

Design of an EMIDA database on European research institutions and their major publication topics

(Version 1 - April 2009)

Jean De Rycke EMIDA WP2 leader

Institut National de la Recherche Agronomique, IASP, Centre de Tours, 37380 Nouzilly, France Tel : 33 (0)2 47 42 75 49 E-mail : <u>jean.derycke@tours.inra.fr</u>

TABLE OF CONTENTS

1. Objective of the database	2
2. Principles of database construction	2
3. Web-based bibliometric resources and compared potentials	2
4. Construction of search queries covering the entire field of EMIDA	3
→ GENERAL DESCRIPTION	3
→ STEP BY STEP DEMONSTRATION OF DESCRIPTORS SELECTION	4
5. Running search queries with controlled descriptors on WoS platform	7
6. Extraction of information on European organisations from the WoS and integration	n in
Excel data base	11
7. Standardisation of organisation names	11
8. Specific CAB search for extraction of data on transversal fields	11
9. Translation of spreadsheet in web-based searchable database	12
10. Limits of the methodology used	13

1. Objective of the database

To map research production in the field of animal health in Europe during the recent years (from 2004 onward), according to scientific <u>topics</u>, research <u>organisations</u>, and <u>number</u> of scientific papers and patents. The focus is on animal infectious and parasitic diseases, the chief concern of EMIDA.

2. Principles of database construction

Through data extraction from ISI Web of Knowledge SM, to build up an Excel spreadsheet displaying the main following fields:

- Name of disease or pathogen
- Search query formulation
- Name of research organisation, City, Country
- Number of source records for each theme and research organization

A detailed description of the source data bases, their specific features and functionalities, can be found on the website of the corresponding platforms, run by Thomson Scientific: <u>Thomson Reuters - ISI Web of Knowledge</u>SM

3. Web-based bibliometric resources and compared potentials

Two complementary source platforms run by the ISI Web of Knowledge SM were used: the CAB Abstracts [®], specialized in agronomic research including animal health, and the Web of Science [®] (WoS) link covering all disciplines. The advantage of CAB Abstracts [®] over the WoS [®] is that each source document is allocated (i) one or more standard classification codes (termed "cabicodes"), indicating the broad subject areas of the paper and (ii) several controlled descriptors (including names of pathogens, diseases and host animals) that are present in a controlled thesaurus. These features lend themselves to normalized procedures of sorting and statistical analysis of source documents, on the basis on topics. However, compared to the WoS, the CAB lacks a function that is essential for the mapping of institutions: sorting on authors' addresses, more specifically on organization names, town of location, and countries. This function is present in the WoS platform. Therefore an interplay between the two platforms was necessary.

Overall procedure

1. Construction of search queries using CAB® controlled descriptors



3. Construction of an Excel [®] spreadsheet with preset fields including: Topic, Animal, Numb. of publications, Institution Name, Town & Country

4. Standardization of institution names

5. Translation of spreasheet in a web-based searchable database



4. Construction of search queries covering the entire field of EMIDA

GENERAL DESCRIPTION

A specific challenge of the present bibliometric analysis was to construct a complete set of search queries that would cover in an exhaustive manner, and without major omissions, the research domain of EMIDA, i.e. animal infectious diseases. A particularity of this research domain is that it can be satisfactorily described by the names of diseases and of causative pathogenic agents, provided that an exhaustive list of these descriptors can be generated, then prioritized.

One way to achieve this goal could have been to exploit existing lists of diseases and of causative pathogenic agents made available by specialized expert groups. But it appears that more exhaustive lists of all the natural descriptors present in the literature could be produced in a systematic manner, then prioritised, using specific functionalities of the CAB Abstracts [®] platform. Very briefly, these are the main steps leading to the formulation of search queries using descriptors controllable in the specific CAB thesaurus.

- (i) Extraction of all the documents, published since 2004, identified by at least one of the three **broad area qualifiers** (called "cabicodes" or CC) that specifically describe the field of infectious and parasitic diseases of animals:
 - LL821 for "Prion, Viral, Bacterial and Fungal Pathogens of Animals";
 - LL822 for "Protozoan, Helminth, Mollusc and Arthropod Parasites
 - LL823 for "Veterinary Pests, Vectors and Intermediate Hosts"
- (ii) Identification and statistical analysis of all the specific descriptors of diseases and pathogens automatically generated by the platform. In the same way, identification of all possible descriptors for host animals in each food animal group.
- (iii) Grouping of descriptors by a specialist of animal infectious diseases on the one hand by "disease cluster" (describing one disease or several closely related diseases) and on the other hand by "animal category"; then formulation of all the search queries basically composed of the two above sets of terms linked by Boolean operators and using appropriate wild cards, such as in the example represented in the diagram below:



Adaptations to this general search scheme was necessary to collect data on a number of **transversal topics** such as "vaccines and vaccination", "quantitative epidemiology of infectious diseases" etc...Such searches were made feasible due to the existence of further specific "cabicodes", used in combination with other broad descriptors. The more useful cabicodes (CC) were the following:

- LL650: Animal Immunology; *
- **HH600:** Host Resistance and Immunity;
- **LL886**: Diagnosis of animal diseases;
- YY700: Pathogens, Parasites and Infectious Diseases (Wild Animals);
- HH405: Pesticides and Drugs: Control.



STEP BY STEP DEMONSTRATION OF DESCRIPTORS SELECTION

To display all possible descriptors of diseases and pathogens for a specified animal category the search query must combine the CC (Cabicode) describing a research field with thesaurus descriptors of this category of production animals"

For instance; the search query to display all possible "disease & pathogen" descriptors of bacterial, viral and prion diseases in pigs should combine the CC=LL821 with animal species CAB thesaurus controlled descriptors "pigs", "wild pigs", "piglets", "sows", "swine diseases". The result of a typical search is illustrated below:

♥ISI Web of Knowledge [v.4.1] - CAB Abstracts Home - Mozilla Firefox	
Eichier Édition_Affichage_Historique_Marque-pages_Outils_?	0
🤕 🔹 🔿 🗸 🕑 🛞 🏠 🗋 http://apps.isiknowledge.com/CABI_GeneralSearch_input.do?product=CABI8search_mode=GeneralSearch&SID=N133Cp8G2ci 🔹 🕨 🖸	Google
🗭 Démarrage 🔝 Dernières nouvelles (en) 🧃 DA: History: General a	
🗶 INTRANET de l'unité IASP - 💿 🔄 Liste alphabétique des bases de 🔄 🔄 CABICODE CABI [v.1] 💿 💽 ISI Web of Knowledge [v.4.1]	•
Signed In My EndNote Web My ResearchertD My Citation Alerts My Journal List	My Saved Searches Log Out Help
ISI Web of Knowledge SM Take the next step All Databases Select a Database CAB Abstracts Additional Resources	Institut National de la
Search Advanced Search Search History Marked List (0)	Recherche Agronomique
CAB Abstracts®	
Search for:	Welcome, De Rycke Jean
LL821 in CABICODES V Q Example: J300 OR Soil Physics Q AND V pigs OR wild pigs OR piglets OR sows OR swine diseases in in Descriptors V Q Example: folder legumes Q AND 2004-2008 in Year Published V Example: 2001 or 1997-1999 Add Another Field >> Search Clear Current Limits: (change Limits and Settings) Timespan-Al Years	Maintenance Alert Attention Proop Sarver Attention Proop Sarver Attentist actacs: Plassa citck here for valuable information regarding prove sarver access to the ISI Web of knowledge. Discover CAB Abstracts Resp track of every area of agricultural science, veterinary medicine, nutrition, and natural resources. • Over 7,400 journals from 140 countries in 50 languages • More than 50 languages • More than 200,000 records added annually Available with up to 8.9 million records to 1910 • Vant to knowr mote?
Please give us your feedback on using ISI Web of Knowledge.	My Web of Knewledge
Acceptable Use Policy Copyright 9000 The Thomason Corporation Copyright 92 2008 CAB International All rights reserved.	My EndNote Web My Citation Alerts

démarrer C Courrer entran... D 2 Microsoft O... **• S** 151 Web of Kno... FR **d C** 15:27 On February 14th 2008, this search selected 4494 source documents (in majority journal articles) since 2004.

