

**Joint Meeting of the Chemicals Committee and the Working Party on Chemicals,
Pesticides and Biotechnology**

**GLOBALLY HARMONISED SUBMISSION AND TRANSPORT STANDARD (GHSTS) FORMAT
SPECIFICATION**

This document is requested for declassification by 17 March 2014.

Ms Sylvie Poret
Tel.: +33 (0) 1 45 24 89 45; E-mail: Sylvie.Poret@oecd.org

JT03352448

Complete document available on OLIS in its original format

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

On 2 September 2013, the *Ad hoc* Expert Group on the Electronic Exchange of Pesticide Data (EGEEPD) endorsed in principle the draft Globally Harmonised Submission and Transport Standard (GHSTS) Schema definition and three other components of the GHSTS (the Format Specification, the Picklist XSD, and the Table of Contents XSD) providing some modifications were made. Taking into account the modifications suggested by delegates, the four GHSTS components were finalised. On 10 October 2013, the Registration Steering Group (RSG) endorsed the draft GHSTS and on 17 January 2014, the Working Group on Pesticides (WGP) endorsed the draft GHSTS.

ACTION REQUIRED: *The Joint Meeting is invited to declassify the four components of the Globally Harmonized Submission and Transport Standard (GHSTS) by 17 March 2014, amended as appropriate:*

- (i) Format specification (ghsts-specification_00-09-04); found on pages 3-138 of this document.*
- (ii) Schema definition (ghsts_00-09-02.xsd); made available on the Joint Meeting password-protected website.*
- (iii) Picklist XSD (ghsts-picklists.xsd [2013-12-06]); made available on the Joint Meeting password-protected website.*
- (iv) Table of Contents XSD (toc_00-09-01.xsd); made available on the Joint Meeting password-protected website.*

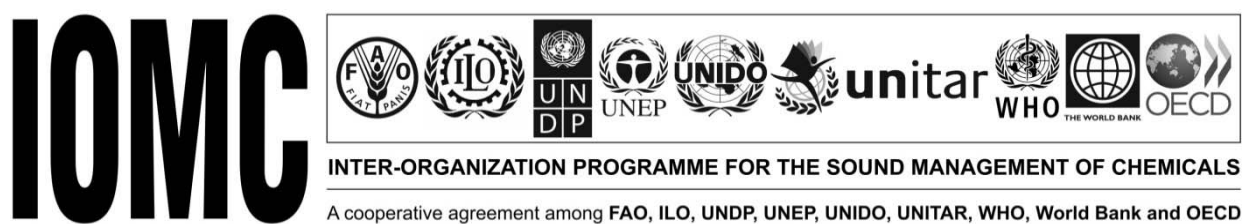
OECD Environment, Health and Safety Publications

Series on Pesticides

No. 75

**GLOBALLY HARMONISED SUBMISSION AND TRANSPORT
STANDARD (GHSTS) FORMAT SPECIFICATION**

An XML-based Interchange Format for Pesticides Registration Applications



Environment Directorate

ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT

Paris 2014

Also published in the Series on Pesticides

- No. 1 *Data Requirements for Pesticide Registration in OECD Member Countries: Survey Results* (1993)
- No. 2 *Final Report on the OECD Pilot Project to Compare Pesticide Data Reviews* (1995)
- No. 3 *Data Requirements for Biological Pesticides* (1996)
- No. 4 *Activities to Reduce Pesticide Risks in OECD and Selected FAO Countries. Part I: Summary Report* (1996)
- No. 5 *Activities to Reduce Pesticide Risks in OECD and Selected FAO Countries. Part II: Survey Responses* (1996)
- No. 6 *OECD Governments' Approaches to the Protection of Proprietary Rights and Confidential Business Information in Pesticide Registration* (1998)
- No. 7 *OECD Survey on the Collection and Use of Agricultural Pesticide Sales Data: Survey Results* (1999) [see also No.47]
- No. 8 *Report of the OECD/FAO Workshop on Integrated Pest Management and Pesticide Risk Reduction* (1999)
- No. 9 *Report of the Survey of OECD Member Countries' Approaches to the Regulation of Biocides* (1999)
- No. 10 *Guidance Notes for Analysis and Evaluation of Repeat-Dose Toxicity Studies* (2000)
- No. 11 *Survey of Best Practices in the Regulation of Pesticides in Twelve OECD Countries* (2001)
- No. 12 *Guidance for Registration Requirements for Pheromones and Other Semiochemicals Used for Arthropod Pest Control* (2001)
- No. 13 *Report of the OECD Workshop on Sharing the Work of Agricultural Pesticide Reviews* (2002)
- No. 14 *Guidance Notes for Analysis and Evaluation of Chronic Toxicity and Carcinogenicity Studies* (2002).
- No. 15 *Persistent, Bioaccumulative and Toxic Pesticides in OECD Member Countries*, (2002)
- No. 16 *OECD Guidance for Industry Data Submissions for Pheromones and Other Semiochemicals and their Active Substances* (Dossier Guidance for Pheromones and other Semiochemicals) (2003)

- No. 17 *OECD Guidance for Country Data Review Reports for Pheromones and Other Semiochemicals and their Active Substances* (Monograph Guidance for Pheromones and other Semiochemicals) (2003)
- No. 18 *Guidance for Registration Requirements for Microbial Pesticides* (2003)
- No. 19 *Registration and Work sharing, Report of the OECD/FAO Zoning Project* (2003)
- No. 20 *OECD Workshop on Electronic Tools for data submission, evaluation and exchange for the Regulation of new and existing industrial chemicals, agricultural pesticides and biocides* (2003)
- No. 21 *Guidance for Regulation of Invertebrates as Biological Control Agents (IBCA's)* (2004)
- No. 22 *OECD Guidance for Country Data Review Reports on Microbial Pest Control Products and their Microbial Pest Control Agents* (Monograph Guidance for Microbials) (2004)
- No. 23 *OECD Guidance for Industry Data Submissions for Microbial Pest Control Product and their Microbial Pest Control Agents* (Dossier Guidance for Microbials) (2004)
- No. 24 *Report of the OECD Pesticide Risk Reduction Steering Group Seminar on Compliance* (2004)
- No. 25 *The Assessment of Persistency and Bioaccumulation in the Pesticide Registration Frameworks within the OECD Region* (2005)
- No. 26 *Report of the OECD Pesticide Risk Reduction Group Seminar on Minor Uses and Pesticide Risk Reduction* (2005)
- No. 27 *Summary Report of the OECD Project on Pesticide Terrestrial Risk Indicators (TERI)* (2005)
- No. 28 *Report of the OECD Pesticide Risk Reduction Steering Group Seminar on Pesticide Risk Reduction through Good Container Management* (2005)
- No. 29 *Report of the OECD Pesticide Risk Reduction Steering Group Seminar on Risk Reduction through Good Pesticide Labelling* (2006)
- No. 30 *Report of the OECD Pesticide Risk Reduction Steering Group: The Second Risk Reduction Survey* (2006)
- No. 31 *Guidance Document on the Definition of Residue* [also published in the series on Testing and Assessment, No. 63] (2006, revised 2009)
- No. 32 *Guidance Document on Overview of Residue Chemistry Studies* [also published in the series on Testing and Assessment, No. 64] (2006, revised 2009)

- No. 33 *Overview of Country and Regional Review Procedures for Agricultural Pesticides and Relevant Documents* (2006)
- No. 34 *Frequently Asked Questions about Work Sharing on Pesticide Registration Reviews* (2007)
- No. 35 *Report of the OECD Pesticide Risk Reduction Steering Group Seminar on "Pesticide Risk Reduction through Better Application Technology"* (2007)
- No. 36 *Analysis and Assessment of Current Protocols to Develop Harmonised Test Methods and Relevant Performance Standards for the Efficacy Testing of Treated Articles/Treated Materials* (2007)
- No. 37 *Report on the OECD Pesticide Risk Reduction Steering Group Workshop "Pesticide User Compliance"* (2007)
- No. 38 *Survey of the Pesticide Risk Reduction Steering Group on Minor Uses of Pesticides* (2007)
- No. 39 *Guidance Document on Pesticide Residue Analytical Methods* [also published in the series on Testing and Assessment, No. 72] (2007)
- No. 40 *Report of the Joint OECD Pesticide Risk Reduction Steering Group EC-HAIR Seminar on Harmonised Environmental Indicators for Pesticide Risk* (2007)
- No. 41 *The Business Case for the Joint Evaluation of Dossiers (Data Submissions) using Work-sharing Arrangements* (2008)
- No. 42 *Report of the OECD Pesticide Risk Reduction Steering Group Seminar on Risk Reduction through Better Worker Safety and Training* (2008)
- No. 43 *Working Document on the Evaluation of Microbials for Pest Control* (2008)
- *Guidance Document on Magnitude of Pesticide Residues in Processed Commodities* - only published in the Series on Testing and Assessment, No. 96 (2008)
- No. 44 *Report of Workshop on the Regulation of BioPesticides: Registration and Communication Issues* (2009)
- No. 45 *Report of the Seminar on Pesticide Risk Reduction through Education / Training the Trainers* (2009)
- No. 46 *Report of the Seminar on Pesticide Risk Reduction through Spray Drift Reduction Strategies as part of National Risk Management* (2009)
- No. 47 *OECD Survey on Countries' Approaches to the Collection and Use of Agricultural Pesticide Sales and Usage Data: Survey Results* (2009)

- No. 48 *OECD Strategic Approach in Pesticide Risk Reduction* (2009)
- No. 49 *OECD Guidance Document on Defining Minor Uses of Pesticides* (2009)
- No. 50 *Report of the OECD Seminar on Pesticide Risk Reduction through Better National Risk Management Strategies for Aerial Application* (2010)
- No. 51 *OECD Survey on Pesticide Maximum Residue Limit (MRL) Policies: Survey Results* (2010)
- No. 52 *OECD Survey of Pollinator Testing, Research, Mitigation and Information Management: Survey Results* (2010)
- No. 53 *Report of the 1st OECD BioPesticides Steering Group Seminar on Identity and Characterisation of Micro-organisms* (2010)
- No. 54 *OECD Survey on Education, Training and Certification of Agricultural Pesticide Users, Trainers and Advisors, and Other Pesticide Communicators: Survey Results* (2010)
- No. 55 *OECD Survey on How Pesticide Ingredients Other than the Stated Pesticide Active Ingredient(s) are Reviewed and Regulated: Survey Results* (2010)
- No. 56 *OECD MRL Calculator User Guide* (2011)
- No. 57 *OECD MRL Calculator MRL Statistical White Paper* (2011)
- No. 58 *Report of the OECD Seminar on Pesticide Risk Reduction Strategies Near/in Residential Areas* (2011)
- No. 59 *Report of the OECD Seminar on Risk Reduction through Prevention, Detection and Control of the Illegal International Trade in Agricultural Pesticides* (2011)
- No. 60 *Guidance Document on the Planning and Implementation of Joint Reviews of Pesticides* (2011)
- No. 61 *OECD Survey on Efficacy & Crop Safety Data Requirements & Guidelines for the Registration of Pesticide Minor Uses: Survey Results* (2011)
- No. 62 *OECD Survey on Regulatory Incentives for the Registration of Pesticide Minor Uses: Survey Results* (2011)
- No. 63 *Guidance Document on Regulatory Incentives for the Registration of Pesticide Minor Uses* (2011)

- *Guidance Notes on Dermal Absorption* - only published in the Series on Testing and Assessment, No. 156 (2011)
- No. 64 *Report of the Second OECD BioPesticides Steering Group Seminar on the Fate in the Environment of Microbial Control Agents and their Effects on Non-Target Organisms* (2011)
- No. 65 *OECD Issue Paper on Microbial Contaminant Limits for Microbial Pest Control Products* (2011)
- No. 66 *Guidance Document on Crop Field Trials* [also published in the Series on Testing and Assessment, No. 164] (2011)
- No. 67 *OECD Guidance to the Environmental Safety Evaluation of Microbial Biocontrol Agents* (2012)
- No. 68 *Report of the OECD Workshop on the Development of Harmonized International Guidance for Pesticide Terrestrial Field Dissipation Studies and Crosswalk of North American & European Eco-Regions* (2012)
- No. 69 *OECD Survey on Integrity of Pesticides at the Manufacturing, Import and Distribution Stages: Survey Results* (2012)
- No. 70 *Report of the OECD Workshop on Integrated Pest Management (IPM) Strategies for the adoption and implementation of IPM in Agriculture Contributing to the sustainable use of Pesticides and to Pesticide Risk Reduction* (2012)
- No. 71 *OECD Guidance on Pesticide Compliance and Enforcement Best Practices* (2012)
- No. 72 *Report of the Third OECD BioPesticides Steering Group Seminar on Characterisation and Analyses of Botanicals for the Use in Plant Protection Products* (2013)
- No. 73 *Guidance Document on Residues in Livestock* (2013)
- No. 74 *Report of the 4th OECD Biopesticides Steering Group Seminar on Trichoderma Spp. for Use in Plant Protection Products: Similarities and Differences* (2013)

Published separately

OECD Guidance for Country Data Review Reports on Plant Protection Products and their Active Substances-Monograph Guidance (1998, revised 2001, 2005, 2006)

OECD Guidance for Industry Data Submissions on Plant Protection Products and their Active Substances-Dossier Guidance (1998, revised 2001, 2005)

Report of the Pesticide Aquatic Risk Indicators Expert Group (2000)

Report of the OECD Workshop on the Economics of Pesticide Risk Reduction (2001)

Report of the OECD-FAO-UNEP Workshop on Obsolete Pesticides (2000)

Report of the OECD Pesticide Aquatic Risk Indicators Expert Group (2000)

Report of the 2nd OECD Workshop on Pesticide Risk Indicators (1999)

Guidelines for the Collection of Pesticide Usage Statistics Within Agriculture and Horticulture (1999)

Report of the [1st] OECD Workshop on Pesticide Risk Indicators (1997)

Report of the OECD/FAO Workshop on Pesticide Risk Reduction (1995)

© OECD 2014

Applications for permission to reproduce or translate all or part of this material should be made to: Head of Publications Service, RIGHTS@oecd.org, OECD, 2 rue André-Pascal, 75775 Paris Cedex 16, France

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 34 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in eleven different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; and Safety of Manufactured Nanomaterials.** More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/ehs/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

**For this and many other Environment,
Health and Safety publications, consult the OECD's
World Wide Web site (www.oecd.org/ehs/)**

or contact:

**OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France**

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

This document has been developed by the OECD *Ad Hoc* Expert Group on the Electronic Exchange of Pesticide Data.

