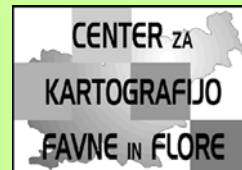


Experience from Slovenia

Mladen Kotarac

Centre for Cartography of Fauna and Flora





Aflatunia
Allium

Allium
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

Anthracis
Anthracis

**The collection:
150.000 sheets**



**The accession:
a couple of thousand sheets/year**



The workplace



1 taxa name

Herbarium sheets are ordered by taxa,
Each taxa having one or more piles of sheets.

Databasing – The Problem

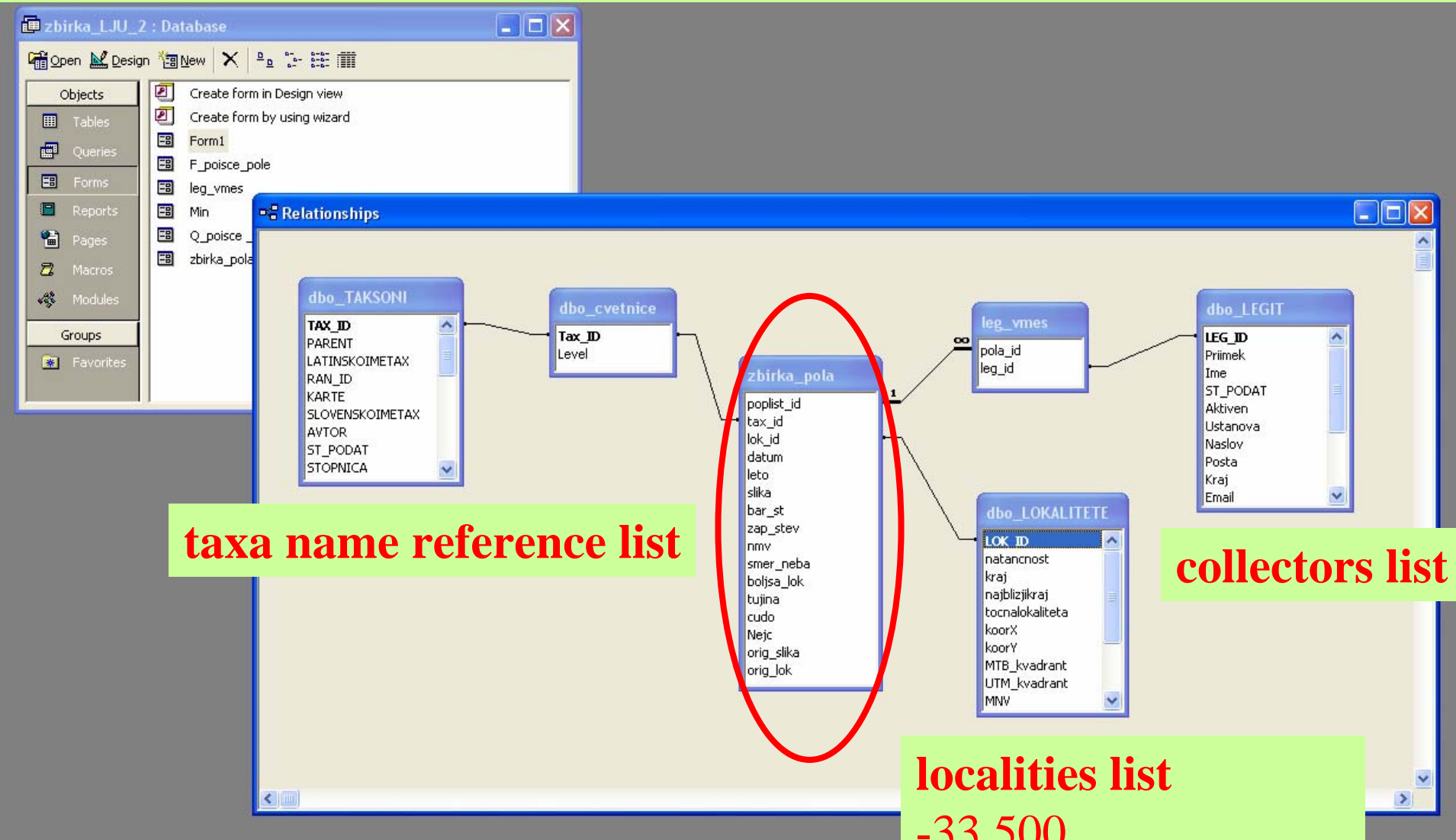
-what is the object of the database?