ISI Cichic	Web of Knowledge [v.4.	1] - CAB Abstracts Results - Mozilla Firefox	
	- Colouri Ameriage Ascologie M	ni querpages Quuis <u>r</u>	
-			
See De	emarrage 🔯 Dernieres nouvelles (en) (D DA: History: General a	
× n	Timespan=All Years, Datab	Ises apriaderque des dases de CABICODE CABI [V.1] ISI Web of Knowle ises=ABSTRACTS.	eage [v.4.1
	Results: 4 494	I≪ ≼ Page 1 of 450 💿 ►►	Sort by. Latest Date
fine	Refine Results	Print (E-mail) (Add to Marked List) (Save to EndNolo Web) more options	📻 Analyze Results
Hide Ref	Search within results for Search • Subject Areas Refine • Intrect Toulos Discasses (44/2) • Ortenent Tab. sciences (4,412) • Advances (4,412) • Advances (4,412) • Advances (4,412) • Advances (4,412) • Decision (4,412) • Advances (4,412)	 I. Title: Postweaning multisystemic wasting syndrome (PMWS) in Sweden. From an exotic to an er Author(s) Waligren, P. Beiak, K. Ehlorsson, C. J., et al. Source Veterinary Quarterly Volume: 29 Issue: 4 Page(s): 122-137 Published: 2007 OFFX 2. Title: A successful national control programme for enzotic respiratory diseases in pigs in Switzer Author(s): Stark, K. D. C., Miserez, R., Siegmann, S., et al. Source: Revue Scientifique et Technique - Office International des Epizooties. Volume: 26 Issue: 3 ③ Title: Swine vesicular disease surveillance and eradication activities in Italy. Author(s): Bellini, S., Santucci, U., Zanardi, G., et al. Source: Revue Scientifique et Technique - Office International des Epizooties. Volume: 26 Issue: 3 ③ Title: The immunogenicity of fusion protein linking the carboxyl terminus of the B subunit of Shiga Author(s): Revin, S. and Loo, Marghen, Lu Jin-Juan, et al. 	ndemic disease. nland. Page(s): 595-606 Published: 2007 Page(s): 595-593 Published: 2007 toxin 2 to the B subunit of E. coli heat-labile enterotoxin.
	BOOK (36) more Authors	Source: vereinnary microsology volume: 122 issue: 112 Page(s): 209-215 Published: 2008	in animal isolates.
	Source Titles	Author(s): Albini, S., Brodard, I., Jaussi, A., et al. Source: Veterinary Microbiology Volume: 127 Issue: 1/2 Page(s): 179-185 Published: 2008	
	Publication Years	ØS-F-X Full Text	
	Descriptors CABICODES Languages For more advanced refine options, use	6. Title: Effect of sow vaccination against Mycoplasma hypopneumoniae on sow and piglet colonizatic Author(s) Siblia, M. Bernal, R., Torrents, D., et al. Source: Veterinary Nicrobiology Volume: 127 Issue: 12 Page(s): 165-170 Publiched: 2009 0:FrX: Fullrect	on and seroconversion, and pig lung lesions at slaughter.
Termi	né	💾 MSP_1_240	108.doc - Microsoft Word
-	démarrer 🔰 🦉 🐸 🔍 🔣 🗉	i 🗽 📴 🧮 🗃 🗁 🦻 😻 🖾 🗃 🍣 🤌 🏧 🔯 🔇) 🔻 😻 ISI Web of Kno 🛛 FR 🛛 📶 🛒 🕅 🖄 🥥 15:31



As can be seen of the left column, descriptors can be sorted according to their frequency in a second step.....

🖲 Results Analysis CAB Abstracts - Mo	ozilla Firefox		
Eichier Édition Affichage Historique Marque-pages	<u>O</u> utils <u>?</u>		\$
 	ge.com/OutboundService.do?action=go&mode=ra	aService&SID=N133Cp8G2ceo9jc7gh@&product=CABI{	• D Google
🌩 Démarrage 🔂 Dernières nouvelles (en) 🦸 DA: History	: General a		
😿 INTRANET de l'unité IASP - 💿 📄 Liste alphab	étique des bases de 💽 📔 CABICODE CAB	I [v.1] 🛛 🕞 Results Analysis CAB A	bs 🖬 🔹
		Signed In My EndNote Web My Citation Alerts	My Journal List My Saved Searches Log Out Help
ISI Web of Knowledge	M Take the next step ⊘		
<<< Back to results list	Analyz	e Results	
4,494 records. CABICODES=(LL821) AND Descriptors Rank the records by this field:	/Broad Descriptors=(pigs OR wild pigs OR piglets Analyze:	DR sows OR swine diseases) AND Year Published=(2004 Set display options:	Sort by:
Author CABICODES Descriptor Document Type	upTo 500 v records.	Show the top 500 🔽 Results. Minimum record count (Threshold): 2	 Record count Selected field
Analyze			
	Please give us your Feedback	on using ISI Web of Knowledge.	
	Acceptab Copyright © 2008 Th	le Use Policy e Thomson Corporation	
Terminé			
🦺 démarrer 🔰 🧖 💆 🛇 🔣 🛒 😒 🗈 💻 🛚	🛛 💽 🗀 🤥 🐸 🗖 🐨 🏐 🥭 📼 🕅 🖉	🎒 Courrier entran 🛛 🖉 2 Microsoft O 🔻	🥹 Results Analysi 🛛 🕅 🚮 🛒 🕅 🙆 🥥 15:33

...which leads to the final display below comprising up to 500 descriptors organized in tabulated format. The window shows the very first descriptors of the list, among which only those describing specific diseases and pathogens are selected. Here the most frequent descriptors referring to a pathogen or a disease are "arterivirus" and, "porcine reproductive and respiratory syndrome", which in fact points to the same disease/pathogen entity. The next most frequent descriptor is "circovirus", etc....