The main focus of the work of the OECD *Ad Hoc* Expert Group on the Electronic Exchange of Pesticide Data is to implement the recommendations of Expert Group Meeting on the Electronic Exchange of Pesticides Data held in the United States from 21 to 23 April 2008. The Expert Group investigates the possibility of harmonization in the information technology used in the pesticide regulatory process and, in particular, has focused on developing ways to harmonise methodologies for submitting documents to regulators using a common transport mechanism based on a harmonized global XML schema. In this context, the Expert Group developed an XML-based interchange format for pesticides registration applications to regulatory authorities – the Global Harmonised Submission Transport Standard (GHSTS).

This document is divided in different sections which describe the GHSTS on both a conceptual level and technical level.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology, which **agreed** that it be declassified and made available to the public.

CONTENTS

1	PREFACE	17
1.1	GHSTS Primer	17
1.1.1	<i>Scope of GHSTS</i>	17
1.1.2	<i>Background and history</i>	17
1.1.3	<i>Supporting stakeholders</i>	18
1.1.4	<i>The role of the OECD</i>	18
1.2	Scope and audience of this document	18
1.3	Organisation of the document	19
1.4	Filename and version	19
1.5	Contacts	20
2	CONTEXT OF GHSTS	21
2.1	Usage of GHSTS in the submission process	21
2.2	GHSTS and different submission Types	23
2.2.1	<i>Joint review</i>	23
2.3	Related standards	24
2.3.1	<i>Submission standards on national / regional level</i>	24
2.3.2	<i>IUCLID</i>	25
2.3.3	<i>OECD Harmonised Templates for Reporting Chemical Test Summaries (OECD Harmonised templates)</i>	25
2.4	GHSTS components	26
2.4.1	<i>GHSTS format specification</i>	26
2.4.2	<i>GHSTS submission schema definition</i>	26
2.4.3	<i>GHSTS viewer</i>	26
2.4.4	<i>GHSTS standalone builder</i>	27
2.4.5	<i>GHSTS validator</i>	27
2.4.6	<i>GHSTS Converter</i>	27
2.4.7	<i>GHSTS Picklists</i>	27
2.4.8	<i>Standard Table of Contents</i>	27
2.4.9	<i>Centrally managed components</i>	28
2.4.10	<i>Conversion tables between dossier structures</i>	29
2.4.11	<i>Lifecycle and dependencies of GHSTS components</i>	29
2.5	Management of GHSTS	30
2.5.1	<i>Information sources for GHSTS elements</i>	30
3	INTRODUCTION TO GHSTS	31
3.1	GHSTS Design Principles	31
3.1.1	<i>Parallel viewpoints</i>	31
3.1.2	<i>Parallel navigation scenarios</i>	32
3.1.3	<i>Document-based organisation</i>	32
3.1.4	<i>Integration of XML-based content</i>	33
3.1.5	<i>Adaptiveness with respect to different standard ToC templates</i>	33
3.1.6	<i>Lifecycle management</i>	33
3.1.7	<i>Persistent identifiers of information concepts</i>	33
3.1.8	<i>Information specific for regulatory authorities</i>	34
3.2	GHSTS logical concepts	34
3.2.1	<i>Logical concepts and XML syntax</i>	34
3.2.2	<i>Overview</i>	35
3.2.3	<i>Metadata</i>	36

3.2.4	<i>Dossier regulatory action</i>	39
3.2.5	<i>Submission</i>	39
3.2.6	<i>Dossier</i>	39
3.2.7	<i>Table of Contents</i>	39
3.2.8	<i>Legal entity</i>	40
3.2.9	<i>Product</i>	41
3.2.10	<i>Substance</i>	41
3.2.11	<i>File</i>	42
3.2.12	<i>Document</i>	42
3.2.13	<i>Document Family</i>	44
3.2.14	<i>File formats</i>	45
3.2.15	<i>Type of content changes</i>	46
3.2.16	<i>Content status changes of files</i>	47
3.2.17	<i>Content status changes of documents</i>	50
3.2.18	<i>Dependencies between content status changes and metadata changes</i>	55
3.2.19	<i>Full and incremental submissions</i>	56
3.2.20	<i>Persistent identifiers</i>	56
3.2.21	<i>Further identifiers on submitter and regulatory authority level</i>	58
3.2.22	<i>Impact of document content status on evaluation price</i>	59
3.2.23	<i>GHSTS Picklists</i>	59
3.2.24	<i>Information for and targeting at individual receivers</i>	61
3.2.25	<i>Internal and external references</i>	62
4	GHSTS SUBMISSION SCHEMA DEFINITION	66
4.1	Overview	66
4.2	Filename and version	66
4.3	Scope	66
4.4	Namespace	67
4.5	Unique keys	67
4.6	Combination of identifier and identifier type	67
4.7	Notation conventions	68
4.7.1	<i>Elements</i>	68
4.7.2	<i>Attributes</i>	69
4.7.3	<i>Occurrence indicators</i>	69
4.8	Boilerplate for element description	69
4.8.1	<i>Child elements</i>	70
4.8.2	<i>Attributes</i>	70
4.8.3	<i>Constraints</i>	70
4.8.4	<i>Example</i>	71
4.9	Elements	71
4.9.1	<i>Overview</i>	71
4.9.2	<i>GHSTS</i>	71
4.9.3	<i>RECEIVERS</i>	73
4.9.4	<i>RECEIVER</i>	73
4.9.5	<i>SENDER</i>	74
4.9.6	<i>PRODUCT</i>	75
4.9.7	<i>PRODUCT_RA</i>	77
4.9.8	<i>INGREDIENTS</i>	77
4.9.9	<i>INGREDIENT</i>	78
4.9.10	<i>DOSSIER</i>	79
4.9.11	<i>REFERENCED_DOSSIER</i>	80

4.9.12	<i>DOSSIER_RA</i>	80
4.9.13	<i>SUBMISSION</i>	82
4.9.14	<i>DOCUMENTS</i>	82
4.9.15	<i>DOCUMENT</i>	83
4.9.16	<i>DOCUMENT_RA</i>	84
4.9.17	<i>OTHER_NATIONAL_GUIDELINE</i>	85
4.9.18	<i>RA_DOCUMENT_NUMBER</i>	86
4.9.19	<i>SUBMISSION_CONTEXT</i>	87
4.9.20	<i>DOCUMENT_GENERIC</i>	87
4.9.21	<i>CONTENT_STATUS_HISTORY</i>	92
4.9.22	<i>REFERENCED_DOCUMENT</i>	93
4.9.23	<i>RELATED_TO_SUBSTANCE</i>	94
4.9.24	<i>REFERENCED_TO_FILE</i>	94
4.9.25	<i>FILES</i>	95
4.9.26	<i>FILE</i>	95
4.9.27	<i>FILE_RA</i>	96
4.9.28	<i>FILE_GENERIC</i>	97
4.9.29	<i>TOC</i>	99
4.9.30	<i>STANDARD_TOC_REFERENCE</i>	100
4.9.31	<i>STRUCTURE</i>	100
4.9.32	<i>LEGAL_ENTITIES</i>	101
4.9.33	<i>LEGAL_ENTITY</i>	101
4.9.34	<i>CONTACT_ADDRESS</i>	103
4.9.35	<i>CONTACT_PERSON</i>	104
4.9.36	<i>SUBSTANCES</i>	106
4.9.37	<i>SUBSTANCE</i>	106
4.9.38	<i>USED_TEMPLATES</i>	107
4.10	Basic types	107
4.10.1	<i>Metadata status type</i>	107
4.10.2	<i>Content status type</i>	108
4.10.3	<i>TOC node type</i>	109
4.10.4	<i>Relative file path type</i>	111
4.10.5	<i>Version number type</i>	112
4.10.6	<i>Submission number type</i>	112
4.10.7	<i>PID type</i>	113
4.10.8	<i>UUID type</i>	114
4.10.9	<i>MD5 type</i>	115
4.10.10	<i>Standard XS and XSI types</i>	115
4.10.11	<i>Restricted string types</i>	116
5	REFERENCED SCHEMA DEFINITIONS	117
5.1	Standard ToC schema definition	117
5.1.1	<i>Overview</i>	117
5.1.2	<i>Filename and version</i>	117
5.1.3	<i>Namespace</i>	117
5.1.4	<i>Unique keys</i>	118
5.1.5	<i>Notation conventions</i>	118
5.1.6	<i>Elements</i>	118
5.1.7	<i>Basic types</i>	121
5.2	GHSTS picklist schema definition	121
5.2.1	<i>Overview</i>	121

5.2.2	<i>Filename and version</i>	122
5.2.3	<i>Picklists</i>	122
5.2.4	<i>Open picklists</i>	123
6	GHSTS PACKAGE DEFINITION	125
6.1	Overview of Folder structure	125
6.1.1	<i>Usage of Content folder</i>	126
6.2	Storage of multiple submissions of the same dossier	126
6.3	Folder naming conventions and constraints	127
6.4	File naming conventions	127
6.5	Platform and media support	128
7	USING GHSTS	129
7.1	Creation and publication of a GHSTS submission	129
7.1.1	<i>Querying the GHSTS Management Body</i>	129
7.2	Conformity testing	129
7.3	Transport of a GHSTS submission	129
7.4	Consumption of a GHSTS submission	129
8	APPENDIX	131
8.1	Glossary of Terms	131
8.2	Standards References	137
8.3	Copyright and Trademark Information	137
8.4	Version history	138

1 PREFACE

1.1 GHSTS Primer

1.1.1 *Scope of GHSTS*

The Global Harmonised Submission Transport Standard (GHSTS) is a standardized set of technical specifications used to assemble electronic files for the electronic submission of pesticide registration applications to regulatory authorities in a predefined manner. The files contained can take any form appropriate for the business needs. Standard text processing and spread sheet formats, PDF and XML data files are just some of the file types which can be transferred using the GHSTS. Once the files are assembled according to the specifications, they can be transferred from one business entity to another with the receiving entity able to extract the files for use in the regulatory process. At present, the majority of the pesticide regulatory process in most nations is centred on documents, for example the “Dossier”, “Monograph”, “study”, “DER”. It is these documents that are required to be exchanged and that are central to the existing processes and policies.

Limited metadata are included in the GHSTS. Only enough information is included to identify who the submitting entity is, the purpose of the files contained within the GHSTS, and how to handle the files once received. The GHSTS itself is not intended to use the content of the files. It is not a standard for the information itself.

The GHSTS is not a tool or software application. Other Information Technology (IT) systems can be constructed to utilize the standard and by doing so, would be able to easily transfer data to each other.

1.1.2 *Background and history*

As increasing numbers of Joint Reviews are processed by regulatory authorities there has been an increased call for harmonization of the review process. With a number of disparate systems and methodologies in place around the world, multinational submissions can be both resource intensive for the industry to assemble and awkward for the authorities to review.

Currently industry has to comply with many different electronic standards for the submission of data to regulatory authorities in the pesticide registration area. Examples of such systems are the Australian Pesticides and Veterinary Medicines Authority data list, CADDY (Computer Aided Dossier and Data Supply)/CADDY-xml in the European Union, e-Index (electronic index) / e-PRS (Electronic Pesticide Regulatory System) in Canada, PRISM (Pesticide Registration Information System) in the United States and IUCLID 5 (International Uniform Chemical Information Database) / REACH IT used in Europe for classification and labelling purposes.

Other regulatory authorities are also developing their own e-submission standards resulting in additional hurdles for companies engaged in submissions globally, particularly in Joint Reviews of an agricultural pesticide submission by two or more regulatory authorities. Often this results in incompatible deliverables for the applicant and in a communication dilemma between regulatory authorities and with the applicant. In addition, most of the information submitted for Joint Reviews to date has been unstructured, making efficient and effective information management very difficult.

The existence of multiple standards means that the applicant must duplicate their submission preparation effort as they need to provide a list of data elements with matching dossier and document metadata such as title, author, guideline number and report number in a proprietary format for every electronic submission to a regulatory authority in support of an application for registration. In some cases regulators will provide a software tool to prepare electronic submissions, but data entry of metadata is usually manual work for the

applicant. Manual compilation of large dossiers can take several man weeks and is an inefficient and error prone process.