SPACE or **TAXA (SPECIMENS)**

-most modern herbaria are organized by taxa, therefore neglecting the space, it is the only feasible way for the human mind

-even more, tearing apart the space relation (usually few/many sheets from the same locality, collector and date) happens at the accession time

Inputs available/provided



localities list

-33.500

-9 levels of accuracy

Digitisation process

1. Imaging
2. Image postprocessing, barcode and taxa name assignment
3. Geocoding

Imaging

imaging of all herbarium sheets

work done exclusively by students working in teams of two

barcode labels attached to the outside of the sheet protective cover
and on the sheet label

images taken by Sony DSC 717 camera, 5Mpixel

sequential order of barcodes AND images STRICTLY enforced

**taxon name of the pile and respective first/last barcode numbers
noted**

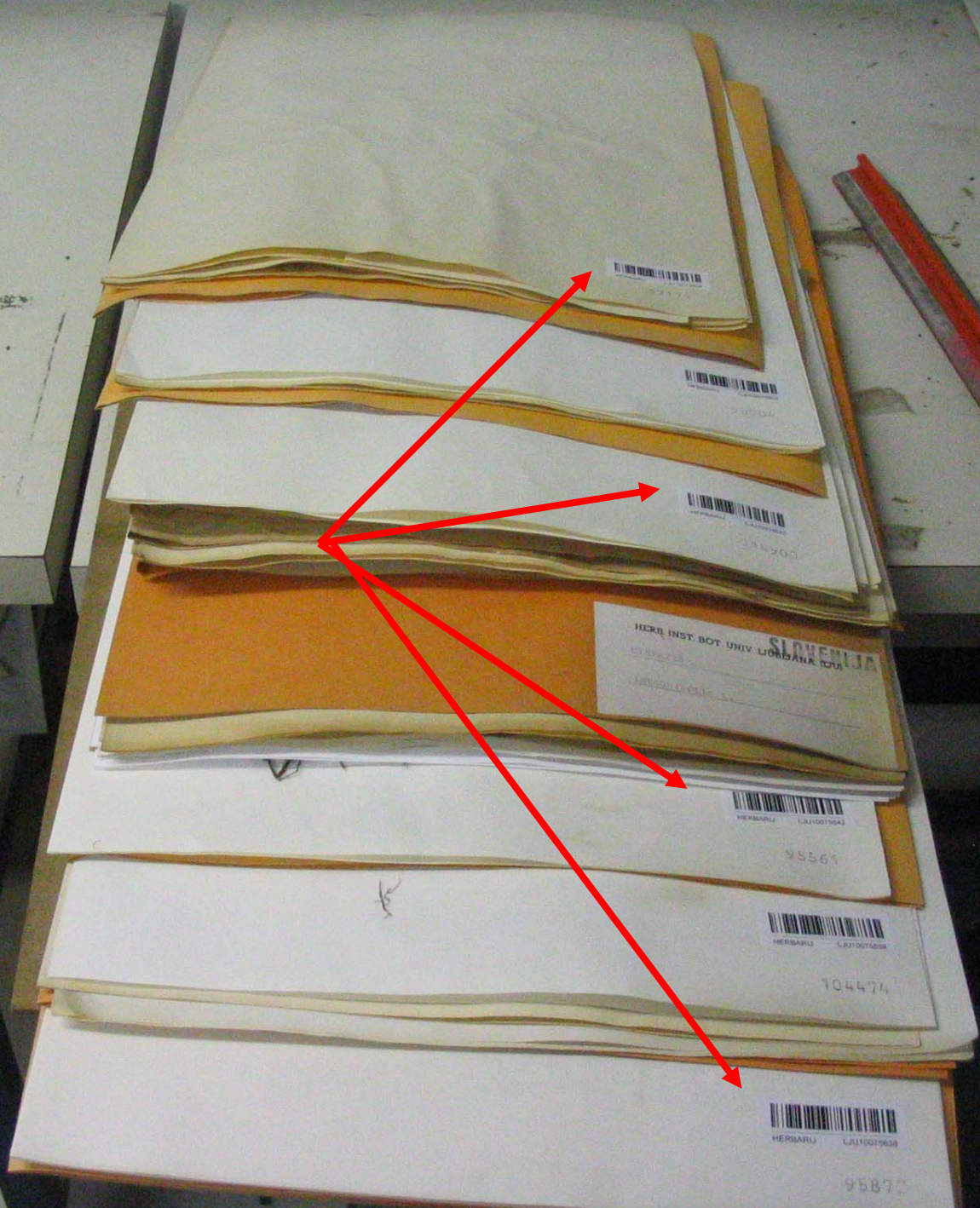


Dr. 0572-02
DNE 00174 (11/11/11)
2. 11/11/11 (11/11/11)

SuperPhoto 4271
SuperPhoto 4271
SuperPhoto 4271
SuperPhoto 4271
SuperPhoto 4271



PHOTOGRAPHY
PHOTOGRAPHY



HERBARIUM LJI1007963

HERBARIUM LJI1007964

HERBARIUM LJI1007965

HERBARIUM LJI1007966
SIKENDAR
JAWABANGSA

HERBARIUM LJI1007967
95567

HERBARIUM LJI1007968
104474

HERBARIUM LJI1007969
95870



HERB. INST. BOT. UNIV. LJUBLJANA (LJU) № 84819



HERBARIJ LJU10075583

ГЕРБАРИЙ ФЛОРЫ СССР
ПОДАРОМЪЮ БОТАНИЧЕСКОМУ ИНСТИТУТУ ИМ. В. И.
ИВАНОВА ИМПУ СССР

5090. *Claytonia joanneana* Roem. et Schult.