🕲 Re	sults Analys	sis CAB Abstracts - Mozilla Firefox						_ 7 🗙
Eichier	r Éditio <u>n A</u> ffich	nage <u>H</u> istorique <u>M</u> arque-pages <u>O</u> utils <u>?</u>						0
- 🏷	🖒 - 💽 🐼	http://pcs.isiknowledge.com/analyze/ra.cgi				▼ ▶ G	- Google	Q
Dái	marrago 🛐 Dorn	ières pervelles (en) 🥵 DA: History: Caparal a						
Ser Der	TRANET de llucité	teres nuuvenes (en) 😰 DA: history, deneral a	DICODE C	NDT [
	Author	Up to SUU V records.	BICODE CA	Show the top 50	Results Analysi Results.	is CAB ADS 🞑 🔄	Record count	
	CABICODES			Minimum record c	ount (Threshold): 2		Selected field	
	Document Type	a 💌						
	Analyze							
	Use the checkbox Note: The number	es below to view the records. • of records displayed may be greater than the listed Record Count						
	if the original set c	ontained more records than the number of records analyzed.						
	View Records	Field: Descriptor	Record Count	% of 500	Bar Chart	Save Analysis D	ata to File	
		PIGS	499	99.8000 %				
		DIAGNOSIS	101	20.2000 %				
		DISEASE PREVALENCE	90	18.0000 %				
		EPIDEMIOLOGY	87	17.4000 %				
		GENES	86	17.2000 %				
		DISEASE CONTROL	64	12.8000 %				
		ARTERIVIRUS	62	12.4000 %	-			
		PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS	62	12.4000 %				
		PIGLETS	60	12.0000 %	_			
		PORCINE CIRCOVIROSES	60	12.0000 %	-			
			59	11.8000 %	-			
			59	11.6000 %	-			
		IMMUNE RESPONSE	58	11.6000 %	_			
		VACCINATION	58	11.6000 %	-			
		DIAGNOSTIC TECHNIQUES	56	11.2000 %	-			
Tennin	-4							<u> </u>
rermin								
🦉 (emarrer			S Courrier	entran 🛛 🔟 2 Microsof	it O 🍸 🥹 Results /	analysi FR 📶 🖏	🛛 🖉 🖉 15:35



To select all disease/pathogen descriptors relevant to the research field searched, the complete table of descriptors is directly copied to an Excel spread sheet. From this list; the animal disease specialist (1) selects the specific descriptors of diseases and pathogens, then (2) groups together those descriptors pertaining to a same disease entity.

The selection and reorganization of controlled descriptors is based on the expertise of the animal health specialist and should be scientifically sound and accountable. Whereas this procedure of descriptors selection secures exhaustiveness and adequate prioritization, the grouping of descriptors (pointing to a same entity) may be discussed in some instances. The total traceability of the procedure (search queries are included in the final EMIDA data base) lends itself to corrections and further improvements, if deemed necessary. The reference manual used to check the relevance of disease/pathogen grouping is the Merck Veterinary Manual, a free comprehensive electronic reference for animal diseases information.

Once selected and reorganized the list of descriptors appears as shown (for pigs diseases covered by CC:LL821)

PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME, ARTERIVIRUS, PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS

SALMONELLA, SALMONELLA TYPHIMURIUM, SALMONELLOSIS, SALMONELLA CHOLERAESUIS, SALMONELLA ENTERITIDIS MYCOPLASMA HYOPNEUMONIAE, MYCOPLASMA PORCINE CIRCOVIRUS, PORCINE CIRCOVIRUSES

SWINE FEVER, SWINE FEVER VIRUS, CLASSICAL SWINE FEVER VIRUS, AFRICAN SWINE FEVER VIRUS, AFRICAN SWINE FEVER ACTINOBACILLUS PLEUROPNEUMONIAE, PLEUROPNEUMONIA

STREPTOCOCCUS SUIS

FOOT AND MOUTH DISEASE, APHTHOVIRUS, FOOT-AND-MOUTH DISEASE VIRUS PASTEURELLA MULTOCIDA, ATROPHIC RHINITIS

AUJESZKY VIRUS, AUJESZKY'S DISEASE, SUID HERPESVIRUS 1, PSEUDORABIES VIRUS, VARICELLOVIRUS ESCHERICHIA COLI

BRACHYSPIRA HYODYSENTERIAE, SWINE DYSENTERY, BRACHYSPIRA

SWINE INFLUENZAVIRUS, AVIAN INFLUENZA, AVIAN INFLUENZAVIRUS, INFLUENZA A, INFLUENZAVIRUS A HAEMOPHILUS PARASUIS, GLASSERS DISEASES PORCINE PARVOVIRUS, INFECTIOUS INFERTILITY HEPATITIS E VIRUS, HEPATITIS E BORDETELLA BRONCHISEPTICA

CAMPYLOBACTER COLI, CAMPYLOBACTER, CAMPYLOBACTER JEJUNI, CAMPYLOBACTERIOSIS

TRANSMISSIBLE GASTROENTERITIS VIRUS, WASTING DISEASE, PORCINE CORONAVIRUS

LEPTOSPIROSIS, LEPTOSPIRA, LEPTOSPIRA INTERROGANS

LAWSONIA (BACTERIA), PROLIFERATIVE ENTEROPATHY, PROLIFERATIVE ILEITIS, LAWSONIA INTRACELLULARIS YERSINIA ENTEROCOLITICA ROTAVIRUS

STAPHYLOCOCCUS AUREUS, STREPTOCOCCUS, MASTITIS, AGALACTIA BRUCELLOSIS, BRUCELLA, BRUCELLA SUIS CLOSTRIDIUM PERFRINGENS, CLOSTRIDIUM



TUBERCULOSIS, MYCOBACTERIUM AVIUM, MYCOBACTERIUM BOVIS ERYSIPELOTHRIX RHUSIOPATHIAE MYCOTOXICOSES, FUSARIUM, FUMONISINS, TIAMULIN JAPANESE ENCEPHALITIS VIRUS

ENTEROCOCCUS FAECIUM, ENTEROCOCCUS, ENTEROCOCCUS FAECALIS STAPHYLOCOCCUS HYICUS

TAENIA SOLIUM, CYSTICERCOSIS, TAENIASIS, TAENIA SAGINATA, ECHINOCOCCUS GRANULOSUS, NEUROCYSTICERCOSIS, ECHINOCOCCOSIS, TAENIA, ECHINOCOCCUS, CYSTICERCI, EUCESTODA, CESTODE INFECTIONS, METACESTODES, CYSTIC ECHINOCOCCOSIS

ASCARIS SUUM, ASCARIASIS, ASCARIS, ASCARIS LUMBRICOIDES TOXOPLASMA GONDII, TOXOPLASMOSIS

TRICHINOSIS, TRICHINELLA SPIRALIS, TRICHINELLA, TRICHINELLA NATIVA, TRICHINELLA PSEUDOSPIRALIS

COCCIDIOSIS, ISOSPORA SUIS, EIMERIA, COCCIDIA, ISOSPORIASIS, TACHYZOITES, ISOSPORA, EIMERIA DEBLIECKI, EIMERIA PERMINUTA, EIMERIA SUIS, ENCEPHALITOZOON INTESTINALIS CRYPTOSPORIDIUM, CRYPTOSPORIDIOSIS