Taking the lead in harmonizing the pesticide review process, for the past number of years the OECD has been engaged with efforts to harmonize the pesticide review process world wide.

To further the efforts of global harmonization, the OECD *Ad hoc* Expert Group on the Electronic Exchange of Pesticide Data was tasked with investigating the possibility of harmonization in the information technology used in the pesticide regulatory process, where that harmonization might occur, and what would be required for the harmonization to take place. As part of its investigation, the Expert Group created the Transport Mechanism Subgroup which looked specifically at the ability to harmonize the various methods used to submit information to the regulatory authorities.

These two groups focused their attention on the development of a common method of electronic submission – the Global Harmonised Submission Transport Standard (GHSTS).

1.1.3 Supporting stakeholders

Participants in the OECD *Ad hoc* Expert Group on the Electronic Exchange of Pesticide Data and the Transport Mechanism Subgroup include members from OECD countries and the European Commission represented by the European Chemicals Agency, the European Food Safety Authority and experts from the pesticide industry.

1.1.4 The role of the OECD

Ownership of the GHSTS lies with OECD, and responsibility for monitoring updates of the GHSTS lie with the OECD *Ad hoc* Expert Group of the Electronic Exchange of Pesticide Data, under the supervision of OECD Registration Steering Group and OECD Working Group on Pesticides.

All components of the GHSTS (XML schema, picklists, OECD and national dossier numbering systems [also known as the Table of Contents] in XML format, guidance documents, etc.) are made available on the OECD public website.

1.2 Scope and audience of this document

The intended audience of this document is as follows:

- Responsible stakeholders on business project level of registrants and authorities that want to adopt GHSTS and need to know details about GHSTS, its usage and main characteristics.
- IT project leads from IT departments that are in charge to guide development projects supporting GHSTS, mostly publication components or ingestion components and that need to evaluate the impact of adopting GHSTS for their in-house dossier management systems.
- Business domain experts that are in charge of analysing generation, validation or import errors.
- IT consultants in charge of GHSTS-related development that require a detailed specification of GHSTS. Expertise in XML and XSD technology is assumed, general knowledge of document management is helpful.

To support the different needs, this document is divided in different sections that describe GHSTS on a conceptual level and then provide technical in-depth details. To ease the use of direct access to specific topics as a reference manual cross-references have been included, also to avoid repetition of statements from the conceptual part in the technical part.

1.3 Organisation of the document

The specification is organized in the following chapters:

- Chapter 2 “Context Of GHSTS” sets GHSTS in the context of the submission process and other related standards. The landscape of GHSTS components next to this specification is introduced. Scenarios for the transition to GHSTS are outlined.
- Chapter 3 “Introduction to GHSTS” explains the design principles behind GHSTS and explains the basic concepts in GHSTS on a logical level, not obscured by technical details.
- Chapter 4 “GHSTS submission schema definition” is the technical reference for the structure of the GHSTS XML backbone file that is the core of submission using GHSTS.
- Chapter 5 “Referenced schema definitions” is the technical reference for the GHSTS Table of Contents (ToC) structure and the GHSTS picklist definition file that are referenced from the GHSTS XML backbone file.
- Chapter 6 “GHSTS Package Definition” specifies the structure and characteristics of the GHSTS package in which all content and additional files are packed for transport.
- Chapter 7 “Using GHSTS” gives some guidance on how GHSTS ought to be used, from creation via testing to viewing.
- Chapter 8 “Appendix” contains a number of appendices. It also contains the version history where the changes between the versions of this specification are detailed.

The document can be read in a sequential way, but also partially to serve as a reference manual. To serve the latter purpose, many cross-references have been introduced to navigate through the document as needed.

1.4 Filename and version

The GHSTS format specification (this document) is versioned using a version numbering scheme of type XX.XX.XX. It is composed of a two digit major, a two digit minor and a two digit patch level number separated by single dots. Major versions are used for global functional changes, minor versions for local changes and patch level versions for error corrections. Patch level versions do not require new versions of other GHSTS components (see chapter 2.4).

The filename of the GHSTS format specification uses the same versioning scheme with the hyphen instead of the dot: The file naming convention of the GHSTS format specification is therefore

ghsts_XX-XX-XX.<format suffix>

for example

ghsts_01-00-00.pdf

1.5 Contacts

Any entity needing assistance regarding the GHSTS should contact the OECD at <mailto:ehscont@oecd.org>. The Secretariat will then direct questions to the OECD body responsible for the maintenance of the GHSTS.

Role	Name	Email
Organisation responsible for the GHSTS	OECD Secretariat	ehscont@oecd.org

Table 1 – Contacts

2 CONTEXT OF GHSTS

2.1 Usage of GHSTS in the submission process

The GHSTS standard supports applicants and regulatory authorities in the course of a dossier regulatory action for a pesticide registration. This chapter explains the context in which GHSTS is applied and gives an overview of the process. Technical details are explained in subsequent chapters.

A dossier regulatory action is a series of events accomplished by a regulatory authority beginning with the submission of an application with data by a registrant and ending with a unique regulatory decision. The following figure shows a sample for one dossier regulatory action, for a given active ingredient or product.

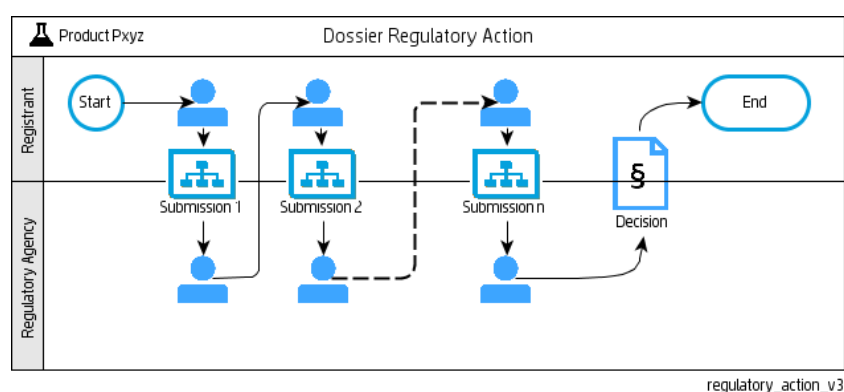


Figure 1 - Dossier Regulatory Action

As illustrated, the registrant initiates the dossier regulatory action with the initial submission. The submission is verified by the regulatory authorities, eventually leading to new submissions to fulfil the regulatory requirements. Finally, the regulatory authority submits the decision for the application to the registrant. This ends this specific regulatory action.

The dossier is the set of documents prepared and continuously maintained by the applicant throughout the lifecycle of a regulatory action. A dossier has at least one submission. A submission is the compilation of documents in a structured form according to the given regulatory requirements. Multiple submissions can be submitted for each dossier, they are differentiated with version numbers.

GHSTS supports the dossier regulatory action by defining a standard for the representation of the individual submissions in a regulatory action. The representation is called a GHSTS package and contains all required information for a submission as well as lifecycle information to set the submission in the correct context of the dossier regulatory action.

GHSTS also supports the Joint submission and the Joint review scenario with multiple senders (registrants) and multiple receivers (regulatory authorities) for the same regulatory action.

The following figure provides details about one single submission step. For information on what depicted components are currently available please see chapter 2.4. The domain of the registrant is depicted on the left side, the transport domain in the middle and the regulatory authority domain on the right side.

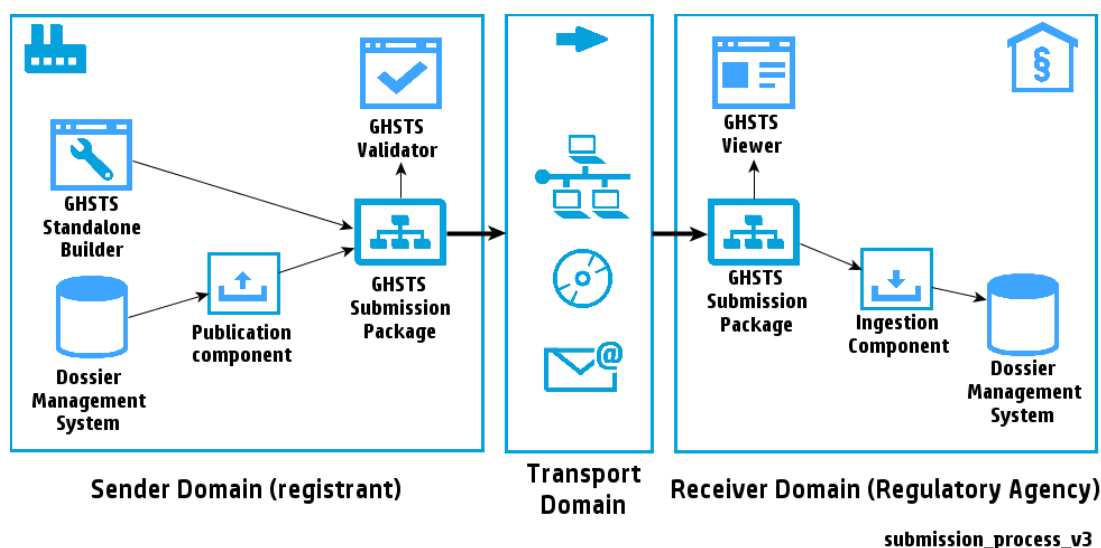


Figure 2 – Submission process

A GHSTS submission package can be created in one of the following ways:

- The required information and content for the submission is kept in a dossier management system and the GHSTS submission package is created by a publication component that generates the GHSTS submission package format.
- A GHSTS standalone builder is used that is capable to create GHSTS submission packages based on information directly supplied to this tool.

After the creation of the GHSTS submission package it must be validated using a GHSTS validator. The submission package cannot be validated directly by a human-reader but requires software components to be processed.

Attention: The GHSTS validator does not currently exist. Its development is under consideration.

The validated package can then be handed over to the regulatory authority as receiver. GHSTS does not restrict the parties on how a package should be transported – for example network protocols, offline media like CD/DVD, USB and hard-drives or contained like email may be used.

The receiver can work with the GHSTS package in at least two ways:

- A GHSTS ingestion component can parse the package and ingest the data into the electronic dossier management system for review and approval. After ingestion the package can be kept for auditing reasons but the further activities will be done with processes and tools outside of GHSTS.
- The GHSTS submission package contains a standalone browser-based GHSTS viewer that can be used for viewing the package in a user-friendly way. The viewer is read-only and does not support additional tools that can be used for review (e.g. notes).

If the submission is not the initial submission of the regulatory action then it will likely contain external references to predecessor packages.

- During ingestion into a document management system these external references need to be resolved.

- The GHSTS viewer can resolve references to predecessor submissions if all packages of the same regulatory action are placed in one common file directory.

Attention: The GHSTS viewer does not currently exist. Its development is under consideration.

The next figure shows a high-level overview of the internals of submission packages:

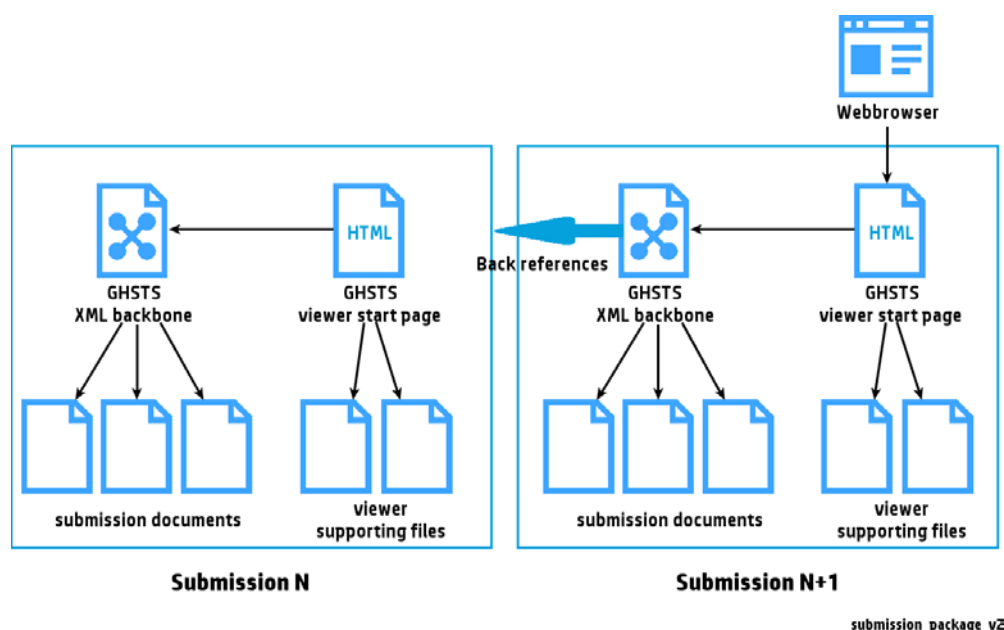


Figure 3 – Submission package overview

A submission package consists of two main parts:

- The GHSTS XML main backbone contains the metadata of the submission as well as references to the submission documents. The GHSTS XML main backbone can also contain external references, for example to a previous submission package.
- The GHSTS viewer is added to the submission package to allow the visualization for a human reader without external support.

2.2 GHSTS and different submission Types

2.2.1 Joint review

The regulatory authorities of two or more countries may share the evaluation work in a “Joint Review” of an agricultural pesticide submission (dossier) for a new active ingredient and/or associated end-use product(s) or an additional use for an existing active ingredient or end-use product. The participating regulatory authorities review the work of the primary reviewers for each particular science discipline, and the end product (ideally a complete monograph or key components of the monograph) is used by all participating countries (and others) as the basis for regulatory decisions.

