

The INCOCOA Website and Virtual Library: Improving information sharing within the cocoa research community

M.J. End¹ and C.J. Turnbull²

¹INCOCOA/INGENIC Secretariat (info@incocoa.org) and Cocoa Research Association Ltd. (michelle.end@cocoaresearch.org.uk)

²*School of Agriculture, Policy and Development, The University of Reading, Whiteknights, Reading, UK, RG6 6A*

Abstract

Good communication and awareness of current and previous research activities are important in reducing unintentional duplication or repetition of previous research and in promoting collaborative activities to the ultimate benefit of the cocoa sector. There are numerous abstracts information services available, which help researchers access recently published literature. However, there is also a wealth of “grey literature”, such as institutional reports and bulletins, newsletters, conference and meeting reports, which is either unpublished or has been published in a non-commercial form and is often more difficult to find. The INCOCOA Virtual Library initiative sets out to improve access to this literature through links to existing bibliographic resources and by encouraging the cocoa community to contribute new resources.

The INCOCOA groups are informal networks created to promote the exchange of ideas and information amongst cocoa scientists. The current groups under the INCOCOA umbrella are INAFORESTA (the international scientific group dedicated to the analysis and improvement of the relationships between people, cocoa, trees, forests and the environment), INCOPED (International Permanent Working Group for Cocoa Pests and Diseases), INCOSOM (International Group on Cocoa and Soil Management) and INGENIC (International Group for Genetic Improvement of Cocoa). The INCOCOA website (www.incocoa.org) provides a portal to each of the groups and includes a news section (Twitter Feed) and a resources section, which incorporates the Virtual Library.

The library provides direct access to material published by the INCOCOA groups and allows users to query various bibliographic databases (using a web services approach) via a single search page. The user can opt to search in one or more of these databases and will be presented with a list of matching references. Depending on the source database and copyright restrictions, access may be provided to a digital version of the paper or a link to the website hosting the database. The databases currently linked include the Penn State database (June 2016, compiled from sources including PubMed and other abstract information service providers), the Cocoa Research UK library (mainly outputs from activities and projects supported by the UK industry), the Gerritsma database (compiled mid-1990’s including many historic and grey literature sources), the ICGD library (literature on cocoa genetic resources) and the Lockwood and Campbell databases (mainly literature on entomology and insect-vectorised diseases).

Progress in expanding the range of linked databases and other developments for the INCOCOA group websites are discussed.

Introduction

Good communication and awareness of current and previous research activities are important in reducing unintentional duplication or repetition of previous research and in promoting collaborative

activities to the ultimate benefit of the cocoa sector. Access to scientific publications via the internet is of great benefit to the research community and many publications can be obtained via publishers' websites, abstracting services, online libraries and academic networking platforms such as ResearchGate and Academia. However, there is also a wealth of "grey literature", such as institutional reports and bulletins, newsletters, conference and meeting reports, which is either unpublished or has been published but is not included by the abstracting services and is consequently often more difficult to find.

Although some of this literature may date from the 1970's or even earlier, it contains a wealth of information that is potentially still of high value to today's researchers. Unfortunately, some of this material is poorly stored and is vulnerable to deterioration, or even total loss, due to humidity, mould and insect damage. Indeed, sometimes materials are neglected and even discarded since their value may not be appreciated by their current custodians. With the scanning technologies currently available, it is possible to transfer physical documents to digital images which can be securely stored and shared. Optical Character Recognition (OCR) can often be used to facilitate searching of this material for keywords or phrases. Bibliographic databases further improve access to the information since additional keywords and notes can be added, for example, to take account of changes to taxonomy and acronyms, or to translate between languages.

The INCOCOA Virtual Library initiative sets out to improve access to this literature through links to existing bibliographic resources and by encouraging the cocoa community to contribute new resources.

The INCOCOA Groups

The INCOCOA groups are informal networks created to promote the exchange of ideas and information amongst cocoa scientists. The current groups under the INCOCOA umbrella are INAFORESTA (the international scientific group dedicated to the analysis and improvement of the relationships between people, cocoa, trees, forests and the environment), INCOPED (International Permanent Working Group for Cocoa Pests and Diseases), INCOSOM (International Group on Cocoa and Soil Management) and INGENIC (International Group for Genetic Improvement of Cocoa). A new joint website for the INCOCOA groups www.incocoa.org was launched in 2013 to provide a portal to the website of each of the groups. Recent developments on the INCOCOA website include the addition of a news section (Twitter Feed) and a resources section, which incorporates the Virtual Library.