TRICHURIS SUIS, TRICHURIS, TRICHURIASIS

SARCOPTES SCABIEI, ECTOPARASITOSES, SCABIES, ECTOPARASITES, PHTHIRAPTERA, DIPTERA, CTENOCEPHALIDES FELIS, DEMODEX, DERMATITIS

OESOPHAGOSTOMUM, OESOPHAGOSTOMUM DENTATUM, METASTRONGYLUS, STRONGYLOIDES RANSOMI, STRONGYLIDAE, HYOSTRONGYLUS, HYOSTRONGYLUS RUBIDUS

TOXOPLASMA , CRYPTOSPORIDIUM PARVUM, GIARDIASIS, GIARDIA, GIARDIA DUODENALIS

TICKBORNE DISEASES, RHIPICEPHALUS SANGUINEUS, METASTIGMATA, VECTOR-BORNE DISEASES, CULEX TRITAENIORHYNCHUS, BABESIA, BABESIOSIS, ANAPLASMA, EPERYTHROZOON, EPERYTHROZOON SUIS

SARCOCYSTIS, SARCOCYSTIS MIESCHERIANA

SCHISTOSOMA JAPONICUM

BALANTIDIUM COLI

LUNGWORMS

With minor modifications, the procedure is run for each CC (LL821, LL822 and LL823) and for each food animal category (pigs, ruminants, poultry, horses, rabbits, bees, fish and shellfish). All the sets of controlled descriptors are now supposed to cover in a satisfactory manner the totality of the diseases and pathogenic agents that can be searched and retrieved in the scientific literature of the domain.

These species-specific sets of descriptors can now be used as topics for further search, using different data bases, particularly WoS. One the main asset of this procedure is that all these descriptors are present in the CAB thesaurus.

5. Running search queries with controlled descriptors on WoS platform

Research institutions having produced publications on each searched scientific topics since 2004 can now be identified using WoS. In contrast to the CAB; the WoS platform can sort and analyse



references using addresses, which include name of <u>countries</u> and of <u>research organisations</u>, this for of all the authors of the source documents sorted. We have decided to include all the documents produced in 2008 at the date of the searches; so as not to exclude relevant newcomers.

For example, the search for references on the porcine reproductive and respiratory syndrome of pigs is instructed as follows

TS=("PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME" OR ARTERIVIRUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=(PIGS OR WILD PIGS OR PIGLETS OR SOWS OR SWINE DISEASES) AND PY=2004-2008

Where "TS" stands for topic and " PY" for years of publication. Note that the disease complex is described by the set of descriptors identified in the previous step, and the production animal category also by a comprehensive set of descriptors picked out from the CAB thesaurus. A WoS advanced boolean search is launched as shown in the screen below :

Web of Knowledge [v.4.1] - Web of Science Advanced Search - Mozilla Firefox		
er Édition_Affichage_Historique_Marque-pages_Outils_?		
🔹 🗼 👻 😢 🏠 🗋 http://apps.isiknowledge.com/WOS_AdvancedSearch_input.do?product=WOS8SID=N133Cp8G2ceo9jc7gh@&search_mode=A 🔽 🕨	G Google	
émarrage 🔂 Dernières nouvelles (en) 🦸 DA: History: General a		
NTRANET de l'unité IASP - 💿 🔄 Liste alphabétique des bases de 💽 📄 CABICODE CABI [v,1] 💿 💽 Results Analysis CAB Abstracts .	🖸 🗋 ISI Web of Knowledge	e (v.4.1 🖸
I Databases Select a Database Web of Science Additional Resources		
Search Cited Reference Search Structure Search Advanced Search Search History Marked List (0)		
Web of Science®		
Advanced Search, Use 2-character tags, Boolean operators, parentheses, and set references to create your query. Results appear in the Search	Field Tage	Booleane
History at the bottom of the page.	TS=Topic	AND
Example: TS=(nanotub*SAME carbon) NOT AU=Smalley RE #11 NOT #2 more examples User the thronal	TI=Title AU=Author Q	OR NOT
TS="PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME" OR APTERIVIDUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY	GP=Group Author Q SO=Publication Name Q	SAME
SYNDROME VIRUS") AND TS=(PIGS OR WILD PIGS OR PIGLETS OR SOWS OR SMINE DISEASES) AND PY=2004-2008	PY=Year Published AD=Address	
	DG=Organization SG=Suborganization	
	SA=Street Address CI=City	
Search	PS=Province/State CU=Country	
Current Limits: (Change Limits and Refined	ZP=Zip/Postal Code	
Timespan-All Years. Databases=Science Citation Index Expanded (SCI-EXPANDED); Social Sciences Citation Index (SSCI); Arts & Humanities Citation Index (A&HCI); Index Cherr	icus	
(IC); Current Chemical Reactions (CCR-EXPANDED [back to 1840])		
Restrict results by any or all of the options below:		
All languages A All document types		
Afrikaans Abstract of Published Item Arabic V Afrikaans V Afrikaan		
Search History		
	Combine Sets Dela	ate Sets
Set Results Open Saved History		lect All
iné		
서émarrer 🖉 🖉 😋 🖏 🖄 🖄 🛤 🗮 🗃 🚔 🖓 🗒 🖬 🚔 승 🛤 🖾 🥚 🖉 🎧 🕼 Cour 🖾 2 M 🔸 🤤 151 👘 CAD	Micro FR 📶 🕏	V 🖻 📿 16

The first results of such a search appears as follows :



🗐 IS	Web of Knowledge [v.4.	1] - Web of Science - Mozilla Firefox	
Eichie	er Éditio <u>n A</u> ffichage <u>H</u> istorique <u>M</u>	tarque-pages Qutils 2	10
4	• 🛶 - 💽 💿 🏠 🖻 http://a	pps.isknowledge.com/summary.do?product=WOS&doc=1&qid=2&SID=N133Cp&G2ceo9jc7gh@&search_mode=Advar 💌 🕨 🔀 🔍 Google	Q
n 🖗	emarrage 🔂 Dernières nouvelles (en)	頂 DA: History: General a	
20 I	VTRANET de l'unité IASP -	🗋 Liste alphabétique des bases de 📄 📄 CABICODE CABI [v.1] 👘 🤉 Results Analysis CAB Abstracts 🔅 📄 ISI Web of Knowledge [v.4.1	
	All Databases Select a Databases Select a Databases Search Cited Reference Search Web of Science® Results TS=(*PORCINE REPI		
	TS=(PIGS OR WILD Timespan=All Years, Datab	PIGS OR PIGLETS OR SOWS OR SWINE DISEASES) AND PY=2004-2008 assesSCLEXPANDED, SSCI, ASHO, IC, CCR-EXPANDED [back to 1840].	
e De	Refine Results	Print E-mail Add to Marked List Save to EndNoi2 Web more options Analyze Results Im Create Citation Report	
Hide Reli	Search within results for Search Search within results for Search Subject Areas Refine VETERINARY SCIENCES (150) MIMUNDLGY (50) VISEN LGY (50)	Title: Failure of an inactivated vaccine against porcine reproductive and respiratory syndrome to protect gilts against a heterologous challenge with with PRRSV Author(s): Scortti M, Prieto C, Alvarez E, et al. Source: VETERINARY RECORD Volume: 161 Issue: 24 Pages: 809-813 Published: ~2007 Times Cited: 0 Os:F-X	
	MICROBIOLOGY (24) MEDICINE, RESEARCH & EXPERIMENTAL (16) more	Title: Construction and immunogenicity of pseudotype baculovirus expressing GP5 and M protein of porcine reproductive and respiratory syndrome virus Author(s): Wang SP, Fang LR, Fan HY, et al. Source: VACCINE Volume: 25 Issue: 49 Pages: 8220-8227 Published: ~2007 Times Cited: 0 Osser.X Full Text	
	ARTICLE (208) REVIEW (7) EDITORIAL MATERIAL (2) NEWS ITEM (1) Authors	3. Title: Emergence of a highly pathogenic porcine reproductive and respiratory syndrome virus in the Mid-Eastern region of China Author(s): U YF, Wang XL, Bo KT, et al. Source: VETERNHARY JOURNAL Volume: 174 Issue: 3 Pages: 577-584 Published: ~2007 Times Cited: 1 5:FrX Full Text	
Term	▶ Source Titlee	4. Title: Use of an experimental model to test the efficacy of planned exposure to live porcine reproductive and respiratory syndrome virus	~
2	démarrer 🔰 🙆 🛤 🗅 🔣 🗉	1 🕅 🖻 🗉 🗑 🐂 🖗 🖉 🖾 🍣 🗧 🔟 C. 🛛 🗳 Cour 🔟 2 M 🔮 151 🖆 CAD 🗷 Micro FR 🚮 💥 V 🖻 Q. 16	5:12

