

ARTYDOC, a new digital archive of meteorological documentation

**Marc Prohom
Mònica Herrero
Albert Aparicio**

marc.prohom@gencat.cat / @MProhom



**Servei Meteorològic
de Catalunya**

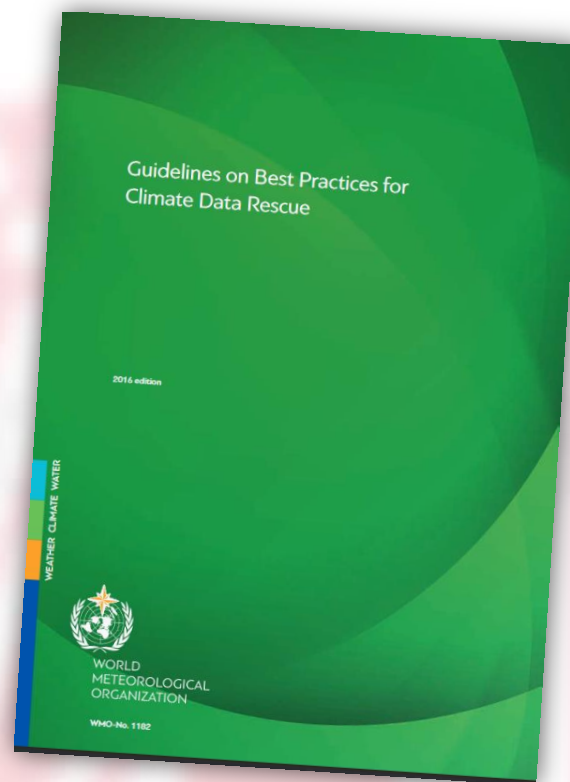
OVERVIEW

- **Background and motivation of the project**
- **Main features of the digital documents**
- **What is ARTYDOC and how it works**

BACKGROUND AND MOTIVATION OF THE PROJECT

In 2016, WMO addressed the four key activities of data rescue:

- cataloguing and storing of paper media to minimize the deterioration;
- imaging paper records for improving preservation and accessibility;
- digitizing data into computer compatible form;
- and continuous updating of digital records to up-to-date media and formats



BACKGROUND AND MOTIVATION OF THE PROJECT

Since 2001, many internal DARE projects have been undertaken at SMC trying to digitize and unify in a single repository as many climate documentation as possible.

Official meteorological archives



Agencia Estatal de Meteorología



Servei Meteorològic
de Catalunya

Historical Scientific Societies



Corporate archives (Electric power companies)



Religious archives (Jesuits and Pairists)



Local, provincial and county archives



Old Newspapers



Private donations



Historical observatories

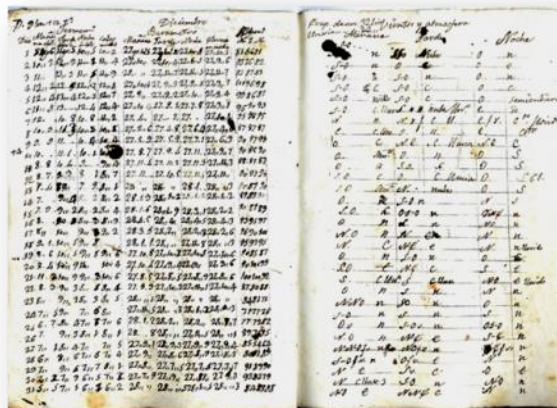


BACKGROUND AND MOTIVATION OF THE PROJECT

Flow-chart strategy 1 (2001-2005)



Identification
of repositories



Document consultation
within the archive



L26 P1 28 8 5. Error de impremta sin duda

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	1825	T1	T2	T3	P1	L	d	P2	L	d	P3	L	d	W1	W2	W3	R	Int
2	1	70	99	79	28	7	3	28	7	0	28	6	7	ss	se	ss	0	
3	2	70	88	71	28	5	0	28	9	0	28	5	8	ss	e	e	0	
4	3	60	91	77	28	5	5	28	5	1	28	5	3	nno	o	n	0	
5	4	72	93	81	28	5	2	28	4	9	28	4	1	no	oso	no	0	
6	5	74	102	85	28	2	8	28	2	0	28	2	5	oso	ene	no	0	
7	6	74	99	85	28	2	7	28	1	0	28	4	5	so	v	n	0	
8	7	59	80	63	28	4	5	28	4	6	28	4	6	no	ese	o	0	
9	8	54	80	70	28	4	0	28	4	0	28	4	1	so	oso	o	0	
10	9	68	93	81	28	4	2	28	3	2	28	4	7	s	o	v	0	
11	10	60	81	67	28	6	0	28	5	8	28	5	7	n	ono	o	0	
12	11	57	79	63	28	5	5	28	5	5	28	5	3	o	nne	so	0	
13	12	55	73	60	28	5	1	28	5	0	28	5	0	so	no	o	0	
14	13	61	73	70	28	4	8	28	4	5	28	4	2	n	no	n	0	
15	14	55	80	70	28	4	1	28	4	0	28	4	0	o	sse	ss	0	
16	15	66	86	72	28	4	0	28	3	3	28	3	3	n	e	n	0	
17	16	64	92	73	28	3	0	28	2	5	28	2	3	nne	se	e	0	
18	17	69	92	76	28	1	0	28	1	5	28	3	0	n	ne	n	1	3
19	18	60	91	72	28	3	9	28	3	5	28	3	1	nne	s	se	0	
20	19	66	86	80	28	1	5	28	1	3	28	0	2	no	ono	v	0	
21	20	60	82	71	27	11	0	27	11	0	27	10	6	ono	nno	o	0	
22	21	59	80	65	27	10	0	27	9	5	27	9	9	ono	so	o	0	
23	22	54	74	65	27	9	2	27	9	3	27	10	4	n	no	v	1	3
24	23	47	60	54	27	11	7	27	11	9	28	0	2	so	s	s	0	
25	24	42	69	52	28	0	4	28	0	2	28	0	9	n	e	o	0	
26	25	39	61	49	28	0	5	28	0	1	28	0	4	o	no	o	0	
27	26	39	75	54	28	0	0	28	1	0	28	2	0	nno	n	o	0	
28	27	35	75	59	28	1	8	28	2	2	28	4	1	no	so	o	0	
29	28	55	92	71	28	3	5	28	2	9	28	2	9	oso	v	e	0	
30	29	62	82	70	28	2	9	28	3	1	28	4	4	ono	ne	ne	0	
31	30	55	85	65	28	2	5	28	2	3	28	4	0	n	ese	ne	0	
32	31	50	87	57	28	4	2	28	4	4	28	5	2	no	sse	o	0	