Cl. n. sp. n. 4564.

Восточная Сибирь, Хангай-Чиндеевское нагорье, Даурский хребет, Быстринский талов. вершина р. Быстрой (приток р. Бурага), на сырых ламповых лужайках. Собр. А. и М. Максимовы. Апрель. М. Максимовы. — *Siberia orientalis, montes Chantai-Goloi, jugum Dauricum, in monte Bystrinsky goletz, ad fontes fl. Bystrajae confluens P. Burkal, in pratulis humidis muscosis*. Leg. A. et M. Maximova. Det. M. Maximova. 1967 VII 3.

Imaging - results

-gigabytes of raw images

-handwritten notes

Imaging – postprocessing

- batch sequential renaming (IrfanView)
- batch autoleveling the colors (Photoshop)
- batch 90° rotate (Photoshop)
- batch crop of the lower 40% extent of the picture (Photoshop)

Barcode and taxa name assignment

4 zaporedje črtnih kod po taksonih: izpolnjuje podajalec [ime taksona, črtna koda prve pole, črtna koda zadnje pole, opombe]
 datum: 14.4. 2004 zap. št.: 39

ime taksona	št. črtne kode prve pole	št. črtne kode zadnje pole	opombe
DANTHUS SYLVESTRIS var. SYLVESTRIS	18752	18762	
D.S. var. UNIFLORUS	18763	18773	
— " —	18774	18788	
D.S. var. BREVICALIX	18789	18789	
D. TERGESTINUS	18790	18798	
— " —	18799	18807	
— " —	18808	18814	
DIGITALIS GRANDIFLORA	18815	18822	
— " —	18823	18831	
— " —	18832	18848	
D. LAEVIGATA	18849	18854	
— " —	18855	18859	
D. PURPUREA	18860	18860	
DIGITARIA AEGYPTICA	18861	18862	
— " —	18863	18864	
D. ISCHAE MUM	18865	18869	
D. SANGUINALIS ssp. RECTINIFORMIS	18870	18881	
D.S. ssp. SANGUINALIS	18882	18905	
DIPHASIMUM ALPINUM	18906	18922	
D. COMPLANATUM	18923	18939	
— " —	18940	18949	
— " —	18950	18959	
— " —	18960	18979	
D. ISSLERI	18980	18981	
D. TRISTACHYUM	18982	18996	
— " —	18997	19005	
DIPLOTAXIS ERUCOIDES	19006	19007	
D. MURALIS	19008	19028	
D. TENUIFOLIA	19029	19044	
DIPSACUS LACINIATUS	19045	19045	
D. SYLVESTRIS	19046	19058	
DITTRICHIA VISCOSA	19059	19062	
DORONICUM AUSTRIACUM	19063	19091	
— " —	19092	19108	
D. COLUMNAE	19109	19109	
D. GLACIALE	19110	19126	
— " —	19127	19140	
D. GRANDIFLORUM	19141	19162	
D. HUNGARICUM	19163	19163	
D. PARDALIANCHES	19164	19164	
DORYCNIMUM GERMANICUM	19165	19187	

57 sheets of the same "taxa" in 4 different piles

Barcode and taxa name assignment

The screenshot shows a Microsoft Access window titled "zbirka_LJU_2 : Database". The main window displays a form with the following fields:

- začetna barkoda: 52763
- končna barkoda: 1 52766
- število pol: 4
- takson: Seseli libanotis
- Record: 50544

A dropdown menu for "takson" is open, showing a list of taxa with their corresponding counts:

Taxon	Count
Seseli libanotis (L.) Koch	0
Seseli libanotis ssp. eu-	0
Seseli libanotis ssp. liba	0
Seseli libanotis ssp. lioca	0
Seseli libanotis var. pyre	0
Seseli malyi A. Kern.	1
Seseli montanum L.	0
Seseli montanum ssp. tr(Rchb. f.) Arcar	0

Red numbers 1 and 2 are overlaid on the form, indicating the specific data points mentioned in the text below.