A Joint Review requires that the pesticide dossier be submitted to all participating regulatory authorities simultaneously. This is a formal process in which timelines and work allocation are negotiated in advance. Data reviews are exchanged and peer reviewed, and there is agreement on both the documentation to be produced and the decision-making target date, i.e. the date on which the decision (which may be a proposed decision) is communicated to the applicant.

Joint review projects are not managed by the OECD Secretariat or any official OECD body. Rather, they are coordinated by various groups of countries (which may or may not all be OECD members) and pesticide companies. The OECD encourages regulatory authorities to adopt the use of the GHSTS and the OECD numbering system for dossiers in order to allow for easier communication among regulatory authorities during Joint Reviews and to reduce the regulatory burden on applicants.

2.3 Related standards

2.3.1 *Submission standards on national / regional level*

As a first step in the development of the GHSTS, the OECD focused on the metadata used in each of the electronic standards already in place to submit information to regulatory authorities and developed a single set of metadata that could be used by all authorities. This set of metadata includes elements common to all the existing systems along with additional elements required for specific systems. Below are the specific electronic submission standards on national and regional levels whose metadata were reviewed.

2.3.1.1 *CADDY*

CADDY is the first electronic submission format for pesticides. It has been developed jointly by regulatory authorities in the European Union and industry since 1995 and is currently the e-submission standard used in the European Union for pesticides. The strategic goal behind CADDY was to facilitate the provision of dossiers for pesticides to regulatory authorities, the long-term archiving of such dossiers, the accessibility of information contained in such dossiers and the examination and assessment of dossiers by regulatory authorities in a cost-effective manner using electronic media.

The initial version of CADDY was focused on replacing paper submissions and used an image-based page standard (TIFF) and associated metadata to allow easy searching, retrieving viewing and printing of documents within a submission. Later the development focus of CADDY shifted towards facilitating the review of submissions by regulators and additional functionality to this effect was added. Since 2005 the CADDY specification has been migrated to a new xml and PDF/A platform, while maintaining backward compatibility of the dossier metadata and providing a migration path to easily convert existing CADDY dossiers to the current CADDY-xml standard v3.0.

2.3.1.2 *e-PRS/e-Index Builder*

Electronic Pesticide Regulatory System (e-PRS) implemented in 2000 of the Health Canada Pest Management Regulatory Agency (PMRA) is an integrated electronic database designed to enable the Agency to manage regulatory submissions and associated information related to pest control products and active ingredients. The e-Index Builder is a standalone software application which facilitates the creation by an applicant of an electronic index (in XML format) that fully describes each document to be submitted in support of an application to register or amend a pest control product registration. The electronic index created by the applicant with the tool includes information on the documents' metadata and the context in which it is being submitted. The e-Index builder supports data coding systems of the OECD, PMRA and United States Environmental Protection Agency (US EPA).

Once indexed, applicants may either attach an electronic document, mail in the paper version of the document, or "cross-reference" the document to one that is already in the PMRA's ePRS database. When the index is completed, the applicant submits it to the Agency as a "PMRA Regulatory ZIP" (PRZ) file. This file includes the e-Index as well as any attached electronic documents.

In summary, the e-Index Builder is an indexing and compiling tool for use by applicants in the preparation of their regulatory submissions, whereas the e-PRS is a database which enables the PMRA to manage and view the many documents associated with the products and active ingredients it regulates.

2.3.1.3 *e-PRISM*

In May 2003, the United States Environmental Protection Agency (US EPA) introduced an electronic submission program where registrants would submit electronic copies of studies after paper copies were submitted and processed for formatting requirements (PR notice 86-5). The intent of the electronic copies was to assist in the review process allowing the scientist to search for content within the study and extract that content for use in the Agency's evaluation, even though the paper copies were considered the official copy of the study.

In July 2008, the United States introduced a revised method for electronic submission that is more advanced and consistent with current technology standards. The electronic submission process, e-PRISM (now e-Dossier), is based on the Canadian e-Index and uses much of the same technical specifications. Even though the U.S. e-PRISM methodology is based on the Canadian approach, some differences still exist.

2.3.2 *IUCLID*

IUCLID (International Uniform Chemical Information Database) is a software tool used to capture, store, maintain, and exchange data on intrinsic and hazard properties of chemical substances according to the format of the OECD Harmonised Templates for Reporting Chemical Test Summaries (OHTs). Distributed free of charge, the software is maintained by the European Chemicals Agency (ECHA).

The software is especially useful to chemical industry companies and to government authorities. IUCLID was built to comply with the European Union REACH Regulation, a European Union law covering the production and use of chemical substances, but is also used in support of the submission requirements of the European Union Biocidal Product Regulation and the OECD Cooperative Chemicals Assessment Programme and is promoted for developments in international harmonization.

Data that can be stored and maintained with IUCLID include information about the chemical substance and its composition, reference information like substance identifiers, classification and labelling, use and exposure, physical and chemical properties, toxicological properties, eco-toxicological properties and the modification history. The IUCLID data model also features Biocides/Pesticides elements.

While the majority of pesticide authorities' processes are built around unstructured data and documents, the XML of the OECD Harmonised Templates for Reporting Chemical Test Summaries contained in IUCLID are used to present structured data. Study summaries in the OECD XML Harmonised Template format may be exported from, exchanged and imported to IUCLID in individual XML files.

2.3.3 *OECD Harmonised Templates for Reporting Chemical Test Summaries (OECD Harmonised templates)*

Beginning in 2002 the OECD undertook an effort to harmonize, where possible by the use of templates and XML tags, tools for submission, evaluation and exchange of chemical data for the regulation of new and existing industrial chemicals, agricultural pesticides and biocides. This led to the establishment of the OECD Expert Group on Harmonising Templates and eventually, the development of standard data formats consisting of end point templates, XML schema, and XML schematrons (rule definition files), for reporting studies done on chemicals to determine their properties or effects on human health and the environment (e.g., hydrolysis, skin irritation, repeat dose toxicity, etc.). These templates can be used for reporting summary results for testing on any type of a chemical (e.g., industrial chemicals, pesticides, biocides).

The Templates are aimed at developers of database systems, as they prescribe the formats by which information can be entered into and maintained in database. The XML schema and schematrons are technical specifications that define the data structure, relation, condition, constraints, and validation rules of that data.

The OECD XML study templates are presently used mainly by the IUCLID software, eChemPortal - The Global Portal to Information on Chemical Substances of the OECD, and the OECD QSAR Toolbox for filling gaps in (eco) toxicity data needed for assessing the hazards of chemicals.

2.4 GHSTS components

This chapter provides an overview of GHSTS existing and future GHSTS components:

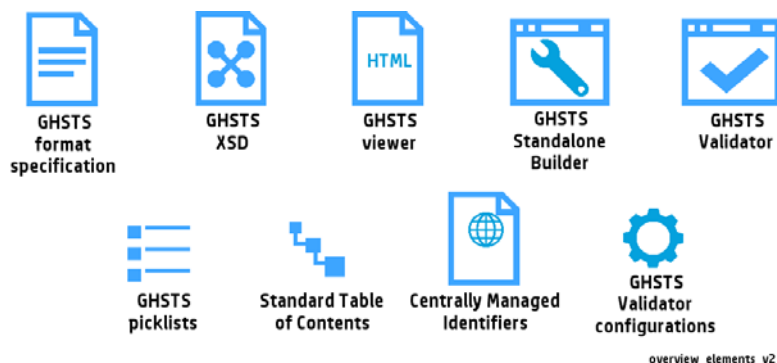


Figure 4 – Overview of GHSTS components

No support will be available from the OECD for other components like publication and ingestion components.

Attention: Please note that with the current version of the Specification the following components are available:

- GHSTS format specification (this document)
- GHSTS XSD (the structure definition of the GHSTS XML backbone)
- GHSTS picklists
- Standard Table of Contents XSD (currently there are no instances of ToC according to this XSD available)

A GHSTS standalone builder will not be supplied by the OECD.
The set of components is currently under review and development.

2.4.1 GHSTS format specification

The GHSTS format specification is the current document. It contains the GHSTS overall description and the definition for the GHSTS schema definition.

2.4.2 GHSTS submission schema definition

The GHSTS submission schema definition (or simply called GHSTS XSD) contains the technical definition of the GHSTS XML backbone file. It is generated from information contained in the GHSTS format specification and supplied in XML schema definition format (XSD).

2.4.3 GHSTS viewer

The GHSTS viewer allows the standalone display of a submission package in a web browser. The viewer will only display a subset of information available in the submission package; the full set of information is available in the GHSTS XML backbone of the submission and the supplied files.

Attention: The GHSTS viewer does not currently exist. Its development is under consideration.

2.4.4 *GHSTS standalone builder*

A GHSTS standalone builder is a client-side standalone software application that can generate GHSTS-compliant submission packages. This type of tool is intended for use for registrants without a dossier management system.

The OECD will not provide a builder application. It will be the responsibility of industry to acquire/develop one or many standalone builder applications.

2.4.5 *GHSTS validator*

The GHSTS validator is a software program that can validate a GHSTS submission according to rules defined in the GHSTS validator configurations. As a result the submission package is deemed compliant to the defined rules.

Attention: The GHSTS validator does not currently exist. Its development is under consideration.

2.4.5.1 *GHSTS validator configurations*

The GHSTS validator applies a configuration (a set of validation rules) to validate a GHSTS submission against the GHSTS specification. Multiple validator configurations can be supplied to accommodate for additional validation rules, for example rules required by a specific regulatory authority.

2.4.6 *GHSTS Converter*

A GHSTS converter converts a different dossier submission format (see e.g. chapter 2.3.1) to GHSTS format. There may be multiple converters from and to GHSTS.

Attention: GHSTS Converters do not currently exist. Their development is under consideration.

2.4.7 *GHSTS Picklists*

Upon generation of a GHSTS submission package the registrant has to supply metadata. Some metadata have to be selected from a list of values supplied in a separate GHSTS picklist file. This file is supplied in XSD format.

2.4.8 *Standard Table of Contents*

The Table of Contents (ToC) used in a GHSTS submission has to have a structure according to a standard Table of Contents defined by authorities. The term “Table of Contents” in this document is used as a synonym for “dossier numbering system”.

These standard ToCs are supplied as XML files that follow a ToC schema definition. They can be used in the following context:

- Import of the standard ToCs into the registrant dossier management system – to use the ToCs for an upcoming submission. These may be GHSTS submissions, but the ToCs may also be used for submissions in different submission formats.
- Usage of the standard ToC information as copy of the hierarchical structure in the backbone XML of the submission package
- Reference of the standard ToC XML file in the backbone XML to be used for validation by the GHSTS validator (compare ToC copy in GHST XML backbone with external ToC XML file).
- Usage in authority in-house systems.

The Table of Contents can be used also outside of the GHSTS contexts, therefore they are deliberately not named using the term GHSTS.

The dependency between ToC instance, dossier management system and GHSTS submission package are depicted in the following figure:

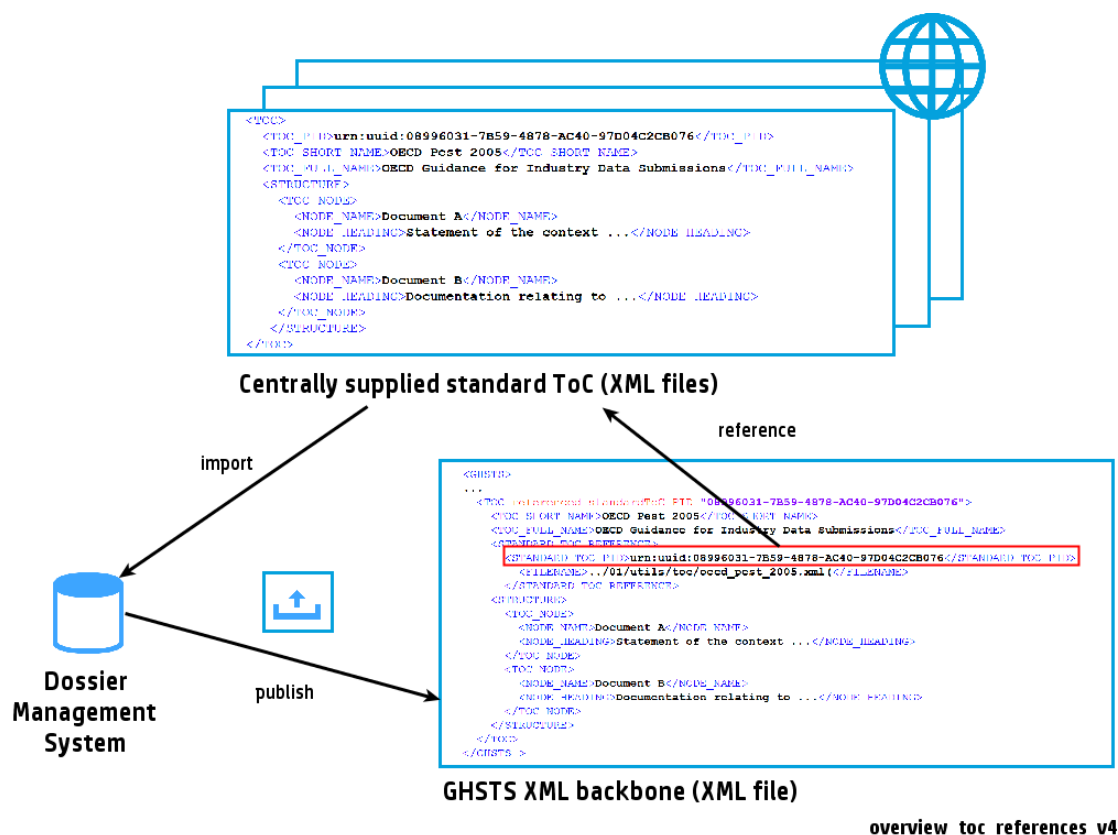


Figure 5 – Overview of ToC references

The top part shows the centrally supplied standard ToC files. These can be imported into the dossier management system. During a GHSTS publication process the resulting GHSTS XML backbone file will contain a logical reference to the used standard ToC in the form of a persistent identifier of the ToC. In addition the standard ToC file is also part of the submission package and physically referenced (not depicted here).