Digitalitization

- Risk of document degradation
- No digital preservation
- Time consuming during QC process

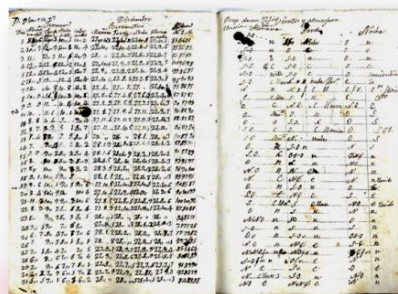
Flow-chart strategy 2 (2005-2013)



Identification
of repositories



Professional
imaging/scanning



Digital document



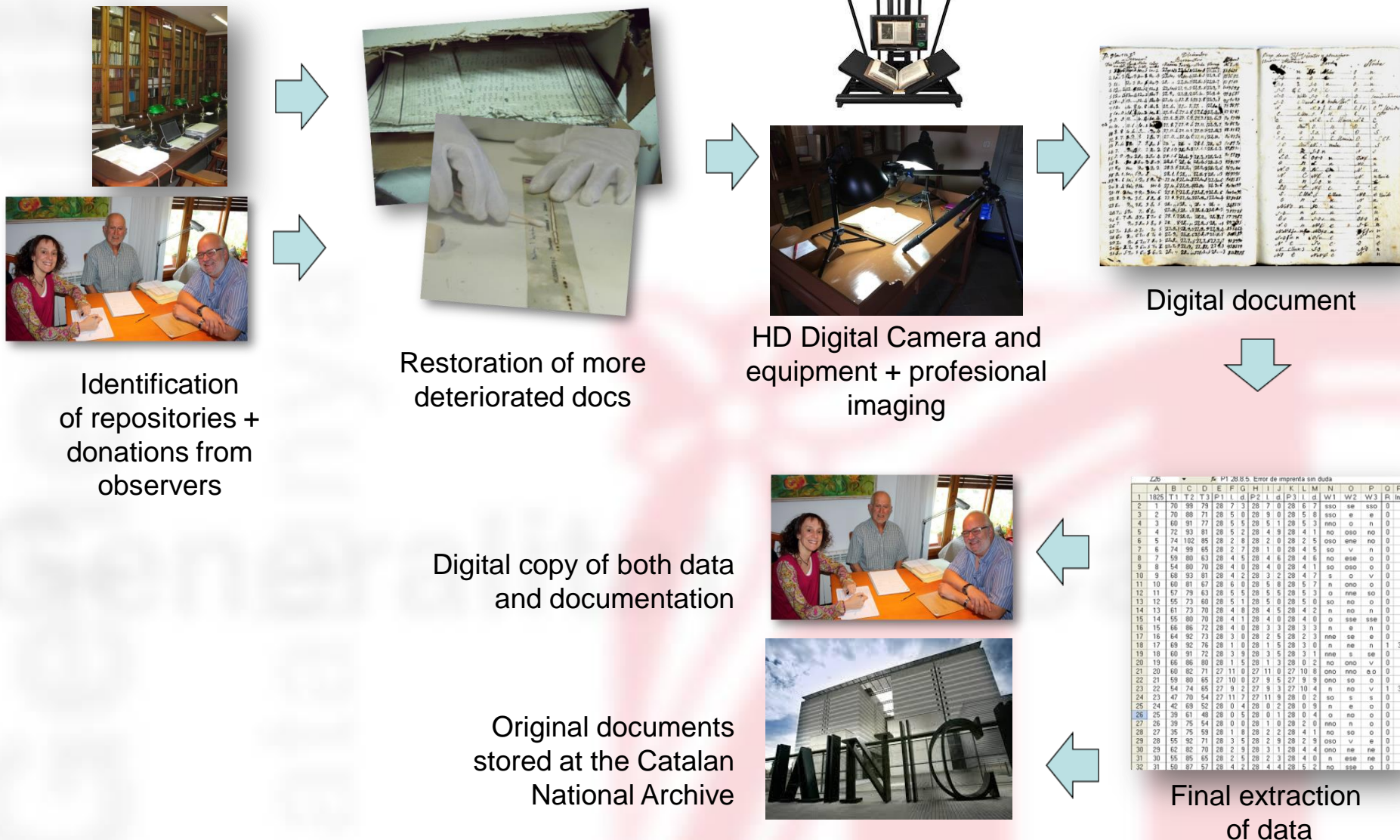
ZS		P1 2015.5 Error de empremta sin duda																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R			
1	1825	T1	T2	T3	P1	I	d	P2	I	d	P3	I	d	W1	W2	W3	R	int			
2	1	70	99	78	28	7	3	28	7	0	28	6	7	ss0	se	ss0	0				
3	2	70	98	71	28	5	0	28	9	0	28	5	8	ss0	e	e	0				
4	3	60	91	77	28	5	5	28	5	1	28	5	3	nn0	o	n	0				
5	4	72	93	81	28	5	2	28	4	9	28	4	1	no	oso	no	0				
6	5	74	102	85	28	2	8	28	2	0	28	2	5	oso	ene	no	0				
7	6	74	99	65	28	2	7	28	1	0	28	4	5	no	v	n	0				
8	7	59	80	63	28	4	5	28	4	6	28	4	6	no	ese	o	0				
9	8	54	80	70	28	4	0	28	4	0	28	4	1	so	oso	o	0				
10	9	60	93	81	28	4	2	28	3	2	28	4	7	s	o	v	0				
11	10	60	81	67	28	6	0	28	5	8	28	5	7	n	omo	o	0				
12	11	67	79	63	28	5	5	28	5	5	28	5	3	o	nne	so	0				
13	12	55	73	60	28	5	1	28	5	0	28	5	0	so	no	o	0				
14	13	61	73	70	28	4	0	28	4	5	28	4	2	n	no	n	0				
15	14	55	80	70	28	4	1	28	4	0	28	4	0	o	sse	ss0	0				
16	15	66	86	72	28	4	0	28	3	3	28	3	3	n	e	e	n	0			
17	16	64	92	73	28	3	0	28	2	5	28	2	3	nne	se	e	0				
18	17	69	92	76	28	1	0	28	1	5	28	3	0	n	ne	n	1	3			
19	18	60	91	72	28	3	9	28	3	5	28	3	1	nne	s	se	0				
20	19	66	86	80	28	1	5	28	1	3	28	0	2	no	omo	v	0				
21	20	60	92	71	27	11	0	27	11	0	27	10	8	omo	mo	0	0				
22	21	59	80	65	27	10	0	27	9	5	27	9	9	omo	so	o	0				
23	22	54	74	65	27	9	2	27	9	3	27	10	4	n	no	v	1	3			
24	23	47	70	54	27	11	7	27	11	9	28	0	2	s	s	0					
25	24	42	69	52	28	0	4	28	0	2	28	0	9	n	e	o	0				
26	25	39	61	48	28	0	5	28	0	1	28	0	4	o	no	o	0				
27	26	39	75	54	28	0	0	28	1	0	28	2	0	nno	n	o	0				
28	27	35	75	59	28	1	8	28	2	2	28	4	1	no	so	o	0				
29	28	55	92	71	28	3	5	28	2	9	28	2	9	oso	v	e	0				
30	29	62	82	70	28	2	9	28	3	1	28	4	4	omo	ne	ne	0				
31	30	55	95	65	28	2	5	28	2	3	28	4	0	n	ese	ne	0				
32	31	50	87	57	28	4	1	28	4	4	28	5	5	no	sse	o	0				