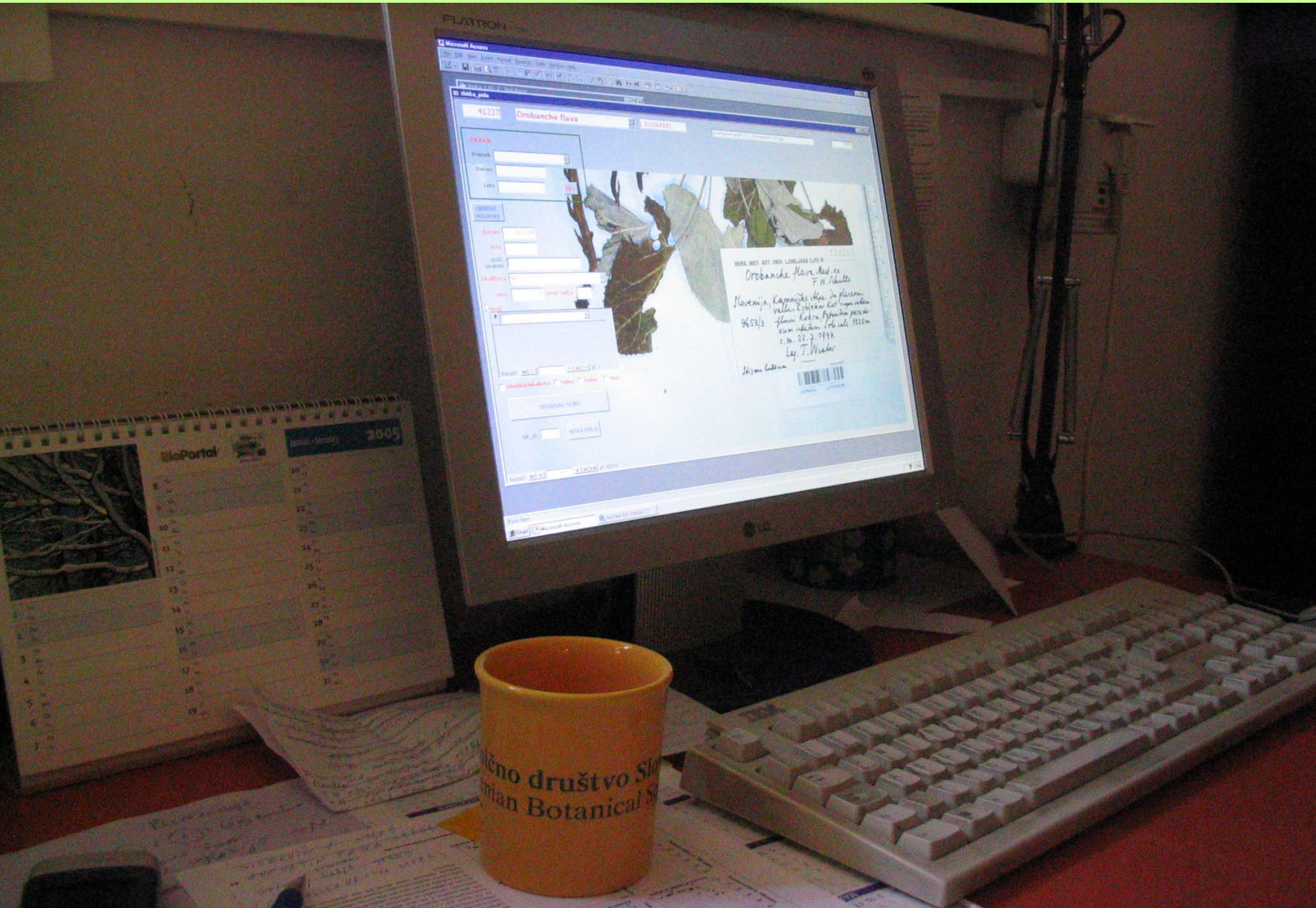
1. Barcode numbers get assigned to the respective sheets.
2. Sheets get taxon identity.

Barcode and taxa name assignment - results

one simple table with three fields:

image_file_name	barcode	taxonomic_name
....
0011740.jpg	LJU10018567	<i>Poa bulbosa</i>
0011741.jpg	LJU10018568	<i>Poa bulbosa</i>
0011742.jpg	LJU10018569	<i>Poa minor</i>
.....

Geocoding



Collector & Date OR Year entered

ISKANJE

Priimek

Datum

Leto

IŠČI

Q_poisce_pole

	datum	Priimek	Ime	Točna lokaliteta	MTB	N	leto	nmv	smer	boljša lok.
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		900		<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5				<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		950		<input type="checkbox"/>
TALE!		Naglič	Dušan	Paka pri Velenju (Velenje)	9656/2	4		440		<input type="checkbox"/>

NOBEDEN NI PRAVI!