To support the use case of submissions where no centrally supplied ToC is available the logical and physical ToC references in the GHSTS XML backbone files are optional. In this case the GHSTS XML backbone will contain a ToC without further references.

2.4.9 Centrally managed components

Attention: Scope and functionality of central management by the OECD is currently under review and development.

GHSTS enforces the use of the following globally unique elements to ensure unambiguous identification on a global level:

- Unique and persistent identifiers and identifying metadata of the GHSTS concepts “Legal entity” (see chapter 3.2.8) and “substance” (pure reference substance) independent on their exact naming and format. For details about the persistent identifiers and which are managed centrally see “Table 11 – Overview of concepts identified by PID” in chapter 3.2.20.

- Available values in picklists. GHSTS provides picklist values that are to be used to enforce global uniqueness. If new values are required then a new version of the GHSTS picklist XSD has to be supplied centrally.
- Standard Table of Contents as defined by OECD including the unique naming of individual ToC nodes. The OECD makes available the following ToC:
 1. Chemical pesticides
 2. Microbials
 3. Pheromones and Other Semiochemicals
- If it is necessary to use national or pan-national (e.g. from the EU) numbering systems parallel to the OECD ToC, the OECD encourages regulatory authorities to publish their national numbering systems supplemented by references into the OECD numbering system on the same platform. So it would be possible to make an (semi-) automatic “cross walk” between OECD ToCs and national ToCs created as member countries.

These components are managed and distributed centrally. When one of the above components is available or updated it is published on the central repository with the relevant version number. Deprecated components are also available on demand.

Currently the responsibility of the management and publication of these components is the *Ad Hoc* Expert Group on the Electronic Exchange of Pesticide Data at the OECD. The remainder of the document will use the term “GHSTS Management Body” (GMB), for the body which centrally manages and publishes the above components within at the OECD.

The tasks of the GMB are:

- Creation and publication and update of persistent identifiers for substances and legal entities. Possibility of (partial) self-service to create/request, query and update identifiers.

Attention: The creation and publication of persistent identifiers for substances and legal entities at a central location is currently under consideration.

- Collection and publication of different types of ToC (standard ToC, national numbering systems) according to the GHSTS Toc XSD. Possibility to download these assets and to subscribe for changes.
- Publication of the GHSTS components like the GHSTS specification and the other GHSTS components (e.g. GHSTS viewer, GHSTS picklists etc.). Information about releases and release changes. Possibility to download these components and to subscribe for changes.

For a sketch of the required GMB functionality during the creation of a GHSTS submission see chapter 7.1.1.

2.4.10 Conversion tables between dossier structures

GHSTS does not provide conversion tables – also called cross-walk tables - between different types of tables of contents. Currently GHSTS focuses on the submission transport aspect with a fixed ToC.

2.4.11 Lifecycle and dependencies of GHSTS components

Each GHSTS component listed above can have its own lifecycle and will be versioned separately from other components using separate version information. However, the components may still be dependent on each other: A new version of the GHSTS backbone XSD requires a mandatory change in the GHSTS format specification. Whether, for example, the GHSTS converter has to be changed depends on the type

of change. In general it has to be documented what version of a specific GHSTS component supports what other component, for example, which XSD version reflects which format specification version.

The dependencies between GHSTS components are documented in the following form:

GHSTS component	Release Package 1 <version and date>	Release Package 2 <version and date>
Format specification	< version and date>	< version and date>
Schema definition	< version and date>	< version and date>
Viewer	< version and date>	< version and date>
Validator	< version and date>	< version and date>
Validator configuration instance X	< version and date>	< version and date>
Picklist XSD	< version and date>	< version and date>
Table of Contents XSD	< version and date>	< version and date>

Table 2 – GHSTS Release Package Matrix example

A new GHSTS release package is generated, when one GHSTS component is versioned. One line is maintained for each GHSTS component. Each cell contains the latest version number and date that supports this release package.

Example: If the schema definition is versioned a new release package is generated. Then, the format specification will also be versioned. If the change of the schema is irrelevant for the viewer and validator then the version will not change compared to the predecessor release package. Comments can be added to the cells, for example, to clarify when a range of versions can support the release package.

Instances of Table of Contents (e.g. OECD ToC, specific national ToC) are independent of GHSTS. The only possible dependency would be, if a change in the ToC XSD would require that the ToC representation would need to be updated to be compliant with the updated XSD.

The latest GHSTS release package matrix for the current version is documented on the GMB website.

2.5 Management of GHSTS

2.5.1 Information sources for GHSTS elements

All the existing centrally managed GHSTS are available on the OECD website at: **XXXXXX**.

3 INTRODUCTION TO GHSTS

This chapter will provide an introduction to GHSTS. It intends to explain GHSTS mainly from the business viewpoint, avoiding many references to the technical chapters that follow afterwards.

- First in chapter 3.1 the design principles behind GHSTS are outlined.
- Then, the logical concepts that are embodied in the GHSTS submission package are explained in chapter 3.2.

3.1 GHSTS Design Principles

3.1.1 *Parallel viewpoints*

A GHSTS submission package supports two different complementary identification view levels, a human readable level and machine readable level:

- The human-readable level is supplied when using the GHSTS viewer. The GHSTS XML backbone file stores the decoded value of usually coded information to allow human-readable display without external references. Of course these values can be used with a different proprietary viewer as well.
- The mandatory “main” content of the documents is supplied in PDF files for human readers.
- The optional “source” content of documents is also intended for human readers.
- For the optional “supplement” content of the document it depends on the file format whether this is intended for human readers or needs processing by a computer program first.
- The machine readable level is supplied by the GHSTS XML backbone file supplying additional technical metadata and references to all supplied files. Most instances of the GHSTS concepts are identified by unique program-readable identifiers. This information can be used by ingestion components to identify and ingest the relevant information for the receiver.
- Additional document content can be supplied in XML format that can be harvested by suitable computer programs, but also displayed for human readers when appropriate display resources (e.g. style sheets) are supplied.

The following figure visualizes the parallel viewpoints:

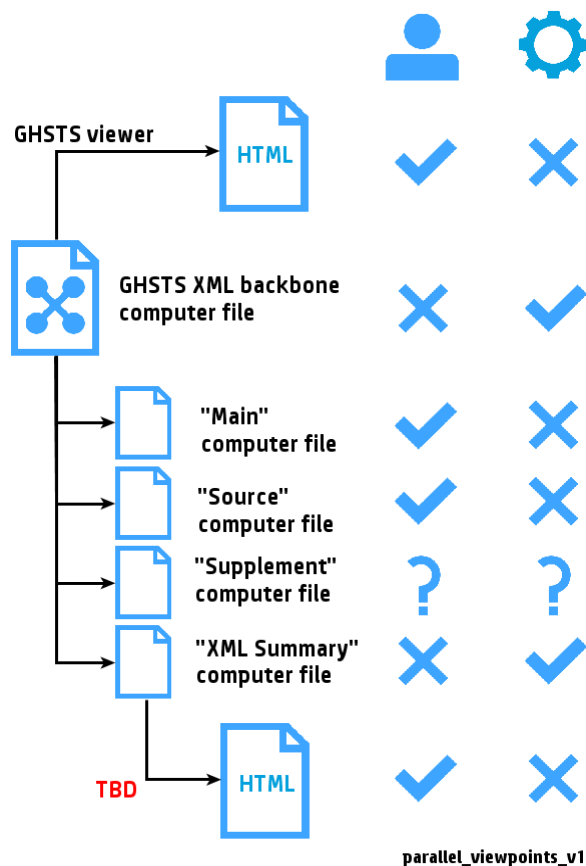


Figure 6 – Support of parallel viewpoints

Attention: The mechanism of a standalone web browser display of OHT XML summaries is under review and not yet available, therefore the figure above shows the sign “TBD” (to be defined). Please note that this mechanism is not dependent on GHSTS itself; however the mechanism would need to be incorporated in the GHSTS viewer.

3.1.2 Parallel navigation scenarios

The GHSTS submission information supports the two common types of navigation scenarios:

- “Show list of documents and display guidelines they refer to”.
- “Display ToC and display attached documents”.

The required information is stored in the GHSTS XML backbone; the GHSTS viewer supports both navigation scenarios in parallel.

3.1.3 Document-based organisation

A GHSTS submission package is organized on two levels:

- A set of documents in which each document subsumes one or multiple computer files in potentially different formats.

- One GHSTS XML file providing administrative and bibliographic metadata for the documents, the submission, and the guideline / ToC and referencing the set of documents.

3.1.4 *Integration of XML-based content*

A GHSTS submission package is open to integrate content in XML format like XML conformant to the OECD Harmonised Templates for Reporting Chemical Test Summaries (OHT). The information in these XML files can be harvested by computer programs. If the XML document supports standard rendering techniques (e.g. XSL, CSS) and the required assets are part of the submission package and correctly referenced, then the information can also be displayed for users by the GHSTS viewer.

The XML content is only one type of content that GHSTS as transport standard can support. Hence GHSTS and OHT are concepts in different domains. Unlike the OECD Harmonised Templates where significant study information is contained within the XML file, limited metadata for content is included in the GHSTS.

3.1.5 *Adaptiveness with respect to different standard ToC templates*

The GHSTS standard is independent of the structure of the Table of Contents. The GHSTS XML backbone can contain an arbitrary Table of Contents (ToC) structure. Optionally the GHSTS validator can check whether the used Table of Contents in the submission package corresponds to one of the existing standard ToC templates.

3.1.6 *Lifecycle management*

The GHSTS standard contains lifecycle information for different embodied information concepts (e.g. product/substance, document metadata, file and document content, ToC etc.) to allow the quick identification of changes between subsequent submissions of this dossier. This information can be used during review or ingestion to focus on the changes in the latest version.

For details about the content lifecycle management see chapter 3.2.14, for the metadata lifecycle management see chapter 3.2.3.3.

3.1.7 *Persistent identifiers of information concepts*

GHSTS supports the use of unique and persistent identifiers of a set of information concepts that are considered to be valid and reusable outside of the scope of one regulatory action, one registrant or one regulatory authority:

- Legal Entity
- Substance
- ToC and ToC node
- Dossier
- Product
- Document and Document Family
- File

These identifiers are called Persistent Identifiers (PID). The term “persistent” underlines that the instances of these information concepts can permanently be identified with this identifier throughout their lifetime and in all contexts where this instance is used. This minimizes the need to send duplicate information by referencing its unique identifier in subsequent submissions.

Attention: Central management by the GMB outside of the context of registrants and regulatory authorities is currently under consideration for the first three identifiers in the list above in order to ensure uniqueness and harmonisation of the data being submitted.

For the discussion of the central management see chapter 2.4.9.

The others are managed by the registrant. The registrant generates a PID once at the creation of the instance of the information concept and will use this across all regulatory actions.

If an instance of one of the above information concepts uses a specific PID that has been used and assigned before, then the consumer of the GHSTS submission (human reader or program) can safely assume that the instance is identical to the one submitted before without the need to check the complete concept in detail.

For example, if a document is submitted with a specific PID as part of a submission package, the receiving regulatory agency can verify if it or another agency has already received and evaluated this document before by simply checking the PID against their dossier management system.

For more information about persistent identifier see chapter 3.2.20.

3.1.8 Information specific for regulatory authorities

GHSTS provides the possibility to supply individual and different metadata to regulatory authorities within the same regulatory action when multiple regulatory authorities are involved. For details see chapter 3.2.24.

3.2 GHSTS logical concepts

This chapter explains the different information concepts embodied in GHSTS on a conceptual and semantic level. The technical reference on how these concepts are represented in the submission package - mainly in the GHSTS XML backbone - is contained in the subsequent chapters.

3.2.1 Logical concepts and XML syntax

Although this chapter is about the conceptual introduction of logical GHSTS concepts, some technical XML terms have to be introduced for better understanding:

An element in the GHSTS XML backbone file is used to represent concepts. The following incomplete XML excerpt shows an example:

```
<documents>
  <document id="1234">
    <document_title>Acute oral LD 50 in the Chinese hamster</document_title>
    <document_author>Sarasin P.</document_author>
    <document_pid>65eb0ea0-7452-11e2-bcfd-0800200c9a66</document_pid>
  </document>
  <document>
    <document_title>Determination of the acute dermal LD 50 to the rat</document_title>
    <document_author>Hurni H; Sachsse K</document_author>
    <document_pid>f72bef30-7455-11e2-bcfd-0800200c9a66</document_pid>
  </document>
</documents>
```

Figure 7 – Example of XML document

An element consists of a start tag (e.g. <documents>) and an end tag (e.g. </documents>). The element <documents></documents> represents the concept of a document list; the nested element

`<document></document>` represents the concept of a document. The element `<document>` has element content, because it only contains elements as content, the element `<document_title>` has text content.