Final extraction
of data

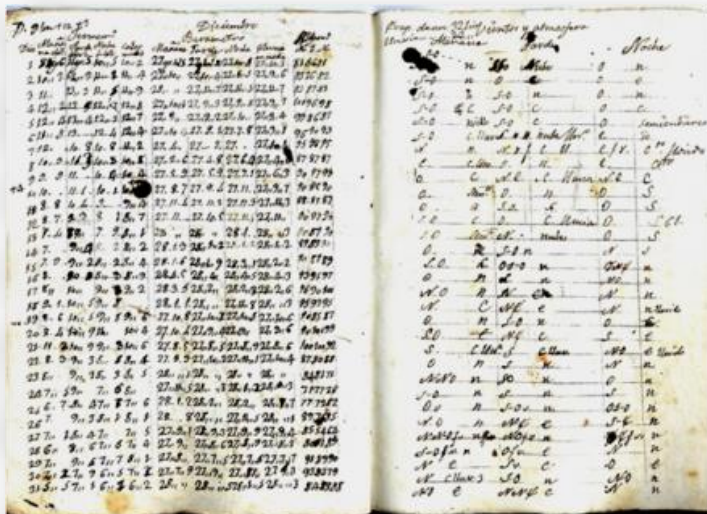
- Professional imaging is very expensive
- The original documents demands restoration in many cases
- Wide range of digital storage platforms: DVDs, CDs, external hard disks...

BACKGROUND AND MOTIVATION OF THE PROJECT

Flow-chart strategy 3 (since 2013)

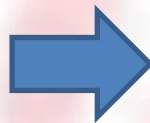


Z26		EJ P1 20.8.5 Error de impremta sin duda																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	1825	T	1	T	2	T	3	P	1	1	6	P	2	1	6	P	3	1
2	1	70	59	78	1	28	7	3	28	7	0	28	6	7	550	58	550	0
3	2	70	68	71	28	5	0	28	9	0	28	5	8	550	58	550	0	
4	3	60	91	77	28	5	5	28	5	1	28	5	3	550	58	550	0	
5	4	72	93	81	28	5	4	28	4	1	28	4	1	550	58	550	0	
6	5	74	102	85	28	2	8	28	2	0	28	2	5	550	58	550	0	
7	6	74	99	85	28	2	7	28	1	0	28	1	0	550	58	550	0	
8	7	59	80	63	28	4	5	28	4	6	28	4	6	550	58	550	0	
9	8	54	60	70	28	4	0	28	4	0	28	4	0	550	58	550	0	
10	9	68	93	81	28	4	2	28	3	2	28	3	2	550	58	550	0	
11	10	60	81	67	28	6	0	28	5	8	28	5	7	550	58	550	0	
12	11	57	79	63	28	5	5	28	5	5	28	5	3	550	58	550	0	
13	12	55	73	60	28	5	1	28	5	0	28	5	0	550	58	550	0	
14	13	61	73	70	28	4	8	28	4	5	28	4	2	550	58	550	0	
15	14	55	80	70	28	4	1	28	4	0	28	4	0	550	58	550	0	
16	15	66	66	72	28	4	0	28	3	3	28	3	3	550	58	550	0	
17	16	64	92	73	28	3	0	28	2	5	28	2	3	550	58	550	0	
18	17	69	92	76	28	1	0	28	1	5	28	1	0	550	58	550	0	
19	18	60	91	72	28	3	9	28	3	5	28	3	1	550	58	550	0	
20	19	66	86	80	28	1	5	28	1	3	28	1	0	550	58	550	0	
21	20	60	82	71	27	11	0	27	11	0	27	11	0	550	58	550	0	
22	21	59	80	65	27	10	0	27	9	5	27	9	5	550	58	550	0	
23	22	54	74	65	27	9	2	27	9	3	27	9	3	550	58	550	0	
24	23	47	70	54	27	11	7	27	11	9	27	11	9	550	58	550	0	
25	24	42	68	52	28	0	4	28	0	2	28	0	9	550	58	550	0	
26	25	39	61	48	28	0	5	28	0	1	28	0	4	550	58	550	0	
27	26	39	75	54	28	0	0	28	1	0	28	1	0	550	58	550	0	
28	27	35	75	59	28	1	8	28	2	2	28	2	1	550	58	550	0	
29	28	55	92	71	28	3	5	28	2	9	28	2	9	550	58	550	0	
30	29	62	82	70	28	2	4	28	2	3	28	2	4	550	58	550	0	
31	30	65	85	65	28	2	5	28	2	3	28	2	4	550	58	550	0	
32	31	50	87	57	28	4	2	28	4	4	28	4	2	550	58	550	0	