Record: 1 of 18

OBNOVI PODATKE

datum

leto

poišči lokaliteto

lokaliteta

nmv smer neba

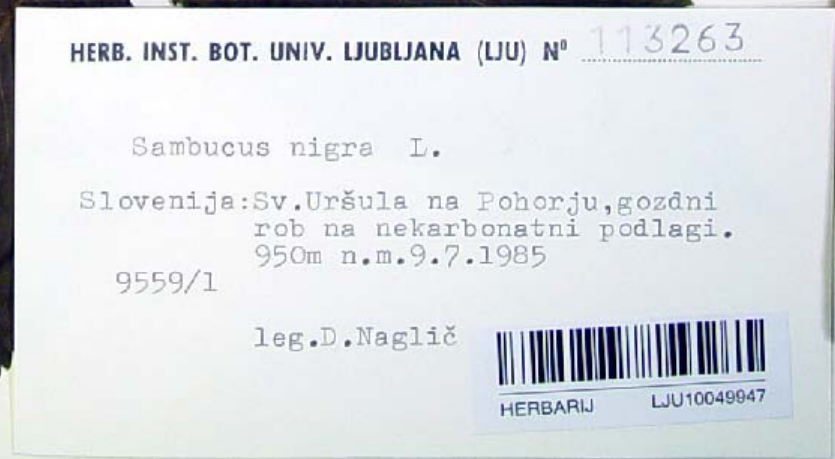
legit

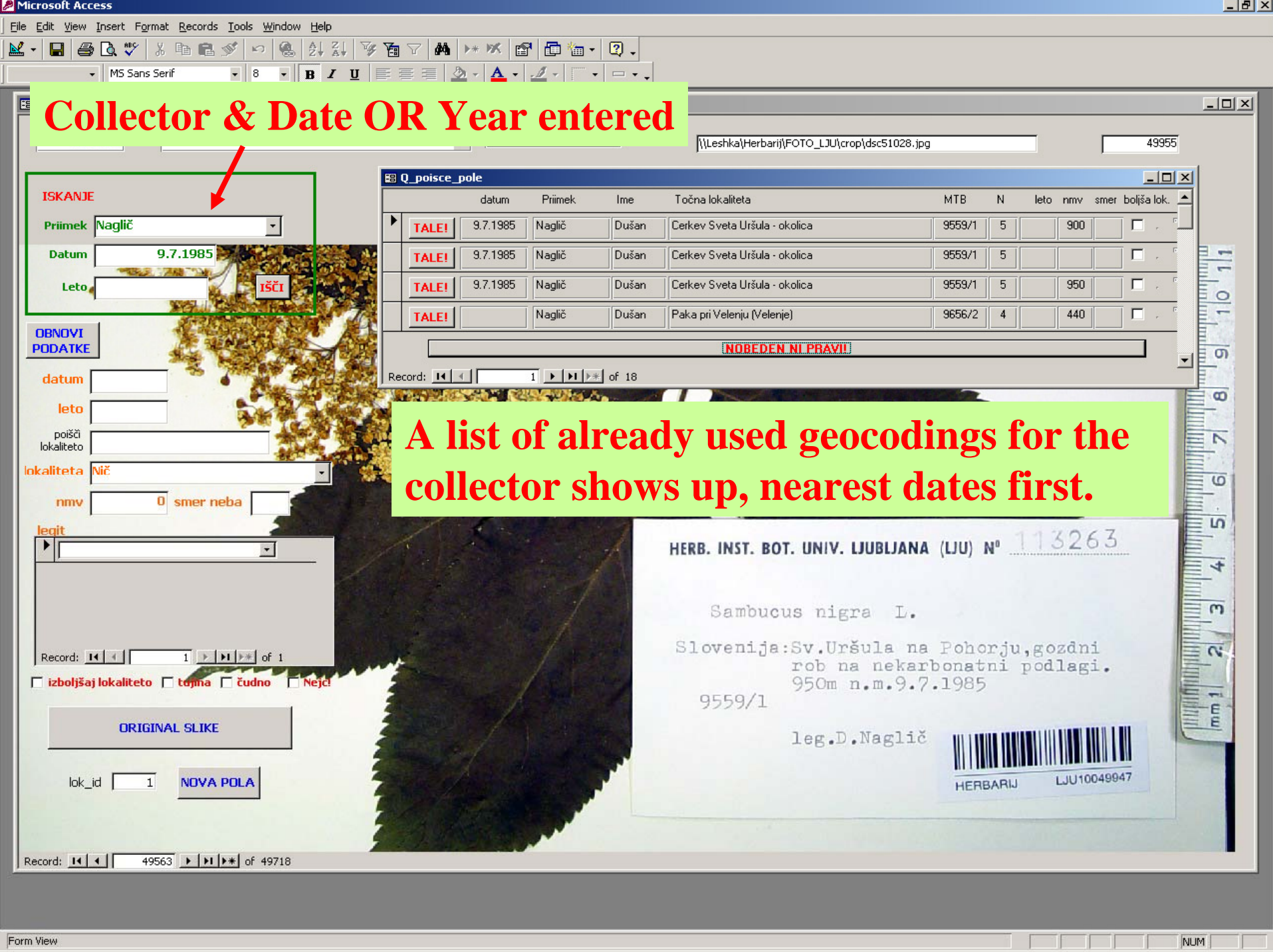
Record: 1 of 1

izboljšaj lokaliteto tojina čudno Nejc!

ORIGINAL SLIKE

lok_id **NOVA POLA**





Collector & Date OR Year entered

ISKANJE

Priimek

Datum

Leto

IŠČI

Q_poisce_pole

	datum	Priimek	Ime	Točna lokaliteta	MTB	N	leto	nmv	smer	boljša lok.
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		900		<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5				<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		950		<input type="checkbox"/>
TALE!		Naglič	Dušan	Paka pri Velenju (Velenje)	9656/2	4		440		<input type="checkbox"/>

Record: 1 of 18

A list of already used geocodings for the collector shows up, nearest dates first.