The element `<document>` is a *direct* child (or subnode) of the element `<documents>`. The element `<document_title>` is an indirect child (or subnode) of the element `<documents>`, because it is also nested within the `<document>` element. Elements that are cross-referenced do not count as direct or indirect children.

An element may also possess attributes that appear in the start tag of the element. The element `document` in the example above possesses an attribute “id”.

Note: The term “attribute” within this document is used only to designate a specific syntax within an XML file. It is not used as synonym for metadata.

3.2.2 Overview

The following figure shows the concepts of GHSTS - as represented in the GHSTS XML backbone - and the internal relationships that are not specific for an individual receiver. Internal relationships are relationships between elements of the same submission, this means within the GHSTS XML backbone file.

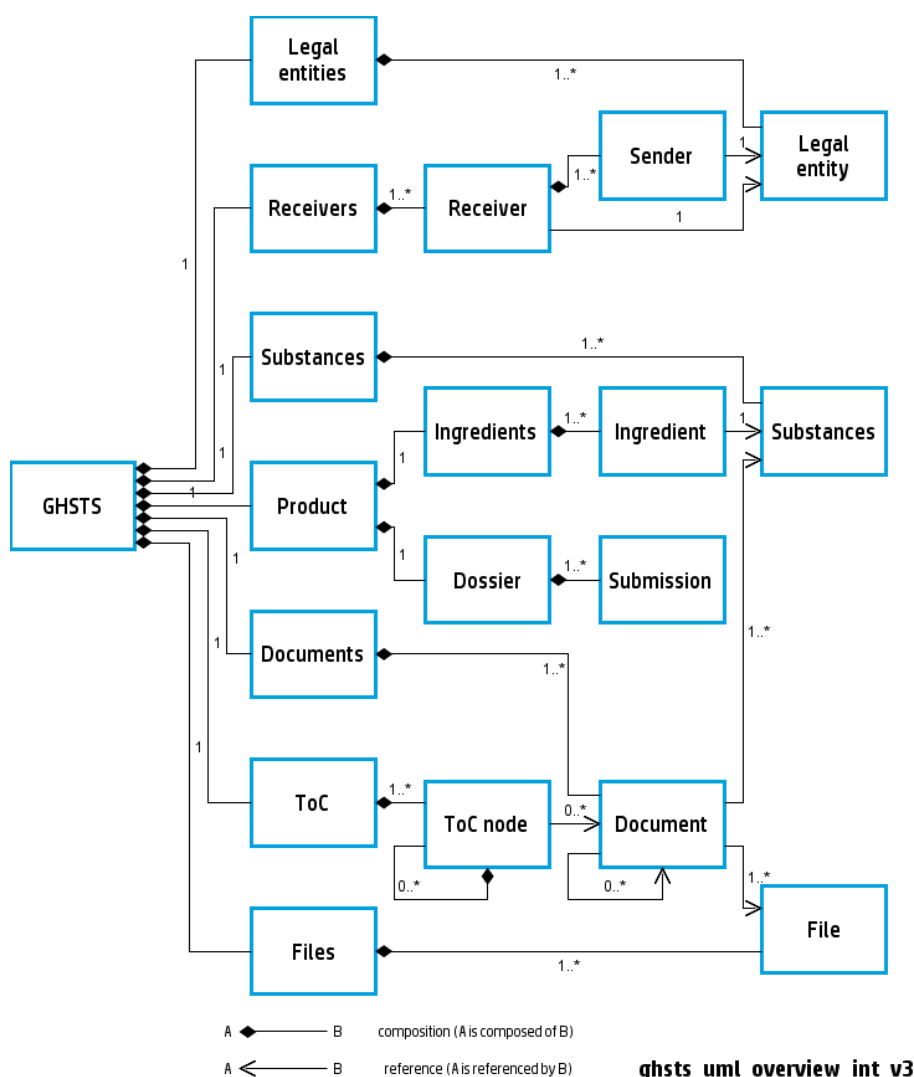


Figure 8 – Overview of GHSTS concepts

The relation “composition” is represented by nesting in the XML backbone file: if A is composed of B then B is syntactically nested in A in the XML file. The relation “reference” is represented by elements or attributes containing the ID of the referenced element.

For an overview about receiver-specific references see chapter 3.2.24, for an overview of external references see chapter 3.2.24.

3.2.3 Metadata

The term “metadata” or more precise “descriptive metadata” is used to denote information about a logical concept, for example a dossier (dossier name), a document (document title) and other concepts of GHSTS. It is set in contrast to the “content” of a concept (see concepts FILE and DOCUMENT, chapter 3.2.9 and 3.2.12).

For XML-based content the notion of metadata and content is blurred, as both the GHSTS XML backbone containing the metadata and the XML content have the same XML format. In general the metadata for FILE and DOCUMENT is used to select and identify the correct computer file and is mostly a subset of the identifying information that is part of the content of the computer file itself, independent on its file format.

In a submission package the metadata is represented as elements or attributes within the GHSTS XML backbone.

3.2.3.1 *Formatting of metadata*

It is recommended to avoid the formatting of metadata. Special characters like tabs, LF, CRLF, NL and multiple white spaces should be avoided. The same applies to supplying XML fragments as text content of GHSTS elements. Such characters can make the display, for example in the GHSTS viewer, unusable. The formatting of the information is up to the receiver, the registrant should only supply the unformatted information.

The only exception to this rule is the element LEGALENTITY_NAME where line feeds are allowed.

3.2.3.2 *Legal metadata characters*

The character encoding that can be used to submit information in the GHSTS XML backbone file is UTF-8. With this encoding all characters from the Unicode character set can be represented. The latest version of Unicode contains a repertoire of more than 110,000 characters covering 100 scripts.

All common web browsers are capable of displaying Unicode characters if the required fonts are installed on the computer. Eventually the missing fonts have to be downloaded and installed. Characters for missing fonts are mostly displayed by the sign „□“.

Please note that Unicode includes a variety of options (e.g. right-to-left languages, Chinese, Japanese, and Korean (CJK) fonts) that require the support not only of the XML backbone file, but also of other GHSTS components (GHSTS validator, GHSTS viewer, publication and ingestion component, dossier management system). Therefore it is recommended that registrant and receiver agree on the features of Unicode that can be used.

3.2.3.3 *Metadata changes*

GHSTS supports the tracking of metadata changes for a number of concepts that are represented in the GHSTS XML backbone. The following values are distinguished:

Value	Description
New	The metadata belongs to a new instance of the concept that has not been part of the previous submission. The value “New” is used for instances that are completely new without predecessor or that are replacing other instances.
No change	The metadata of the instance of concept did not change compared to the previous submission.
Modified	The metadata of the instance of the concept changed compared to the previous submission; the instance itself is still part of the submission.

Table 3 – Values for metadata change

The tracking is done by setting one of the above values to the element METADATA_STATUS of the concept. Usually the metadata status is set for all elements in the GHSTS XML backbone that are within the same block or its children. However there are some exceptions to this rule. The following table lists the concepts within GHSTS that possess the element METADATA_STATUS and explains the scope of this element.

Concept	Description of concept	Scope of tracked metadata changes
LEGAL_ENTITY	Complete Information for receiver and sender see chapter 3.2.8	Any element within LEGAL_ENTITY
RECEIVER	Specific additional metadata for receiver and sender, added to information in LEGAL_ENTITY	Any element within RECEIVER, including the child SENDER concept.
SUBSTANCE	Concept for substance	Any element within SUBSTANCE
PRODUCT	Concept for product, referencing the substance concept	Any element within PRODUCT and all direct and indirect children, except for the SUBMISSION element. This includes changes to the set of used ingredients and their quantity. The SUBMISSION element changes by default for every submission; hence these changes are not reported.
FILE_RA	Concept for RA-specific metadata on the FILE level	Any element within FILE_RA
FILE_GENERIC	Concept for generic metadata on the FILE level	Any element within FILE_GENERIC except for the block CONTENT_STATUS that contains the information for content changes.
DOCUMENT_RA	Concept for RA-specific metadata on the DOCUMENT level see chapter 3.2.12	Any element within DOCUMENT_RA
DOCUMENT_GENERIC	Concept for generic and bibliographic metadata on the DOCUMENT level see chapter 3.2.12	Any element within DOCUMENT_GENERIC except for a) the block CONTENT_STATUS_HISTORY that contains the information for content changes b) references to the FILE concept (REFERENCED_TO_FILE) that can change with every (full) submission. The scope includes changes to references to other documents or references to substances
TOC	see chapter 3.2.7	Any element within the TOC concept. This includes changes in the actual ToC structure as well as any changes in the references between ToC nodes and documents, e.g. when documents are “added” to a ToC node.

Table 4 – Concepts that track metadata changes

3.2.4 *Dossier regulatory action*

A dossier regulatory action is a series of events accomplished by a regulatory authority beginning with the submission of an application with data by a registrant and ending with a unique regulatory decision. During one regulatory action the applied standardised Table of Contents (e.g. EU, OECD, national) has to remain the same. It is up to the communicating parties to decide what kind change of

- Registrant
- Regulatory authority (abbreviated as “RA”)
- Substance
- Product

could be covered under the same regulatory action or requires a new regulatory action.

Within GHSTS a dossier regulatory action is identified with a PID called DOSSIER_PID which remains unchanged throughout all submissions. In addition GHSTS supports the identification of a dossier with an identification assigned by the regulatory authority using the element PROJECT_ID_NUMBER.

3.2.5 *Submission*

A submission is the compilation of documents in a structured form according to the given regulatory requirements, as part of a dossier regulatory action.

The physical representation of a submission on a media is called submission package.

In GHSTS a submission within a dossier regulatory action is identified with the SUBMISSION_NUMBER.

The concept of submission in the GHSTS XML backbone file is represented by the set of metadata, consisting of submission number, submission version date, submission title and the information whether the submission is incremental or not (see chapter 3.2.19).

Each submission package contains this information for the current and all preceding submissions – if any - of the regulatory action.

3.2.6 *Dossier*

A dossier is the set of information set up by the registrant for the purpose to serve a specific regulatory action and updated according to the process of the regulatory action. A GHSTS submission captures the state of a dossier at a certain milestone in GHSTS format.

For each regulatory action the registrant has to create a separate dossier with a new DOSSIER_PID.

3.2.7 *Table of Contents*

The Table of Contents (ToC) that is used as part of the GHST XML backbone is identified by the element TOC_SHORT_NAME.

The registrant has to use the adequate ToC from the predefined ToC (see chapter 2.4.8). It is not permitted to change the ToC within the regulatory action.

Each GHSTS backbone XML file contains a complete copy of the standard Table of Contents, even if not all ToC nodes are filled. This requirement is the pre-condition to read an old dossier with the help of the GHSTS viewer years later with exactly the same context, which was valid at submission time. That means

“freezing” the full ToC into each GHSTS submission package makes GHSTS non-sensitive regarding the version management of standard TOC.

The GHSTS backbone can contain an optional reference to the XML file of the applied standard ToC (see **Figure 5 – Overview of ToC references**). This reference can be used by a validator to check whether the embodied ToC in the GHSTS backbone XML is identical to the applied standard ToC. The reference is represented by a logical PID, by the TOC_SHORT_NAME and the physical reference to a ToC XML file (see chapter 2.4.8) that has to be part of the submission package. However the GHSTS validator will preferably validate the Toc in the XML backbone against the ToC instance on the centrally managed platform, using an online connection.

3.2.7.1 *OECD tables of Contents*

The OECD encourages regulatory authorities to adopt the use of the OECD Table of Contents for dossiers in order to allow for easier communication among regulatory authorities during Joint Reviews and to reduce the regulatory burden on applicants. The OECD Table of Contents is utilized for both joint and many national submissions.

The OECD makes available different dossier ToC in XML format according to the ToC XSD described in chapter 5.2 in order to be incorporated easily into the meta-data of an electronic pesticide package built according to GHSTS. For the list of available ToC and where they can be found please see chapter 2.4.9.

3.2.7.2 *National tables of Contents*

Though the OECD encourages regulatory authorities to adopt the use of the OECD numbering system for dossiers, other numbering systems exist and national/regional applications of different types (e.g. end-use product authorisations with active substances already evaluated, for use extensions of existing products, for MRL setting, for EU zonal authorisations or mutual recognition) remain a major workload for most OECD member countries. GHSTS serves as a submission tool for these other types of applications and can incorporate national numbering systems or the OECD numbering system in the meta-data of an electronic pesticide package to describe the files being transported.

To this end, Regulatory authorities have the responsibility to make their national ToCs available in XML format to the OECD, if they choose to have it available to be incorporated easily into the meta-data of an electronic pesticide package built according to the GHSTS. For the list of available ToC and where they can be found please see chapter 2.4.9.

It is recommended that, if a dossier is submitted to more than one regulatory agency, the appropriate OECD ToC (for chemicals, micro-organisms or pheromones) should be used rather than using a national ToC. Ultimately, regulatory authorities decide which ToC should be used.

3.2.8 *Legal entity*

The concept of legal entity summarizes all contact information both for the registrant(s) and regulatory authorities.