And analysed statistically according to countries :

esults Analys	is Web Of Scienc	e - Mozi	lla Firefox						
er Éditio <u>n A</u> ffich	age <u>H</u> istorique <u>M</u> arque-pa	iges <u>O</u> utils	2						
- 🗼 - 🥑 😣) 🏠 📄 http://pcs.isiknov	vledge.com/a	analyze/ra.cgi				• • G	Google	Q
émarrage 🔂 Derni	ères nouvelles (en) 🗴 DA: H	listory: Gener	ral a						
NTRANET de l'unité	IASP -	Iphabétique	des bases de (CABI [v.1]	Results Analysis 0	AB Abstracts 🔝 📑	Results Analysis We	b Of 🖸
<<< Back to results	list			Анар	Ze Nesulis				
218 records. TS=	PORCINE REPRODUCTIVE	AND RESPIR.	ATORY SYNDRO	ME" OR ARTERIVIRUS	OR "PORCINE REPRODUCT	IVE AND RESPIRATOR	Y SYNDROME VIRUS")	AND TS=(PIGS OR WILD PIC	38
OR PIGLETS OR S	SOWS OR SWINE DISEASES)	AND PY=2004	-2008						
Rank	the records by this field:		Ana	lyze:		Set display options:		Sort by:	
Author	^	upTo	500 🔽 rec	ords.	Show the top 500 🔽 F	lesults.		Record count	
Country/Territor	40				Minimum record count (T	hreshold): 2		Selected field	
Institution Name									
Analyze									
Use the checkboxe Note: The number	is below to view the records. of records displayed may be an	aatar than the	listed Record Co	aunt					
if the original set co	ontained more records than the	number of re	cords analyzed.	Jun					
View Records	Field: Country/Territory	Record Count	% of 218	Bar Chart	Save Analysis Data to File)			
	USA	117	53.6697 %						
	PEOPLES R CHINA	23	10.5505 %						
	CANADA	19	8.7156 %						
	SPAIN	16	7.3394 %						
	GERMANY	8	3.6697 %						
	JAPAN	8	3.6697 %	1 C C					
	DENMARK	7	3.2110 %						
	BELGIUM	6	2.7523 %	1.0					
	POLAND	6	2.7523 %	1.0					
	CHILE	5	2.2936 %	1.0					
	NETHERLANDS	5	2.2936 %	1					
iné									
démarrer	@ 24 G 🖬 🖬 🗞 🖎	😐 🛛 🗿	👝 😳 😜 🗖	🗟 😂 🖻 🕅 (2 🛛 🎯 Cour 🕅 :	2 M 👻 😢 Resul	C:\D	Micro FR 📶 🛒 🕅	V 🖻 🙆 16:1-

and to research organizations:



1	e - Mozilia Pirelox						_
Edition Affichage Historique Marque-pa	iges <u>O</u> utils <u>?</u>						
🔶 + 🥑 💿 🚮 🗋 http://pcs.isiknov	vledge.com/analyze/ra.cgi				•	▶ G• Google	
arrage 🔝 Dernières nouvelles (en) ß DA: H	listory: General a						
RANET de l'unité IASP - 💿 📄 Liste a	Iphabétique des bases de 🖂 📔 C	ABICODE C	ABI [v.1]	🛛 🗋 Re	sults Analysis CAB Abstracts	🖸 🗋 Results Analysis Web O)f
		\sim					
< Back to results list		Analyz	e Results				
218 records. TS=("PORCINE REPRODUCTIVE / OR PIGLETS OR SOWS OR SWINE DISEASES) /	AND RESPIRATORY SYNDROME" OR AR AND PY=2004-2008	TERIVIRUS 0	R "PORCINE REF	RODUCTIVE AI	ND RESPIRATORY SYNDROME	VIRUS") AND TS=(PIGS OR WILD PIGS	
Rank the records by this field:	Analyze:			Set d	lisplay options:	Sort by:	
Language 🔨	upTo 500 💌 records.		Show the top 5	00 🔽 Result:	3.	Record count	
Publication Year Source Title			Minimum recor	l count (Thresh	old): 2	Selected field	
Subject Area 💌							
Analyze							
loo the sheel deeves heleve to view the records							
lse the checkboxes below to view the records. ote: The number of records displayed may be gr	eater than the listed Record Count						
Jee the checkboxes below to view the records. ote: The number of records displayed may be gr the original set contained more records than the	eater than the listed Record Count number of records analyzed.	Record	% of				
Ise the checkboxes below to view the records. tote: The number of records displayed may be gr the original set contained more records than the <u>View Records</u>	eater than the listed Record Count number of records analyzed. Field: Institution Name	Record Count	% of 218	Bar Chart	Save Analysis Data to File		
Jee the checktooxes below to view the records. otie: The number of records displayed may be gr the original set contained more records than the <u>View Records</u>	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA	Record Count 53	% of 218 24.3119 %	Bar Chart	Save Analysis Data to File		
See the checkboxes below to view the records. Net "The number of records displayed may begin the original set contained more records than the <u>View Records</u>	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV	Record Count 53 17	% of 218 24.3119 % 7.7982 %	Bar Chart	Save Analysis Data to File		
See the checkbores below to view the records. where The number of records displayed may begin the original set contained more records than the View Records	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV UNIV ILLINOIS	Record Count 53 17 12	% of 218 24.3119 % 7.7982 % 5.5046 %	Bar Chart	Save Analysis Data to File		
Ise the checkboxes below to were the records. were the number of records displayed may begin the original set contained more records than the <u>View Records</u> 	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV UNIV ILLINOIS S DAKOTA STATE UNIV	Record Count 53 17 12 9	% of 218 24.3119 % 7.7982 % 5.5046 % 4.1284 %	Bar Chart	Save Analysis Data to File		
Ise the checkboxes below to were the records. where the number of records displayed may begin the original set contained more records than the view Records 	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV UNIV ILLINOIS S DAKOTA STATE UNIV USD ARS	Record Count 53 17 12 9 9	% of 218 24.3119 % 7.7982 % 5.5046 % 4.1284 % 4.1284 %	Bar Chart	Save Analysis Data to File		
See the checkboxes below to view the records. Net The number of records displayed may begin the original set contained more records than the view Records 	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV UNIV ILLINOIS S DAKOTA STATE UNIV USDA ARS UNIV GUELPH	Record Count 53 17 12 9 9 9 8	% of 218 24.3119 % 7.7982 % 5.5046 % 4.1284 % 4.1284 % 3.6697 %	Bar Chart	Save Analysis Data to File		
See the checkboxes below to view the records. Net The number of records displayed may begin the original set contained more records than the (View Records)	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNESOTA IOWA STATE UNIV UNIV ILLINOIS S DAKOTA STATE UNIV USDA AS UNIV GUELPH UNIV RERASKA	Record Count 53 17 12 9 9 9 9 8 8 8	% of 218 24.3119 % 7.7982 % 5.5046 % 4.1284 % 4.1284 % 3.6697 % 3.6697 %	Bar Chart	Save Analysis Data to File		
Ise the checkboxes below to were the records. where the number of records displayed may begin the original set contained more records than the view Records 	eater than the listed Record Count number of records analyzed. Field: Institution Name UNIV MINNEBOTA IOWA STATE UNIV UNIV STATE UNIV USDA ARS UNIV GUELPH UNIV NEBRASKA UNIV COMPLUTENSE MADRID	Record Count 53 17 12 9 9 9 8 8 8 8 7	% of 218 24.3119 % 7.7982 % 5.5046 % 4.1284 % 3.6697 % 3.6697 % 3.2110 %	Bar Chart	Save Analysis Data to File		