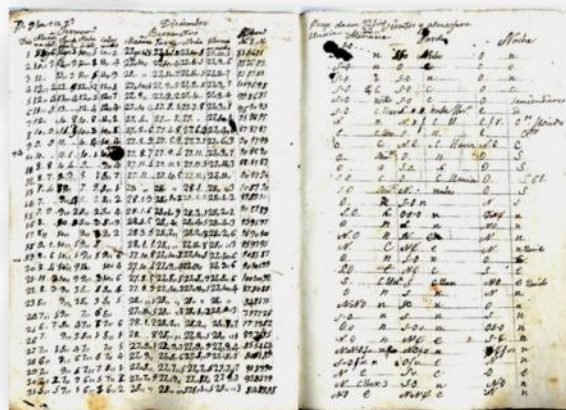
**HUGE AMOUNT OF DIGITAL
INFORMATION !!!!**

**6.1 TB ... and growing!
(August 2018)**

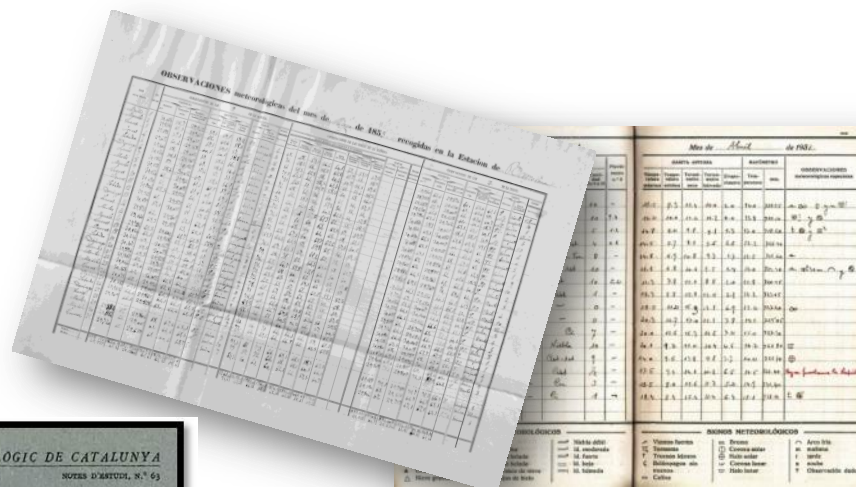


- Need for a unique and save storing system: NAS (Network Attached Storage) with 14 TB of capacity
- Need for a usable and reliable cataloguing tool