Whether an entry for a legal entity serves as information for sender or receiver is specified separately in the concepts sender and receiver by cross-referencing the corresponding entry. Sender and receiver only contain few additional metadata to specify their role.

GHSTS supports the definition of multiple receiver and for each receiver the definition of individual multiple sender.

3.2.9 *Product*

A regulatory action serves to receive a regulatory decision for a certain product. There are two kinds of products:

- A product as a formulated product / composition of multiple ingredients. An ingredient in the GHSTS backbone XML is the reference to a substance (see chapter 3.2.10) specifying its quantity in the product.
- A product as Technical Grade Active Ingredient (U.S.)/Active Substance (EU). In this case the product name is set to the common name of the Active Ingredient / Active Substance and the only ingredient of the product is this active substance.

These two kinds of products are reflected by two types of dossiers with different studies:

- Dossiers for Technical Grade Active Ingredient (U.S.)/Active Substance (EU)
- Dossiers for formulated products

Each product is identified by a PID, also for a TGAI (Technical Grade Active Ingredient). For example, the TGAI “saflufenacil” will get a PID and will get the generic product name “saflufenacil” in the GHSTS submission. In addition the registrant can supply specific product names for each receiving regulatory authority, for example.

- BAS 800 H
- saflufenacil
- saflufénacil
- saflufenacil TC
- benmihuangcaoan

Within the GHSTS XML backbone the product is in both cases identified with a unique identifier that will be created by the submitting company. It is not allowed to change this identifier during the ownership of the company; hence the identifier has to be reused for all regulatory actions concerning this product. The identifier of the product designates the ingredients in the product and is independent of the product name, so the unique identifier will remain stable even when the product name changes. A new product identifier has to be generated when the composition of active ingredients / the active ingredient is changed.

3.2.10 *Substance*

In the GHSTS XML backbone for a submission the registrant has to list all relevant reference-substances that are part of the product - which are the active ingredients – as well as the metabolites and impurities of the product. Impurities of active ingredients and co-formulants do not have to be supplied as part of the GHST XML backbone.

The substances are referenced by the product in the GHSTS XML backbone, this reference is called ingredient. Substances can also be optionally referenced by single documents, for example for metabolites and impurities.

The registrant supplies the common name of the substance if available, otherwise an internal name.

Next to the CAS number or other substance identifier types the substance is also (optionally) identified with a unique identifier that is centrally managed (see chapter 3.2.20).

3.2.11 File

In the context of an electronic submission the term “file” is usually used as synonym for computer file, as being the basic resource for storing information on a computer. Each computer file possesses a content plus some additional file attributes (e.g. size, file creation date). The latter are managed by the operating system.

For GHSTS this concept of “file” is not sufficient. It is not possible to easily model additional GHSTS-specific metadata and to track the lifecycle of a computer file during a submission. Therefore GHSTS contains the concept of a FILE (in capital letters).

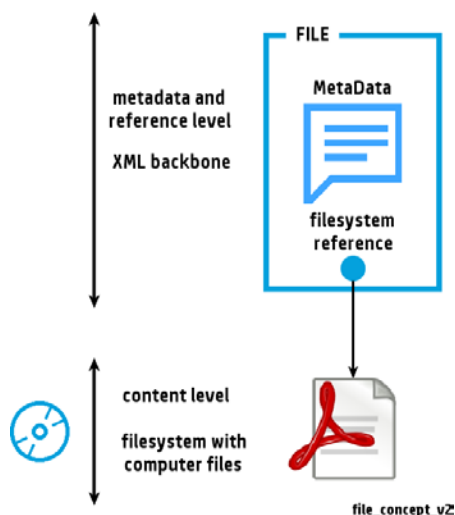


Figure 9 – logical concept of FILE

As shown in the preceding figure, a FILE is a concept within the XML backbone and is composed of

- Metadata (e.g. a file ID supplied by the company)
- One single reference from the XML backbone to one file on the file system (here a PDF file)

In the remainder of this document, the term “FILE” will be used to identify this GHSTS concept, whereas “file” will be used as synonym for computer file.

3.2.12 Document

In times of digital technology the term “document” has to be defined functionally and from a semantic point of view, as the technical representation of a document may change (e.g. having a new PDF representation including bookmarks) whereas for a human stakeholder the “document” has not changed its content.

To support the idea of a document from a functional viewpoint (whatever functions as a document *is* a document) with multiple representations GHSTS has developed the concept of a DOCUMENT (in capital letters).

A sample instance of a DOCUMENT is shown in the figure below:

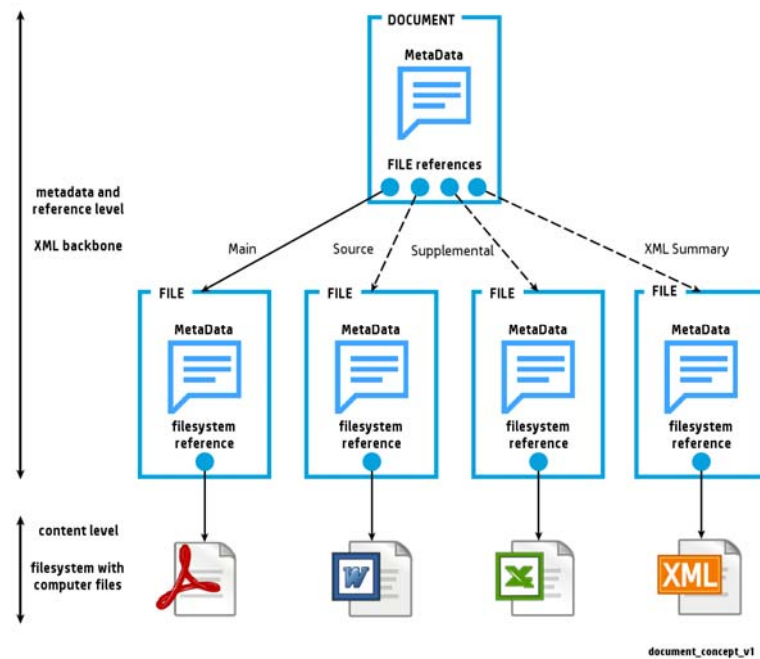


Figure 10 – logical concept of DOCUMENT

An instance of a DOCUMENT is a compound object of multiple representations, modelled in GHSTS in the FILE concept plus the metadata of the DOCUMENT itself. The metadata of the DOCUMENT propagates to the referenced FILE instances: If, for example, the data protection claim is set in a document instance then this information applies to all referenced FILE instances.

The content of a document is comprised of the set of computer files it references via the FILE concept. Each file is assigned one file type (element FILE_TYPE); this specifies the role for the document from which it is referenced:

Type of file	Classification	Interpretation
Main	core representation for the regulatory review, containing the full content	The FILE instance references a computer file in PDF/A format. It contains the full content information. The evaluator has to start reading the document content with this file. A document can only have one reference of type “Main”.
Source	Attachment An attachment can contain either full, partial or supplemental content information compared to the content of the core representation.	The FILE instance references a computer file containing the same full content and layout as the “Main” computer file in editable format, e.g. in Microsoft Word format.
Supplemental		The FILE instance references a computer file with supplemental information for the review, e.g. the raw data for an evaluation, high resolution pictures, data files for statistic programs, Excel-Sheets, etc.
XML Summary		The FILE instance references a computer file with an XML Summary in the OECD Harmonised Template (OHT) XML format for Reporting Chemical Test Summaries

Table 5 – Overview of types of FILE

Please note the following constraints:

- The element FILE_TYPE is part of the FILE – so the role of each file is static and not dependent on references.
- Every DOCUMENT instance must be contained only once in the submission. Each DOCUMENT instance must be associated with one to many ToC nodes (data points).
- Every FILE instance must be referenced from exactly one DOCUMENT instance. A reuse of FILE instances between multiple DOCUMENT instances is not permitted. Eventually these FILE instances have to be submitted multiple times.
- If it is necessary to split one document into different files to minimize the file sizes, for example of large toxicological study reports, it is allowed to differ from the above interpretation and to mark only the first file as “Main” and the subsequent parts as “Supplemental”.

3.2.13 Document Family

A registrant can group sets of documents in its dossier management system that are common in the following domains:

- The documents have evolved from each other (the term “version” is not used in GHSTS, but one may think of a version lifecycle of a document in a document management system to yield multiple document instances).

- There are tight semantic relationships between the documents, so if one document evolves it is likely that the other document also will evolve. An example for such a relation is a translation.

All document instances that share these characteristics are called a document family. Within a dossier a document instance can only be replaced with another document instance from the same document family (see chapter 3.2.17.)

GHSTS supports the concept of a document family with two depending elements as synonyms:

- The element DOCUMENT_FAMILY carries a short textual identifier.
- The element DOCUMENT_FAMILY_PID carries a unique identifier for the document family.

For each document instance it is recommended to use the value of the element DOCUMENT_FAMILY as leading string of the value for the element DOCUMENT_NUMBER.

Both elements carry the same additional information that can be used during viewing or machine processing. They are not required to uniquely express relations between documents, neither in the evolution nor the semantic aspect.

However with the help of the document family concept, it is possible to filter the original document and its translation in the GHSTS viewer if both documents are submitted in the same GHSTS submission package and both documents have the same document family value.

It is strongly recommended to administrate the values for DOCUMENT_FAMILY_PID and DOCUMENT_FAMILY synchronically that both can act as unique identifiers on the regulatory action level.

3.2.14 File formats

The term file format describes the encoding of information in a computer file. GHSTS is no archiving standard, but a transport standard, however also the aspect of long-term usability of the content in the computer files should be considered.

GHSTS recommends supplying the “Main” file types in a specific PDF format called GHSTS PDF profile. The characteristics of this PDF format are described in a GHSTS PDF profile in the next subchapter.

For the computer files of other file types there is no imposed constraint with respect to the file format. If possible GHSTS recommends using file format types that are established, standardized, widespread and support a long-term visualization and lossless migration. It is expected that word processing, spreadsheet and XML files will be used.

The following characteristics of file formats are not recommended:

- A computer files that is zipped
- A computer file representing an email message (e.g. an Outlook .msg file)
- A password protected, encrypted or digitally signed computer file

Regional business rules can add or alter those constraints.

If a user opens the computer file on a computer where the file extension is not known usually a message box will pop up and the user is requested to select an appropriate viewer application. Therefore the registrant should only use file formats where he is sure that the receiver has the appropriate viewer / editor software already installed.

3.2.14.1 GHSTS PDF profile

The GHSTS PDF profile describes some characteristics of the PDF that is recommended for the “Main” file type. It is based on the PDF/A-1 format, as defined in ISO-19005-1:2004. The profile defines additional constraints on top of this base standard and partly relieves constraints from the base standard. The profile is not a complete definition of the file format that allows strict format verification, but a set of recommendations.

Please note that there is currently no tool available that allows verifying the compliance of a PDF document with the PDF/A standard 100% accurately. It is currently not planned to include a complete file format validation into the GHSTS validator component.

The following table defines the GHSTS PDF profile in detail.

	GHSTS PDF profile
Allowed	External references Restriction: Only hyperlinks to computer files that are part of the current submission package in the content folder (see chapter 6.3) are to be used. The hyperlinks in the PDF files should adhere to the relative path definition in chapter 4.10.1. Hyperlinks that refer to individual chapters or pages of a target PDF document are not recommended, as this functionality is not stable.
Recommended	Embedding of all fonts Colour spaces specified in a device-independent manner Inclusion of document structure Language specification Use of metadata
Not recommended	Transparency / overlay Audio / video content Encryption and password protection Digital signatures Embedding of arbitrary file formats Usage of dynamic content (JavaScript) and executable file launches

Table 6 – GHSTS PDF profile

3.2.15 Type of content changes

During a regulatory action the computer files attached to instance of FILE - and indirectly DOCUMENT - can change. GHSTS distinguishes two types of content changes:

- Changes on the technical level that have no influence on the content of a computer file, as evaluated by a human being. This type of change is called “Modification”. Examples of Modifications would be:
 4. Regeneration of a PDF with different settings (e.g. bookmarks)
 5. Correction of typos on the letter level
 6. Change of format for a supplemental file (e.g. using a later Office version)
- Changes on the content level as perceived by a human being that may lead to a different evaluation. This type of change is called “Replacement” as it leads to a new instance of a FILE or DOCUMENT that replaces the preceding instance.

These types of changes cannot be distinguished on a technical level. It is therefore within the responsibility of the registrant to clearly and carefully distinguish those two types of changes.

Both the concepts of DOCUMENT and FILE possess an element CONTENT_STATUS that is used to track the content changes. This element is also called content status indicator.

3.2.16 Content status changes of files

The following figure shows a scenario with content changes for a FILE instance.