OBNOVI PODATKE

datum

leto

poišči lokaliteto

lokaliteta

nmv smer neba

legit

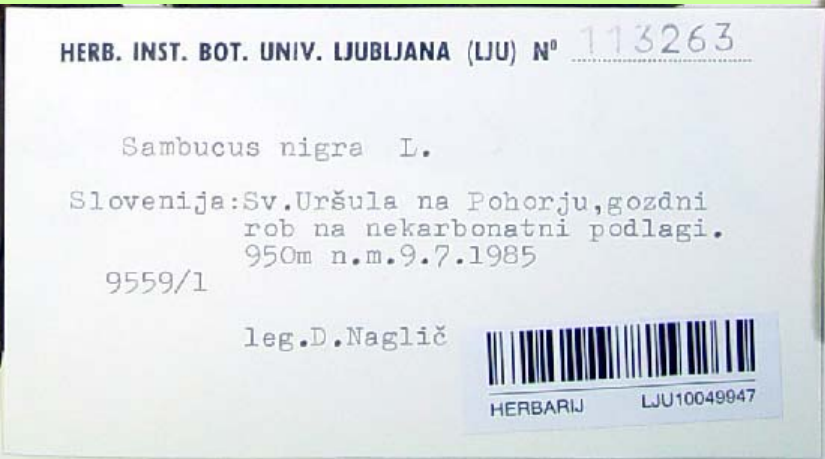
Record: 1 of 1

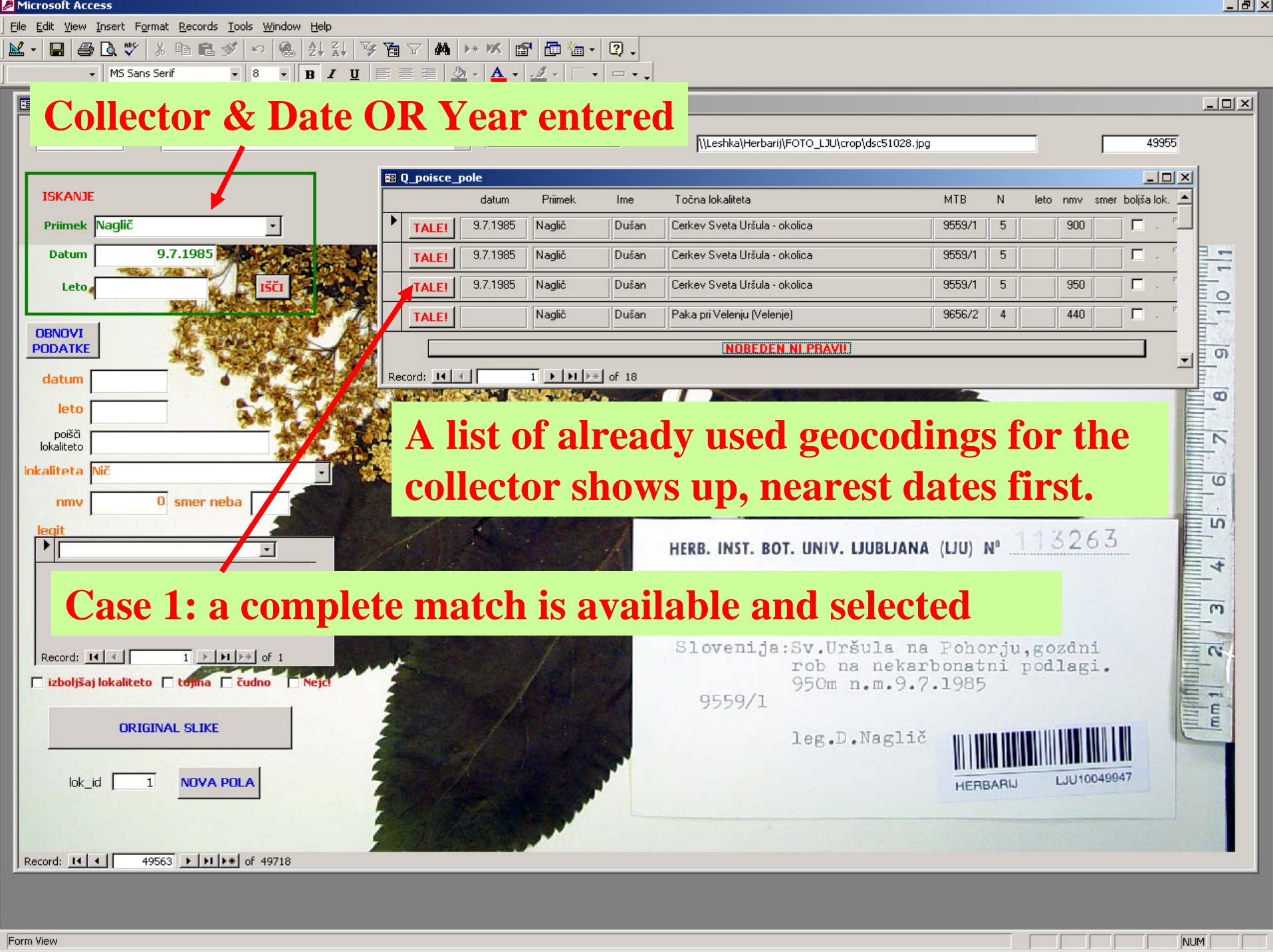
izboljšaj lokaliteto tojina čudno Nejc!