Using the "refine" function one can display the table of institutions working on a specific disease or pathogen in specified countries, and their respective number of publications during the period. As an example, here is for Spain the list of organizations working on the porcine reproductive and respiratory syndrome of pigs :

Edition	sults Analysis	Web Of Science -	Mozilla I	Firefox				
Control of the control of t	Edition Affichage	 Historique Marque-pages 	Outils ?					
marage Demières nouvelles (en) D DA: History: General a TRANET de l'Unité IASP - CAB sightadétique des bases de CABLOCCE - CABI [V:1] Results Analysis - CAB Abstracts Results Results Analysis - CAB Abstracts Results Results Results Analysis - CAB Abstracts Results Resourd Results Results Results Results Resou	🕪 • 🕑 😒 •	😭 🕒 http://pcs.isiknowledg	je.com/analyz	:e/ra.cgi		Ŧ	Google	
TREAMET de l'unité 14SP- Luste alphabétique des bases de CABICODE - CABI (v.1) Results Analysis - CAB Abstracts Results Analysis - Web 01 CARACTOR SUMPLY Analyze Results Analyze Results Analyze Results I records. TS-(*CORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME' OR ARTERMIRUS OR *PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS') AND TS=(PIGS OR WILD PIGS OR PIGLETS OR SOWNS OR SOWNS DIBLASES) AND PY=2008-2008 Set to records by this field: Image: Set display options: Set display options: Set to record so to this field is the record so to this field is the record so to this field is the record so to this field field Image: Set display options: Set displ	marrage 🔝 Dernière	s nouvelles (en) 🗴 DA: Histor	y: General a					
Analyze Results Search to results list Analyze Results Image: Counting SPC ORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME 'OR ARTERMIRUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS') AND TS=(PIOS OR WILD PIOS ON Analysis): Countries/Territories=(SPAIN) Teack the records by this field: Analyze: Set display options: Sort by: Country Territories=(SPAIN) Show the top 500 Results. Or Record count Country Territories=(SPAIN) Show the top 500 Results. Or Record count Description: Sort by: Or Record count Description: Show the top 500 Results. Or Record count Description: Show the top 500 Results. Or Record count Description: Show the top 500 Results. Or Record count Use the checkboxes below to view the records. Mater than the listed Record Count The number of records. Note: The number of records analyzed UNIV COMPLUTENSE MADRID 7 43,7500 % ONIV AUTONOMA BARCELONA 6 37,500 % ONIV AUTONOMA BARCELONA 6 37,500 % ONIV AUTONOMA BARCELONA 6 37,	TRANET de l'unité IAS	SP - 🔹 📄 Liste alpha	oétique des ba	ases de 💽	CABICODE C	ABI [v.1] 🔄 📄 Results Analysis CAB Abstract	s 📴 🗋 Results Analysis Web Of	(
Celeants test Analyze Results Analyze Results Seconds. T5e-(PORCOLE REPRODUCTIVE AND RESPIRATORY SYNDROME" OR ARTERNIRUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND T5e-(PIOS OR WILD PIOS OR Analysis: Countrities/Territories=(SPAIN) Image: Countrities/Territories=(SPAIN) Image: Set display options: Set display opting set display options: Set display opting set display o		Ŭ						
If secords. TS=_CPORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME* OR ARTERIVIRUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=_(PIOS OR WILD PIOS OR WILD PIOS OR SWINE DISEASES) AND PY=2004-2008 Analysis: Collective SPAND) OPECATION OF SPAND PY=2004-2008 Analysis: Collective SPAND) Analysis: Collective SPAND Analysis: Collective SPAND OPECATION OF SPAND PY=2004-2008 Analysis: Collective SPAND OPECATION OF SPAND PY=2004-2008 Analysis: Set of Spand Py State Spand Py S	<< Back to results list	t			Analy	ze Results		
If records. TS=CPORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME" OR ARTERIVIRUS OR "PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=(PIOS OR WILD PIOS OR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=(PIOS OR WILD PIOS OR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=(PIOS OR WILD PIOS OR PORCINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS") AND TS=(PIOS OR WILD PIOS OR VILD PIOS OR WILD PIOS OR WI								
Rank the records by this field: Analyze: Set display options: Sort by: Author upTo 000 records. Show the top 000 results. Show the top 000 results. Peccord count CountryTerritory Document Type Image: Count of the top 000 records. Show the top 000 records. Peccord count Selected field Analyze	OR PIGLETS OR SON Analysis: Countries/	VS OR SWINE DISEASES) AND Territories=(SPAIN)	PY=2004-2008	SYNDROME OF	RARTERIVIRUS	R PORCINE REPRODUCTIVE AND RESPIRATORY STNDRUM	VIRUS") AND TS=(PIGS OR WILD PIGS	
Autor Up To 500 records. Show the top Solo Results. Immur record count Selected field Country Type Immur record count (Threshold): 2 Selected field Selected field Selected field Analyze Selected field Selected field Selected field Selected field Use the checkboxes below to view the records. Note: The number of records displayed may be greater than the listed Record Count Selected field Selected field View Records Field: Institution Name Record % of Bar Chart Save Analysis Data to File UNIV COMPLUTENSE MADRID 7 43.7500 % Immur record Save Analysis Data to File UNIV MURCIA 2 12.5000 % Immur record Save Analysis Data to File UNIV MURCIA 2 12.5000 % Immur record Save Analysis Data to File (13 Institution Name value(s) outside display options.) Save Analysis Data to File Save Analysis Data to File	Rank the	records by this field:		Analyze:		Set display options:	Sort by:	
Minimum record count (Threshold): 2 C Selected field Instruction Name Image: Count (Threshold): 2 C Selected field	Author	•	upTo 500	💌 records.		Show the top 500 💌 Results.	Record count	
	Document Type					Minimum record count (Threshold): 2	 Selected field 	
Analyze Use the checkboxes below to view the records. Note: The number of records displayed may be greater than to instead Record Count if the original set contained more records than to more records analyzed. View Records Fiekt Institution Name Record Count % of 1 Bar Chart Save Analysis Data to File UNIV COMPLUTENSE MADRID 7 43.7500 % Image: Crees A 2 12.5000 % Image: Crees A 2 12.5000 % UNIV COMPLUTENSE MADRID 7 43.7500 % Image: Crees A 2 12.5000 % Image: Crees A 2 Image: Crees A 2 12.5000 % Image: Crees A 2 Ima	Institution Name	2						
Use the checkboxes below to view the records. Note: The number of records displayed may be greater than the listed Record Count if the original set contained more records sinalyzed. View Records VIEW MARCH Provided Market Count VINIV COMPLUTENSE MADRID 7 43.7500 % UNIV COMPLUTENSE MADRID 7 43.7500 % UNIV AUTONOMA BARCELONA 6 37.5000 % CRESA 2 12.5000 % UNIV MURCIA 2 12.5000 % VINIV MURCIA 2 12.5000 % CRESA 12 12.5000 % CRESA 12 12.5000 % VINIV MURCIA 2 12.5000 % CRESA 12 12.5000 % VINIV MURCIA 12 12.5000 % VINIV VINIV VINIV VINIV VINIV VINI	Analyze							
Use the checkbores below to view the records. Note: The number of records displayed may be greater than the listed Record Count if the original set contained more records than the number of records analyzed. View Records VIIV COMPLUTENSE MADRID 7 43.7500 % UNIV COMPLUTENSE MADRID 7 43.7500 % UNIV AUTONOMA BARCELONA 6 37.5000 % UNIV AUTONOMA BARCELONA 6 37.5000 % CRESA 2 12.5000 % UNIV MURCIA 32 30.5000 % UNIV MURCIA 32 30.5000 % UNIV MURCIA 32 30.5000 % UNIV MURCIA 30.								
View Records Field: Institution Name Record Count % of 16 Bar Chart Save Analysis Data to File Image:	Use the checkboxes b Note: The number of r if the original set conta	elow to view the records. ecords displayed may be greate alned more records than the nun	r than the listed	d Record Count analyzed.				
UNIV COMPLUTENSE MADRID 7 43,7500 % UNIV AUTONOMA BARCELONA 6 37,500 % CCRESA 2 12,500 % UNIV MURCIA 2 12,500 % UNIV MURCIA 2 12,500 % Field: Institution Name Record Count %	View Records	Field: Institution Name	Record Count	% of 16	Bar Chart	Save Analysis Data to File		
UNIV AUTONOMA BARCELONA 6 37.5000 % CRESA 2 12.5000 % UNIV MURCIA 2 12.5000 % UNIV MURCIA 2 12.5000 % View Records Field: Institution Name Record (19 Institution Name value(s) outside display options.) Save Analysis Data to File	1U 🔲	NIV COMPLUTENSE MADRID	7	43.7500 %				
CRESA 2 12.5000 % UNIV MURCIA 2 12.5000 % Field: Institution Name Record Count % of Count Bar Chart Save Analysis Data to File	1U 🗌	IV AUTONOMA BARCELONA	6	37.5000 %				
UNIV MURCIA 2 12.5000 % View Records Field: Institution Name Record Count % of 6% of 16% Bar Chart Save Analysis Data to File		CRESA	2	12.5000 %				
View Records Field: Institution Name Record Count % of 16 Bar Chart Save Analysis Data to File (19 Institution Name value(s) outside display options.)		UNIV MURCIA	2	12.5000 %	-			
(19 Institution Name value(s) outside display options.)		Field: Institution Name	Record	% of 16	Bar Chart	Save Analysis Data to File		
A	View Records		Count					
	View Records	(19 Institution Name value(s)	outside displa	ay options.)				