MAIN FEATURES OF THE DIGITAL INFO: TYPOLOGY



Manuscripts



Meteorological
sheets and logbooks



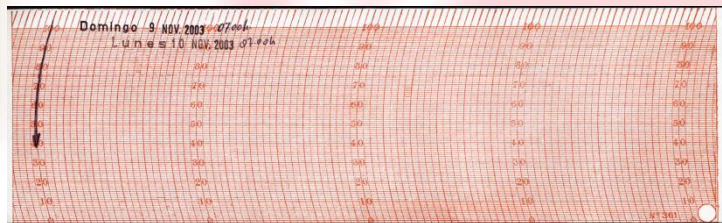
Publications



Photos



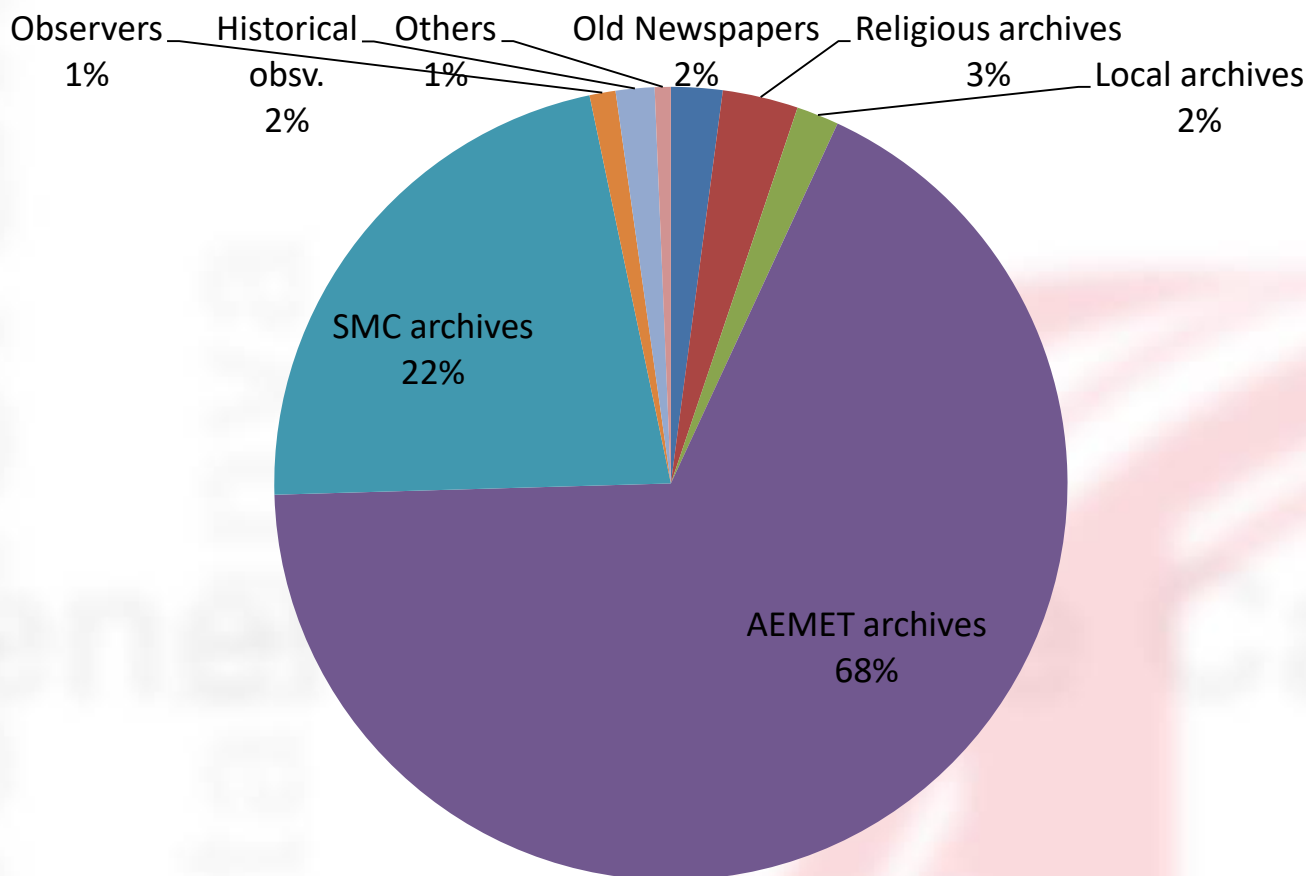
Historical Press



Strip charts

MAIN FEATURES OF THE DIGITAL INFO: FILES

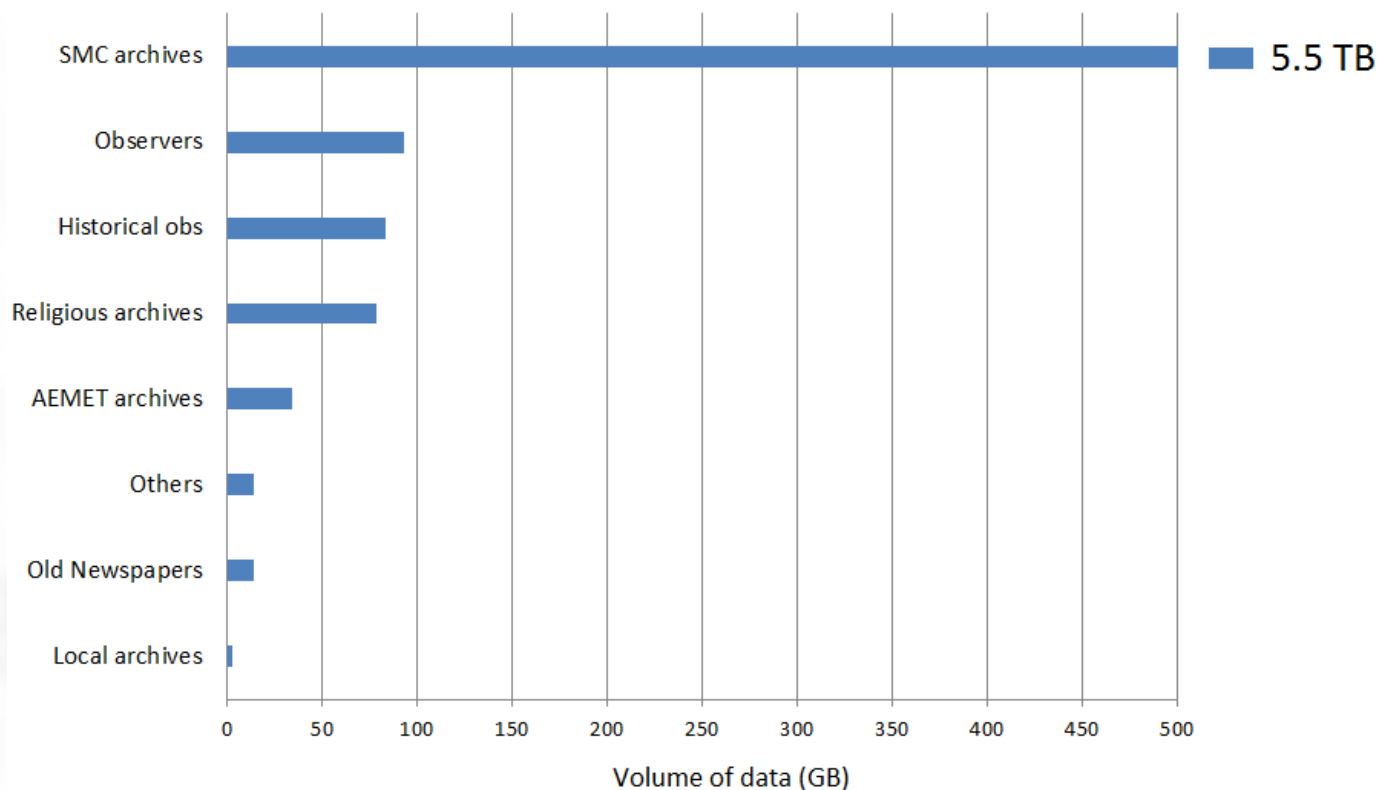
- **542,000 digital files** in a variety of formats (png, tiff, bmp, jpg, pdf) and of different quality/resolution.



Where do digital files come from?

MAIN FEATURES OF THE DIGITAL INFO: VOLUME

- 86% of the volume comes from SMC archives (higher resolution)



Volume of digital data (in GB)

WHAT IS ARTYDOC AND HOW IT WORKS

- ARTYDOC is the solution for storage, management, cataloguing, description and dissemination of the climate digital information at SMC.
- Web-based software
- Software developed by ARTYPLAN a leading Catalan company highly specialized in any print and copy service (www.artyplan.com). In 2003 started a new division: Document. This division linked to the world of the archives, biblioteconomics, documentation and TIC.
- ARTYDOC for climatic digital archives is being developed jointly with SMC.



WHAT IS ARTYDOC AND HOW IT WORKS


- DESIGN: the main point is how the information is internally classified according to the variety of info and digital formats.
- CATEGORIES:
 1. Numerical data (sheets)
 2. Publications and historical press
 3. Correspondence
 4. Photographies: weather stations, weather phenomena
weather instruments, phenology stages
 5. Strip charts (thermographs, pluviographs...)