ORIGINAL SLIKE

lok_id **NOVA POLA**

Record: 49563 of 49718





Collector & Date OR Year entered

ISKANJE

Priimek

Datum

Leto

IŠČI

Q_poisce_pole

	datum	Priimek	Ime	Točna lokaliteta	MTB	N	leto	nmv	smer	boljša lok.
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		900		<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5				<input type="checkbox"/>
TALE!	9.7.1985	Naglič	Dušan	Cerkev Sveta Uršula - okolica	9559/1	5		950		<input type="checkbox"/>
TALE!		Naglič	Dušan	Paka pri Velenju (Velenje)	9656/2	4		440		<input type="checkbox"/>

Record: 1 of 18

A list of already used geocodings for the collector shows up, nearest dates first.

Case 1: a complete match is available and selected

OBNOVI PODATKE

datum

leto

poišči lokaliteto

lokaliteta

nmv smer neba

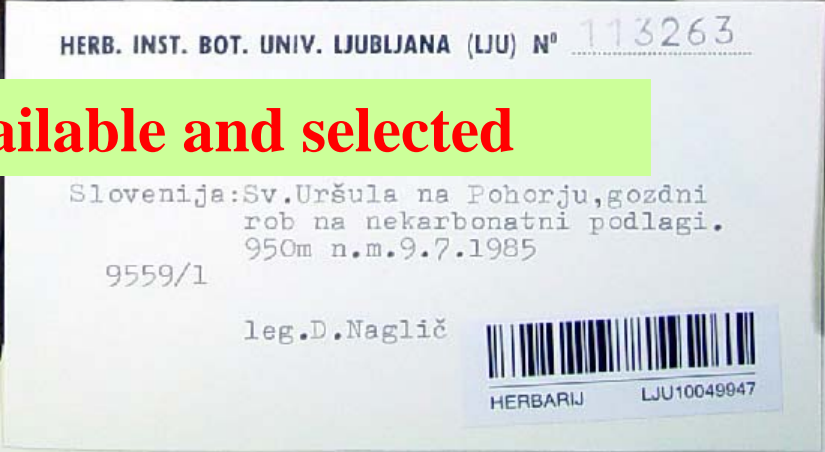
legit

Record: 1 of 1

izboljšaj lokaliteto tojina čudno Nejc!

ORIGINAL SLIKE

lok_id **NOVA POLA**



Record: 49563 of 49718





zbirka_LJU_2 : Database

Objec zbirka_pola

Open Design New

Tabl
Qu
For
Rep
Pag
Mac
Mok
Group
Fav

2916 Alchemilla glaucescens LJU10001508 1507

ISKANJE

Primek

Datum

Leto IŠČI

OBNOVI PODATKE

datum 21.7.1956

leto

poišči lokaliteto

lokaliteta Gora Mali Golak (1480)

nmv 1400 smer neba

legit

Martinčič

*

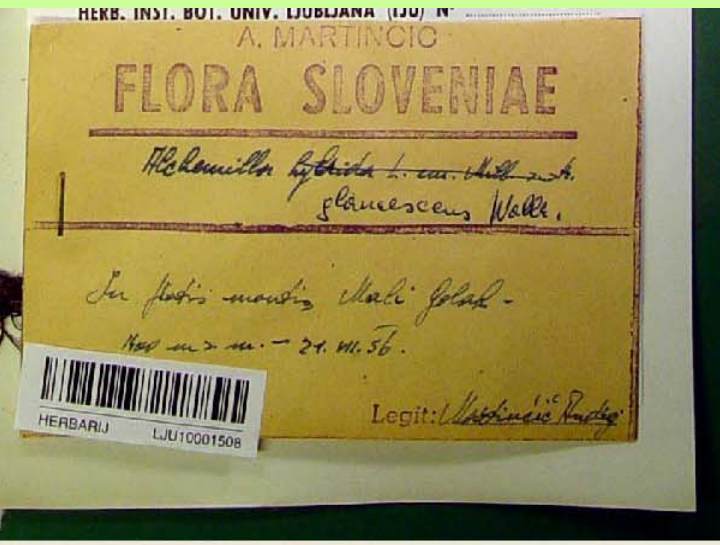
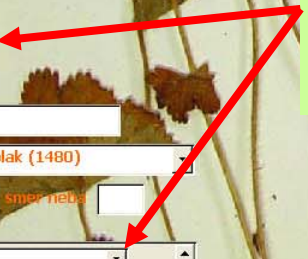
Record: 1 of 49718