For EMIDA this procedure has been applied to all countries belonging to the European continent, plus Turkey, as a EC candidate country, and Israël as associate partner of the project: AUSTRIA, BELARUS, "CZECH REPUBLIC", HUNGARY, MOLDOVA, POLAND, SLOVAKIA, UKRAINE, ESTONIA, LATVIA, LITHUANIA, DENMARK, FINLAND, ICELAND, NORWAY, SWEDEN, ALBANIA, "BOSNIA-



HERCEGOVINA", BULGARIA, CROATIA, "REPUBLIC OF MACEDONIA", ROMANIA, SLOVENIA, YUGOSLAVIA, GIBRALTAR, GREECE, ITALY, MALTA, MONACO, PORTUGAL, "SAN MARINO", SPAIN, VATICAN, ANDORRA, BELGIUM, "IRISH REPUBLIC", UK, FRANCE, GERMANY, LIECHTENSTEIN, LUXEMBOURG, NETHERLANDS, SWITZERLAND, RUSSIA, TURKEY, ISRAEL

6. Extraction of information on European organisations from the WoS and integration in the EMIDA Excel data base

This information is extracted from the WoS tabulated outputs above, only for those organizations with at least 2 papers during the period surveyed. As the organizations of all the authors of a given source document are identified in the WoS data base, this threshold of 2 papers appear reasonable to select organizations with a significant impact in the topic searched.

In one step, the selected WoS tabulated data are copied into an Excel spreadsheet containing eventually the following fields upon completion of the procedure:

- 1. Search query
- 2. Pathogen/ Disease
- 3. Animal
- 4. Source
- 5. Date search
- 6. Number of publications
- 7. Institution Name
- 8. Town
- 9. Country
- 10. Country code (3 Letters)
- 11. Country code (2 Letters)

7. Standardisation of organisation names

The WoS platform has a standardized procedure for abbreviating institution names of all co-authors of a given publication. The developed name can therefore be easily restored.

However, a same institution can be referred to by two or more different names names in different publications. It was therefore essential to standardise these names so that one institution is referred to by one and single name in the database.

For all the countries that participate in EMIDA, this standardization was entrusted to partner representatives of the project. For the other countries, this task was done by the designer of the database.

8. Specific CAB search for extraction of data on transversal fields

Complementary search on specific transversal fields is possible in the CAB abstracts data base, due to the existence of CABICODES (CC), which indicate the broad subject areas of the publications, and/or of controlled descriptors (DE) to complement CABICODE information.

Search procedures

Considering the opportunities offered by these CAB functionalities, the five following transversal fields were searched, with their associated search query:

Immunology of food animal species

CC=LL650 AND DE=...(animal).....



Genetic resistance to infectious diseases CC=(LL240 AND (HH600 OR LL821 OR LL822 OR LL650)) AND DE =....(animal).....