WHAT IS ARTYDOC AND HOW IT WORKS



LOGIN PAGE

Usuario

 mcanas_ext@meteo.cat


Contraseña



☐ Permanecer registrado

ENTRAR

WHAT IS ARTYDOC AND HOW IT WORKS



Montse Cañas
ADMINISTRADOR

MENÚ ADMINISTRADOR

- Usuaris
- Grups
- Permisos
- Estadístiques

CATEGORIES

- Registres SMC
- Grups de sèries
- Sèries
- UDS Registres meteorològics
- Publicacions i premsa històrica

Categories Grups de sèries

Searcher of digital files

[Inici](#) / [Grups de sèries](#)

Buscador/filtre

Columnes

Navegador visual **Taula de registres**

Codi de referència	Nom	Data Creació	Nº Doc
CAT_SMC_FD	Fons amb dades d'observacions meteorològiques o fenològiques	27/06/2016	0
CAT_SMC_FF	Fons fotogràfic	27/06/2016	0

Registres 1 a 2 de 2

[Exportar a CSV](#) [Editar tots els registres](#)

Registres nous

UDS Registres meteorològics	Esparreguera - 4	2017/14/07
UDS Registres meteorològics	Esparreguera - 4	2017/14/07
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28
UDS Registres meteorològics	Central Nuclear de Vandellòs	2017/06/28

Resum de la classificació

	Registres	Imatges
Registres SMC		
Grups de sèries	2	0
Sèries	2	0
UDS Registres meteorològics	330679	360015




Navigation menu



Last entries

WHAT IS ARTYDOC AND HOW IT WORKS



Montse Cañas
ADMINISTRADOR

MENÚ ADMINISTRADOR

- Usuaris
- Grups
- Permisos
- Estadístiques

CATEGORIES

- Registres SMC
 - Grups de sèries
 - Sèries
- UDS Registres meteorològics
- Publicacions i premsa històrica

Montse Cañas
català

Categories UDS Registres meteorològics

Inici / UDS Registres meteorològics

Buscador/filtre

Variables

conté

Temperatura Màxima

Data inicial

major que

1958/09/01

Comarca

igual

Berguedà

+

Buscar

Buscar...

Nou

Columns

Navegador visual

Taula de registres

Codi artydoc	Codi estació SMC	Codi estació AEMET	X UTM	Y UTM	Z	Nom de l'estació meteorològica	Municipi	Comarca	Variables	Resolució temporal
CAT_SMC_FD.01.312593	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312594	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312595	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312596	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312597	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312598	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312599	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària
CAT_SMC_FD.01.312600	BG033	0142	397793	4644775	725	Serrateix	Viver i Serrateix	Berguedà	Temperatura Màxima - Temperatura Mínima	Diària

< 1 2 >

Registres 1 a 8 de 701

Exportar a CSV

Editar tots els registres

Several filters can be implemented

Results in a list format

WHAT IS ARTYDOC AND HOW IT WORKS

[illegible]

General		
Codi artydoc		
CAT_SMC_FD.01.00		
Codi estació SMC		
BC005		
Codi estació AEMET		
00021		
X UTM	Y UTM	Z
320783	4536219	34
Nom de l'estació meteorològica		
Central Nuclear de Vandellòs		
Municipi		
Vandellòs i l'Hospitalet de l'Infant		
Comarca		
Baix Camp		
Observador/s		
Sr. Soria, 1974-1976?Josep Roca, 1977(gen)-1984(nov)Q. Borràs, 1984(des)-1985(abr)EMA, 1985(mai)		
Variables		
Precipitació - Meteors - Direcció del Vent		
Resolució temporal		
Diària		
Format		
Targetes		
Productor		
Agència Estatal de Meteorologia		
Font de procedència		
Banco Nacional de Datos (INM)		
Informador		
Administradors		