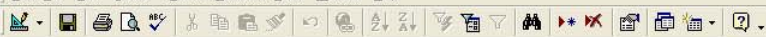
izboljšaj lokaliteto tujina čudno Nejl

ORIGINAL SLIKE

lok_id 15267 NOVA POLA

Case 2: no match was available
-already entered data are automatically transferred (collector & date OR year)





zbirka_LJU_2 : Database

Objec: zbirka_pola

Open Design New

Objec: zbirka_pola

Tabl: zbirka_pola

Qu: zbirka_pola

For: zbirka_pola

Rep: zbirka_pola

Pag: zbirka_pola

Mac: zbirka_pola

Mok: zbirka_pola

Group: zbirka_pola

Fav: zbirka_pola

2916 Alchemilla glaucescens LJU10001508

1507

ISKANJE

Primek

Datum

Leto IŠČI

**OBNOVI
PODATKE**

datum 21.7.1956

leto

poišči lokaliteto

lokaliteta Gora Mali Golak (1480)

nmv 1400 smer: neba

legit

▶ Martinčič

*

Record: 1 of 49718

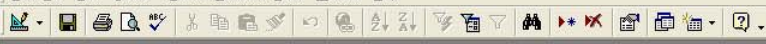
izboljšaj lokaliteto tujina čudno Nejl!

ORIGINAL SLIKE

lok_id 15267 NOVA POLA

Case 2: no match was available
-already entered data are automatically transferred (collector & date OR year)
-locality is entered OR searched





zbirka_LJU_2 : Database

Objec: zbirka_pola

2916 Alchemilla glaucescens LJU10001508 1507

ISKANJE

Priimek

Datum

Leto

OBNOVI PODATKE

datum 21.7.1956

leto

pošči lokaliteto

lokaliteta Gora Mali Golak (1480)

nmv 1400 smer neba

legit

Martinčič

*

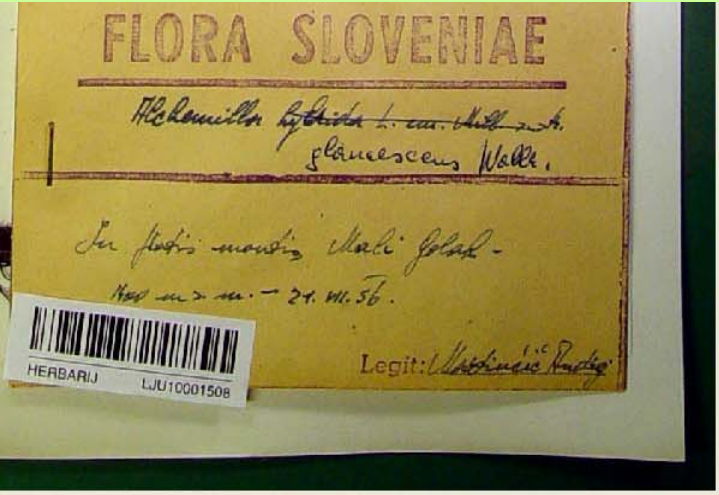
Record: 1 of 49718

izboljšaj lokaliteto tujina čudno Nejl

ORIGINAL SLIKE

lok_id 15267 NOVA POLA

Case 3: no geocoding is possible at the moment
-the sheet is left behind
-new locality is made in GIS using topographic maps and attached by an other geocoder



Geocoding - results

a (re)-established one to many relationship between space (& collector & date) on one side and specimens on the other side

simple relational database allowing retrieval/querying on all transcribed/normalized fields

The figures -workforce

2 students fulltime one year

160.000 images

1 geocoder fulltime one year

80.000 sheets

1 geocoder for new localities: 3 months

1 oversight: 1 month a year

images & database manipulation: 2 days per 10.000 sheets

The figures - quantity

Images production and barcode labeling:

aprox. 800 images/day

Geocoded sheets:

aprox. 380 sheets/day

New localities and geocoding:

aprox. 90 localities/day

The figures - money

Imaging costs – labour: 14.600 €

0,09 €per image/sheet

Geocoding costs – labour: 35.000 €

0,45 €per image/sheet

Oversight and manipulation:

0,06 €per image/sheet

equipment and facilities taken for granted

The mistakes

-The camera broke 4 times within first 75.000 sheets, first time after 18.000 sheets, no spare equipment was prepared. Approximately 3 months were waisted waiting for the repair!!!

-Regular control of the students and transfer of the images was neglected immediately after the repair breaks resulting in higher number of mistakes such as:

- missing or double images**
- missing or messed-up barcode labels**
- wrong or missing sequence/taxa notes**



ToDo list

- prepare or select/use an interface for the digitisation of the new accessions
- prepare a Web access interface for the database

Lessons learned

- atomize the work as much as needed/possible**
- assign the different work phases to DIFFERENT people according to the needed skills, costs and possible location of the work**
- maintain quality control in all phases**
- constantly think about lowering the costs without sacrificing the quality/usage**

Hopefully someday you will have the work done. 😊