensional]	12345678	Parameters to be monitored 1 letter code or 0 (zero) if not monitored
Sample Code	SampleCode	text	[adimensional]	CMC-PPP-NNN-(1-8)	TrialCode+SampleNumber+ParamsCode



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 817946.*

	D1.4 - Database on soil mapping and physico-chemical characteristics of experimental fields	Page	12/12
		Rev	4
		Date	28-01-2021

Sample Information Code	SampleInfo	text	[adimentional]	D-S-R-(GPS)	Date+SoilType+Replicate+(Lat Lon)
-------------------------	------------	------	----------------	-------------	-----------------------------------

**Tab.6 Soil chemical data example**

Date	Soil Type	Replicate	Layer	pH	Na Cl/l	N-NO3	P	K	C	Mg	TrialCode	TrialNumber	Latitude	Longitude	ParamsCode	SampleCode	SampleInfo
2019 1112	B	1	A	6 .5	0.2 7	57	4 7. 3	2 4 8	5 5 1	6 6	AOF- INH- 006	6	51.95 4799	20.15 8310	12000 000	AOF-INH- 006- 12000000	20191112-B-1- (51.954799,20.1583 1)
2019 1112	B	1	Bg 1	6 .2	0.1 1	17	3 3. 3	5 5 5	4 9 5	4 9	AOF- INH- 005	5	51.95 4799	20.15 8311	12000 000	AOF-INH- 005- 12000000	20191112-B-1- (51.954799,20.1583 11)
2019 1112	B	1	Bg 2	6 .5	0.1 1	17	5 1. 1	1 5	6 2 8	5 9	AOF- INH- 004	4	51.95 4798	20.15 8311	12000 000	AOF-INH- 004- 12000000	20191112-B-1- (51.954798,20.1583 11)
2019 1112	B	1	A	5 .4	0.3 3	33	2 8. 1	1 1 9	3 4 7	4 5	AIF- INH- 001	1	51.95 4798	20.15 8312	12000 000	AIF-INH-001- 12000000	20191112-B-1- (51.954798,20.1583 12)
2019 1112	B	1	Bg 1	5 .9	0.1 1	18	2 2. 7	3 8	3 4 7	4 5	AIF- INH- 003	3	51.95 4797	20.15 8312	12000 000	AIF-INH-003- 12000000	20191112-B-1- (51.954797,20.1583 12)
2019 1112	B	1	Bg 2	6 .3	0.1 2	14	3 2. 1	9 8	4 7 6	4 9	AIF- INH- 002	2	51.95 4797	20.15 8313	12000 000	AIF-INH-002- 12000000	20191112-B-1- (51.954797,20.1583 13)
2019 1112	B	1	Ap	5 .6	0.1 8	35	4 1. 8	5 7 4	3 8 4	3 9	SIG- INH- 009	9	51.95 4797	20.15 8314	12000 000	SIG-INH-009- 12000000	20191112-B-1- (51.954797,20.1583 14)
2019 1112	B	1	3	5 .7	0.2 2	40	4 1. 1	6 2 2	4 2 4	3 9	SIG- INH- 007	7	51.95 4796	20.15 8315	12000 000	SIG-INH-007- 12000000	20191112-B-1- (51.954796,20.1583 15)
2019 1112	B	1	Bg 5	5 .5	0.1 9	30	2 0	5 4	5 7 5	4 6	SIG- INH- 008	8	51.95 4796	20.15 8316	12000 000	SIG-INH-008- 12000000	20191112-B-1- (51.954796,20.1583 16)