Editar
Clonar
Nou
Eliminar

Dades tècniques

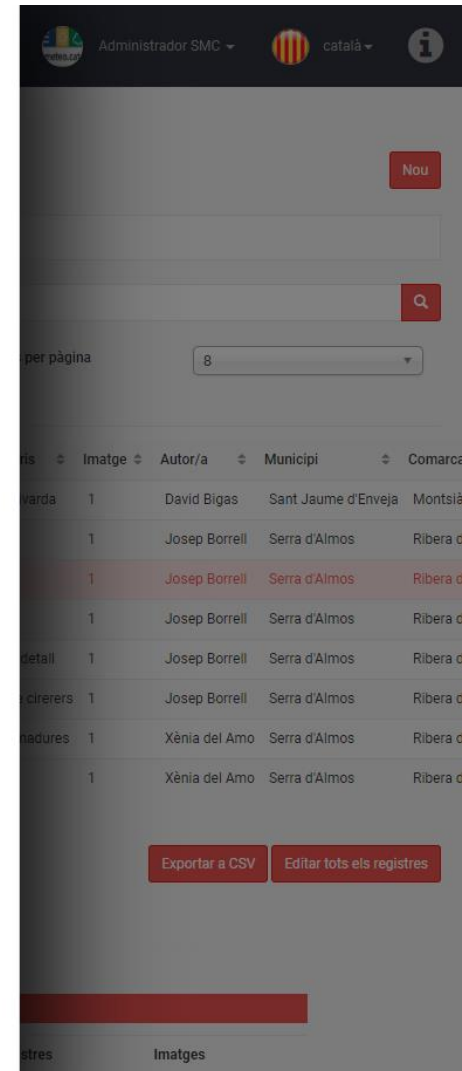
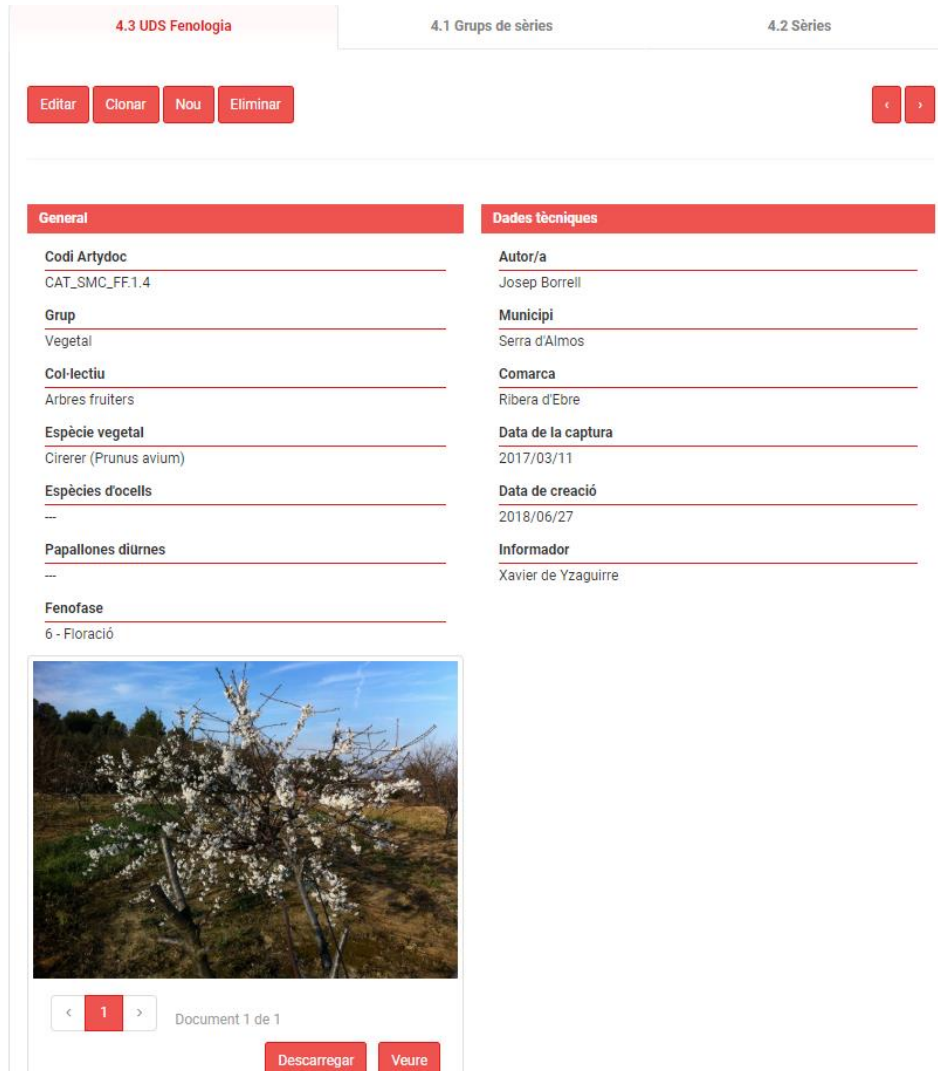
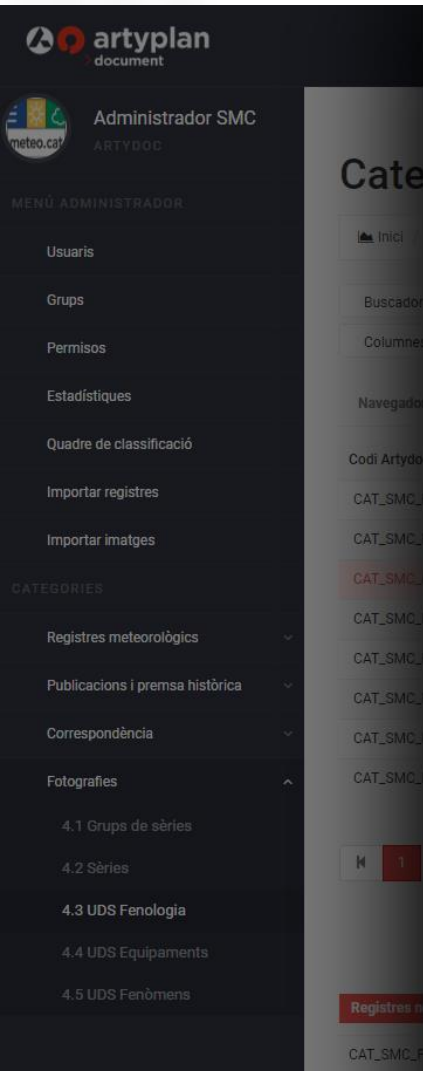
Montse Cafias

català

Nou

ador ⇅	Nº Doc ↓	Data inicial ⇅	Data final ⇅	Data Creació ⇅
adors 1		01/02/1974	28/02/1974	28/06/2017
adors 1		01/09/1974	30/09/1974	28/06/2017
adors 1		01/03/1982	31/03/1982	28/06/2017
adors 1		01/06/1974	30/06/1974	28/06/2017
adors 1		01/10/1975	31/10/1975	28/06/2017
adors 1		01/12/1975	31/12/1975	28/06/2017
adors 1		01/01/1983	31/01/1983	28/06/2017
adors 1		01/02/1983	28/02/1983	28/06/2017

WHAT IS ARTYDOC AND HOW IT WORKS



Handwritten letters

SMC publications (1921-1938)