The INCOCOA Virtual Library

The library provides direct access to material published by the INCOCOA groups and allows users to query various bibliographic databases via a single search page using a web services approach. Some of these databases may be hosted on different websites/organisations but can exchange information using a common data format. The user can opt to search in one or more of these databases and will be presented with a list of matching references. Depending on the source database and copyright restrictions, access may be provided to a digital version of the paper or a link provided to the reference on the website hosting the database, where further information may be available on how to access the publication. The databases currently include:

Cocoa Bibliography: a database of over 7100 references largely sourced from PubMed and other online databases by Mark Gultinan and Siela Maximova in June 2016.

Cocoa and Chocolate Bibliography: a database of some 9000 references compiled by Wouter Gerritsma in 1997-9 as a joint initiative between University of Wageningen and the Dutch cocoa and chocolate industry.

Entomology and Virology Bibliographies: Databases of information on cocoa's pests and pathogens. Includes a bibliographic database of kindly contributed by Colin Campbell (>2000 References) and a database of nearly 3200 insects and mites associated with cacao growing including pests, their natural enemies and pollinators kindly contributed by Mike Bigger (Bigger, 2012). A database of references on CSSV and mealy bug vectors, compiled by Rob Lockwood with support from Mars is also included.

ICGD and University of Reading collection: The bibliographic database maintained by the International Cocoa Germplasm Database www.icgd.rdg.ac.uk, comprising over 800 references associated with collecting, characterising and evaluating cocoa germplasm and additional literature held at the University thanks to contributions from BCCCA, Harold Owen, Mike Thresh, A. Vernon and former staff, associates and students.

UK Cocoa & Chocolate Industry Research: The virtual library accessible through www.CocoaResearch.org.uk which includes 150 references largely resulting from projects supported by the UK Biscuit, Cake, Chocolate, Confectionery Association (to 2008), the Cocoa Research Association Ltd (CRA Ltd), The Ghana Cocoa Growing Research Association Ltd. (GCGRA Ltd) and Cocoa Research UK Ltd (CRUK Ltd). It also includes the references and linked pdf files for the 235 articles in the Cocoa Growers' Bulletin (1963 to 2000) which have been kindly contributed by Mondelez International.

Keywords

A set of keywords is being developed in order to facilitate access to the information. This is especially important where there have been changes to the taxonomy of organisms, institute names and acronyms over the years. For example, the addition of a Keyword "Witches' Broom" may help locate matching references whether the organism is referred to as *Marasmius perniciosus*, *Crinipellis perniciosus* or *Moniliophthora perniciosus* in the publication itself. Keywords are being developed to facilitate searching by country, institute, and subject. Input into this process of ensuring that the most appropriate keywords are being used is welcomed.

Next Steps

There is an ongoing process of checking the references and adding keywords where appropriate. Given the volume of information, this process will take time and users are asked for forbearance whilst this work continues.

There are further sources of information scheduled for inclusion including the Proceedings of the Conferences organised by the Cocoa, Chocolate and Confectionery Alliance (1945-1955) and other industry supported meetings. Discussions are underway with a number of individuals who have expressed an interest in making their personal collections of cocoa literature available to the initiative. We have also initiated discussions with several organisations and institutions who might be interested in linking their databases using the web server approach.

We would be grateful for further contributions of information or offers of linkages to other bibliographic databases. We are especially interested in publications which have a limited distribution or where the hard copies are at risk of disposal/damage in their current location: it would be a great service to the cocoa community, if such priority materials could be scanned and catalogued to help minimise the risk that the physical copies are lost before they can be made more accessible to a wider audience.

Acknowledgements

We are indebted to all those individuals and organisations who are contributing to this initiative. The ICGD project, the development of the INCOCOA website and virtual library is supported by CRA Ltd, a non-profit scientific research association supported by contributions from Mars, Mondelez International and ICE Futures Europe.

References

Bigger, M. 2012. A GEOGRAPHICAL DISTRIBUTION LIST OF INSECTS AND MITES ASSOCIATED WITH COCOA, DERIVED FROM LITERATURE PUBLISHED BEFORE 2012. (Pers. Comm.)