Quantitative epidemiology of infectious diseases

CC=(LL821 OR LL8222 OR LL823 OR LL800 OR YY700 OR EE117 OR LL650) AND DE=("RISK FACTORS" OR "MATHEMATICAL MODELS" OR "SIMULATION MODELS" OR "RISK ANALYSIS" OR "ECONOMIC IMPACT" OR "STOCHASTIC MODELS" OR "ECONOMIC ANALYSIS" OR "COST BENEFIT ANALYSIS" OR "TEMPORAL VARIATION" OR "AGRICULTURAL ECONOMICS")

Veterinary vaccines

DE=vaccine* AND CC=((LL821 OR LL822 OR LL823 OR LL650 OR HH600 OR LL882) NOT VV*) AND DE=...(animal)

Resistance to antibiotics CC=(HH410 AND LL821)

Resistance to antiparasitic drugs CC=(HH410 AND (LL822 OR LL823))....

Wildlife parasitic diseases

CC=(YY700 AND (LL822 OR LL823))

Wildlife bacterial diseases CC=(YY700 AND LL821)

Restriction to European countries

To this effect the names of the relevant European countries is introduced in the address field search of the CAB query. Here is the list of the countries: AUSTRIA, BELARUS, "CZECH REPUBLIC", HUNGARY, MOLDOVA, POLAND, SLOVAKIA, UKRAINE, ESTONIA, LATVIA, LITHUANIA, DENMARK, FINLAND, ICELAND, NORWAY, SWEDEN, ALBANIA, "BOSNIA-HERCEGOVINA", BULGARIA, CROATIA, "REPUBLIC OF MACEDONIA", ROMANIA, SLOVENIA, YUGOSLAVIA, GIBRALTAR, GREECE, ITALY, MALTA, MONACO, PORTUGAL, "SAN MARINO", SPAIN, VATICAN, ANDORRA, BELGIUM, "IRISH REPUBLIC", UK, FRANCE, GERMANY, LIECHTENSTEIN, LUXEMBOURG, NETHERLANDS, SWITZERLAND, RUSSIA, TURKEY, ISRAEL

Retrieval of institutions addresses

In contrast to WoS, it is not possible to sort data according to institutions names in CAB. For each specific search, all records were therefore copied in an EndNote file to be sorted by institution name. Besides, as CAB records contain the names and addresses of the first authors of papers only, we selected all institutions and not only those with at least two papers in the period searched (2004-2008).

9. Translation of spreadsheet in web-based searchable database

The output spreadsheet was then turned into a web searchable database after importation using SQL language (Yeoconcept ®). The resulting search interface was being made available on the EMIDA website in April 2009. Simplified selection and sorting of information can be performed starting from the following entries : Topics, Institutions, Countries. Institutions can also be located on maps.



10. Potentials and limitations of the methodology used

The information on recent research output was drawn from the ISI Web of Knowledge [®] platform, namely the CAB Abstracts[®] and the Web of Science[®]. Some features of our original collection process should be recalled here to point out some potentialities but also some limitations of this methodology.

(i). <u>Thematic coverage of EMIDA research field</u>. The main challenge, before any extraction of data, was to define the field of EMIDA (animal infectious diseases) through a complete set of search queries purported to exhaustively cover this field without major omissions. This challenge was made possible by two inner functionalities of the CAB Abstracts ®, a platform specialized in agronomy and animal health: (i) broad descriptors (the "cabicodes") indexing specific sub-fields of animal health and (ii) statistical analysis of specific controlled descriptors on diseases, pathogenic agents, and animal host names, present in a specific thesaurus.

These particular assets of the CAB [®] make it unlikely (but not impossible) that major topics (i.e. topics with significant number of publications in the recent years in Europe) escaped our attention, provided that they be referenced in the source platform. However minor research topics (precisely with less than 2 records referenced in the WoS [®] during the 3.5 years period surveyed and at world scale) can obviously be absent. It could be interesting, at the stage of gap analysis, to identify missing topics and analyse if their absence in Europe can easily be accounted for.

(ii). <u>Construction of search queries and disease clustering</u>. Search queries were constructed by grouping together keywords that pertain to the same disease or "disease cluster" in a given animal category. Although performed by an animal health specialist (JDR, the author of this study), this clustering may in some cases be considered excessive, or even unjustified from a mere scientific standpoint. To fully clarify this issue and prevent any misinterpretation, all the query searches used to extract data from ISI Web of Science [®] will be accessible in the EMIDA output database. Comments on the query search formulations and on the descriptors grouping scheme will thus be possible on the part of database users to improve any next version of EMIDA output database.

(iii). <u>Transversal research fields</u>. Survey of these transversal fields was made reliable thanks to the existence, in CAB Abstracts [®], of broad descriptors covering a large part of them. The combination of the 9 transversal fields presented in Annex VII covers about 60% of the CAB records that were otherwise selected by the combination of the three "cabicodes" defining the field of infectious and parasitic diseases of animals (LL821, LL822 and LL823). In spite of what can be considered as a satisfactory coverage of applied and finalized research topics, major generic topics dealing with basic research, such as pathogenesis and the molecular dissection of pathogenic agents, are missing. The analysis of research data on these topics is essential, in particular with a view to identify new research fronts, and improved approaches should be found in the next phase of EMIDA WP2 to study more efficiently output in basic research in animal infectious diseases.

(iv). <u>Period surveyed</u>. The period of survey starts from 2004 so as to take into account recent research output up to the date of collection. For comparison, a term of 4 years corresponds roughly to the minimum time of completion of a large research programme. The majority of information originated from the ISI Web of Knowledge ® was collected in July 2008, thus covering a 4.5 years term. The specific analysis on publication sources or on co-authorship (section K and M3, respectively) theoretically include data of the full five years 2004 to 2008, having been collected during the fist trimester of 2009. In any case, the dates of extraction are indicated as required in the report, and all the comparisons



presented in this study were done from data sets collected contemporaneously. It is also worth mentioning that the information available in ISI platforms is any kind constantly evolving, even for years already completed (here 2008).

(v). <u>Record counts</u>. In its conception, the primary aim of this database is qualitative and not quantitative. Although counted records reliably reflect the volume of the research output for each entry (topic, country, institution), some features of the data collation process should be recalled to prevent any over-interpretation of such quantitative information.

→ the original "entry" of data extraction process from the WoS [®] is not the publication itself but the research institution. For any given entity (country, topic, research institution), the cumulated "number of records" corresponds to the cumulated number of lines to which this entity is associated in the EMIDA database, and not strictly to the number of publications during the period surveyed. This is due to two main reasons:

• any co-authored publication provides as many lines (or records) as the number of author institutions of the publication;

• many publications are related to both a "disease topic" and one (or even several) "transversal topics", thus generating as many lines in the database.

→ extraction in WoS has been restricted to institutions with at least two publications during the 3.5 year period of survey. However, it appears that an institution name can have several different wordings in the WoS ®, resulting in the splitting of records under these several names. Consequently, and in spite of careful attention, a few number of eligible institutions (i.e. with two publications or more) may have been overlooked.

-----END OF DOCUMENT------