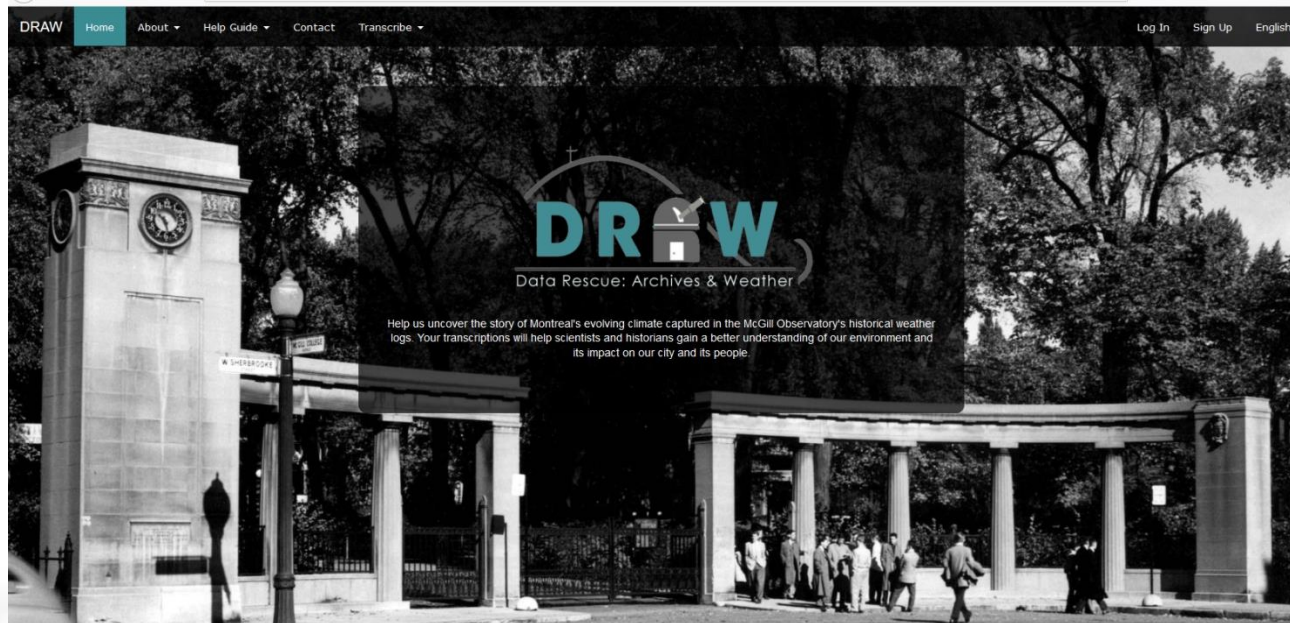
Phenology stages

GOOD NEWS

- ARTYDOC fits well SMC demands for digital storing and cataloguing
- Useful tool for preservation and consultation
- Is fully transparent: the administrator has full capacity for adding new requirements
- Very adequate for QC purposes and metadata sources
- All kind of digital documents can be catalogued: weather stations, phenological stages, weather phenomena, instrumentation...
- About 70% of the images are already catalogued
- The platform will be publicly accessible soon (hopefully)

MORE TO COME...

- Explore a similar experience as DRAW project (Data Rescue: Archives and Weather)
<https://citsci.geog.mcgill.ca>



Project allows volunteers to participate in the transcription of historical weather logs captured at the McGill University McLeod Observatory from 1863 to 1963 (Canada) = **CITIZEN SCIENCE**

Data from the **Observatory at the University of Barcelona** (1855-1936)

WHAT IS ARTYDOC AND HOW IT WORKS: SUMMARY

Observaciones de la Estación meteorológica de

1 ^h del de 19...		(7 ^h u 8 ^h) del de 19...		13 ^h del <u>24</u> de <u>enero</u> 19 <u>35</u>		18 ^h del de 19...	
Barógrafo	Lectura directa.....	41.7	Termómetro unido.....	10.4	Termómetro unido.....	10.6	Termómetro unido.....
	Corrección ±	0.0	Lectura directa.....	72.45	Lectura directa.....	771.56	Lectura directa.....
	Lectura corregida.....	71.7	Corrección a 0° (I).....	13	Corrección a 0° (I).....	1.3	Corrección a 0° (I).....
Termógrafo	Lectura directa.....	6.0	Corrección instrumental.....	71.15	Corrección instrumental.....	770.26	Corrección instrumental.....
	Corrección ±	5.3	Reducción al nivel del mar.....	0.05	Reducción al nivel del mar.....	0.05	Reducción al nivel del mar.....
	Lectura corregida.....	9.3	Idem a milibares (III).....	71.20	Idem a milibares (III).....	770.31	Idem a milibares (III).....
Higrógrafo	Lectura directa.....	56	Reducción al nivel del mar.....	74.80	Reducción al nivel del mar.....	773.9	Reducción al nivel del mar.....
	Corrección ±	14	Idem a milibares (III).....	330	Idem a milibares (III).....	1031.8	Idem a milibares (III).....
	Lectura corregida.....	70	Variación en 3 ^h anter. a la ob.	0.0	Variación entre 10 ^h y 30 ^h	-1.3	Variación entre 15 ^h y 18 ^h
Textos regis- tradores..	Dirección del viento...		Reducción a milibares (II).....	0.0	Reducción a milibares (II)	-1.7	Reducción a milibares (II)
	Velocidad.....		Forma de la variación.....	7	Forma de la variación.....	8	Forma de la variación.....
OBSERVACIONES EN 24 HORAS		Viento.—Dirección..... Fuerza.....		Viento.—Dirección..... Fuerza.....		Viento.—Dirección..... Fuerza.....	
Temperatura máxima del día.....		Psicrómetro. { Termómetro seco.....		Psicrómetro. { Termómetro seco.....		Psicrómetro. { Termómetro seco.....	
Idem mínima del día.....		Idem húmedo.....		Idem húmedo.....		Idem húmedo.....	
Lluvia total en milímetros.....		Diferencia.....		Diferencia.....		Diferencia.....	
Evaporación.....		Tensión del vapor.....		Tensión del vapor.....		Tensión del vapor.....	
Recorrido total del viento en kms.....		Humedad.....		Humedad.....		Humedad.....	
Temp. ^a mínima a 0,10 m. del suelo.....		NUBES BAJAS { Clase.....		NUBES BAJAS { Clase.....		NUBES BAJAS { Clase.....	
Idem Id. junto al suelo.....		Dirección..... Cantidad.....		Dirección..... Cantidad.....		Dirección..... Cantidad.....	
Horas de sol en el día h. m.....		NUBES ALTAS { Clase.....		NUBES ALTAS { Clase.....		NUBES ALTAS { Clase.....	
		Dirección..... Cantidad.....		Dirección..... Cantidad.....		Dirección..... Cantidad.....	
		Cantidad total de nubes.....		Cantidad total de nubes.....		Cantidad total de nubes.....	
Pequeñas precipitaciones y fenómenos lumi- nosos en el día.....		Lluvia recog. desde la ob. ant. m/m.....		Lluvia recog. desde la ob. ant. m/m.....		Lluvia recog. desde la ob. ant. m/m.....	
		Estado del mar.....		Estado del mar.....		Estado del mar.....	
		Visibilidad al I.....		Visibilidad al I.....		Visibilidad al I.....	
		Id. al II.....		Id. al II.....		Id. al II.....	
		Id. al III.....		Id. al III.....		Id. al III.....	
		Id. al IV.....		Id. al IV.....		Id. al IV.....	

*Eclipse parcial de sol
Se observa el tono especial
de la luz de eclipse 19.1*

Nagyon
köszönöm !