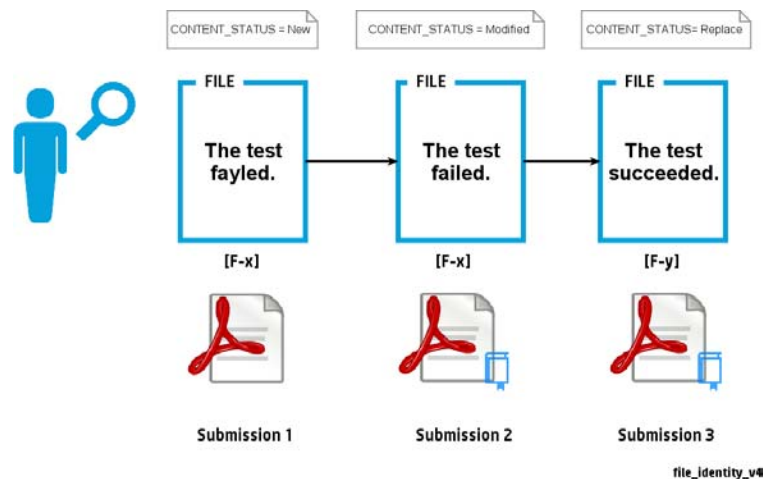


Figure 11 – content change scenario for FILE instance

- In submission 1 the file is new, so the value of CONTENT_STATUS is set to New. The file has a unique identifier [F-x].
- In submission 2 the computer file of the FILE instance is modified, a typo is corrected and the PDF is generated with bookmarks. The registrant has classified this as modification; hence the CONTENT_STATUS is set to Modified. The previous identifier [F-x] is kept for a modification.
- In submission 3 the content of the computer file is changed in a way that alters its meaning. Therefore the change is classified as replacement and CONTENT_STATUS is set to Replace. The identifier is changed to [F-y]; the new FILE instance is considered a different file instance than its predecessor [F-x].

The symbols for the person and the magnifier are added to highlight that the differentiation between “Modified” (no change of semantics) and “Replace” (change of semantics that may lead to a different evaluation) can only be done by a human being.

In total there are 5 different values for the content status indicator. The following table lists the values and their meaning for the FILE concept:

Value	Description
New	New file instance with new computer file instance that had not been part of the dossier before.
No change	Referenced computer file did not change on a binary level compared to the previous submission. An import component can safely skip this computer file.
Modified	The referenced computer file has been modified on a binary level, but without influence on its semantic content and evaluation result. An example would be the correction of a typo without influence on semantics. The unique identifier of the modified FILE instance does not change.
Replace	The referenced computer file has been altered in a way that the semantic content has changed and the evaluation of the content may lead to a new result compared to its predecessor. The replacing FILE instance has a different unique identifier and is considered a different instance than its predecessor. The predecessor FILE instance and the current FILE instance have to be referenced from DOCUMENT instances that are part of the same document family.
Retired	The content of the computer file is no longer considered relevant for the dossier.

Table 7 – Content status indicator for FILE

In the following sections the content status indicators are explained in more detail. In the figures the right column stands for the latest submission where the content status indicator in the element CONTENT_STATUS has to be set. The left column represents its predecessor submission.

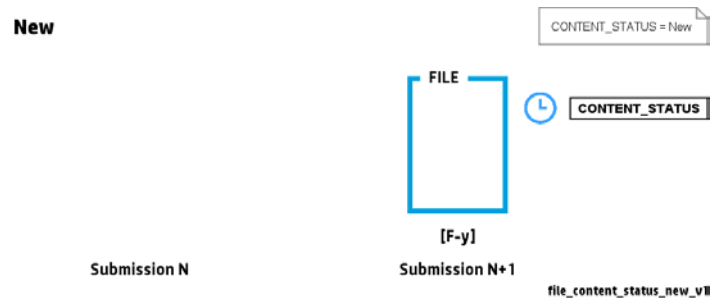


Figure 12 – FILE - content status indicator “New”

The CONTENT_STATUS “New” is set for FILE instances with computer files that have not been part of the dossier before and that are not replacing another FILE instances (see below).

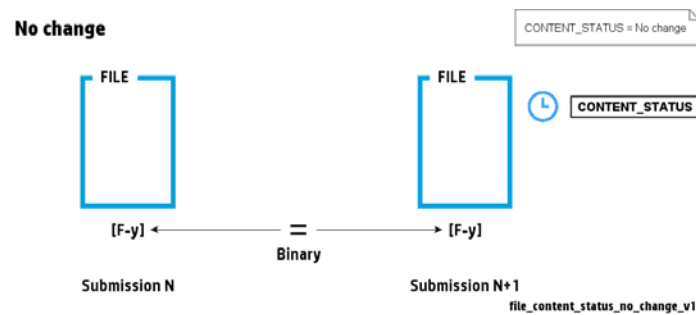


Figure 13 – FILE - content status indicator “No change”

The CONTENT_STATUS “No change” is set for computer files that are identical on the binary level compared to their predecessor FILE instance.

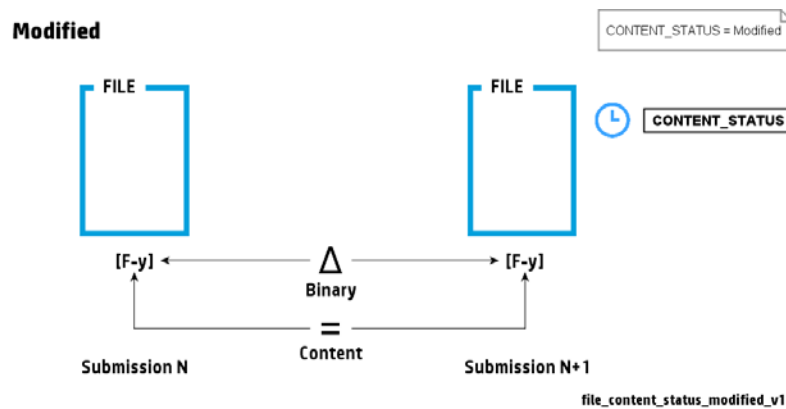


Figure 14 – FILE - content status indicator “Modified”

The CONTENT_STATUS “Modified” is set for computer files that are different on a binary level but identical on the content level compared to their predecessor computer file. The identifier is not changed between the submissions; a content modification does not trigger the creation of a new FILE instance. The sign “Δ” in the figure stands for the Greek character “delta” and denotes changes.

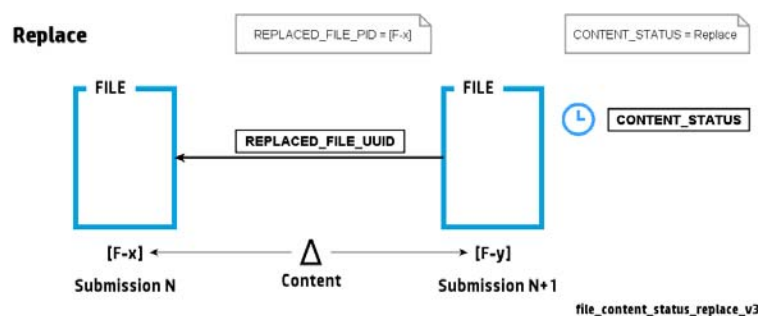


Figure 15 – FILE - content status indicator “Replace”

The CONTENT_STATUS “Replace” is set for computer files that have changed on the content level compared to their predecessor computer file. A new file instance with a new identifier [F-y] is created.

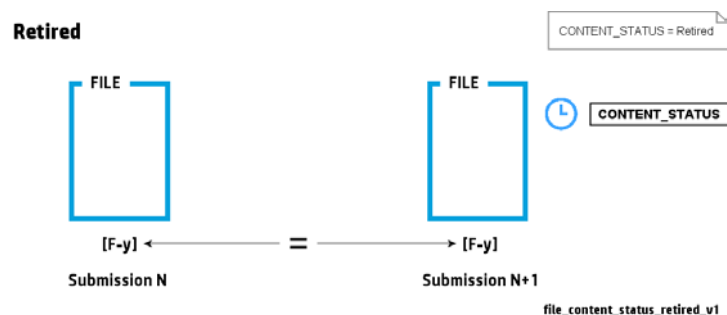


Figure 16 – FILE - content status indicator “Retired”

The CONTENT_STATUS “Retired” is set for computer files that did not change but are to be removed from the dossier. Content and metadata of the FILE instance does not change.

The FILE concept possesses the element MD5CHECKSUM that contains a checksum of the computer file. Please note that the value of this element serves only as check to detect transmission errors, it is no indicator of content changes. The responsibility to correctly set the content status indicator is within the registrant, independent on the MD5 value.

3.2.17 Content status changes of documents

As a DOCUMENT itself does not carry content directly the content status of a DOCUMENT is derived from the value of the content status indicators of its referenced FILE instances. Please remember that there are two types of FILES, “Main” files and “attachment” files (see “Table 5 – Overview of types of FILE”).

Different to the FILE instance where only the last content change is kept, each DOCUMENT instance keeps a complete content status history throughout its existence. The following figure shows a summary of the content status information for FILE and DOCUMENT, as well as the metadata information.

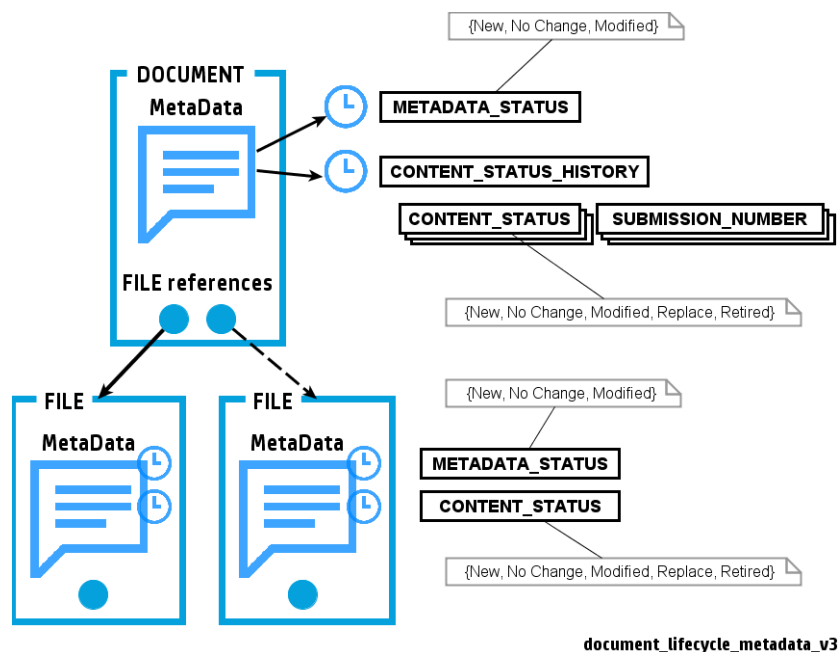


Figure 17 – Lifecycle information for FILE and DOCUMENT

The element CONTENT_STATUS_HISTORY keeps track of the content changes since this DOCUMENT instance with this particular identifier has been added to the dossier. Each entry in CONTENT_STATUS_HISTORY consists of a set of associated values for CONTENT_STATUS and SUBMISSION_NUMBER. For each change of the document instance a new set of values is added.

As a first rule of thumb the content status indicator of a DOCUMENT instance is set to the value of the content status indicator of its “Main” FILE instance. Before discussing the exceptions the following figure shows a scenario following this rule:

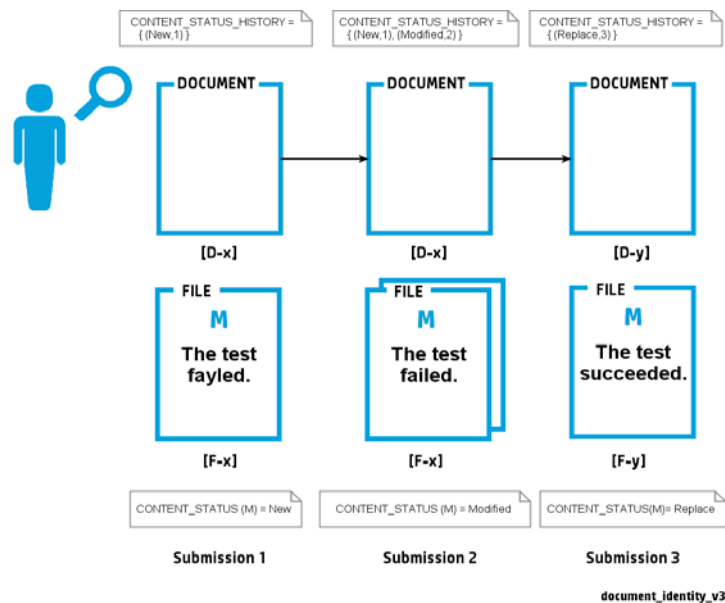


Figure 18 - content change scenario for DOCUMENT instance

The DOCUMENT instances are shown in the first row, the associated FILE instances directly below in the second row. The letter **M** stands for the “Main” FILE instance of the DOCUMENT instance. The CONTENT_STATUS (M) represents the content status indicator of the “Main” FILE instance.

- In submission 1 both FILE and DOCUMENT instance are new.
- In submission 2 the “Main” file has been Modified and an additional FILE instance has been added as attachment (with CONTENT_STATUS = New). The content status of the DOCUMENT instance [D-x] is set to Modified, a new entry (Modified,2) is added to the CONTENT_STATUS_HISTORY.
- In submission 3 the “Main” file has been Replaced, in addition the attachment has been removed (CONTENT_STATUS=Retired). The replacement triggers the creation of a new document instance with identifier [D-y]. The new instances possesses a fresh CONTENT_STATUS_HISTORY with the initial entry (Replace,3)

This rule of thumb is not correct for all possible combinations of content status indicator for “Main” and “attachment” FILE instances. The following table shows, how the content status indicator of a DOCUMENT instance is to be set depending on the content status indicator of its “Main” and “attachment” FILE instances.

Cells contain derived values of document content status		Content status of “attachment” file(s)				
		New	No change	Modified	Replace	Retired
Content status of “Main” file	New	New	n.a.	n.a.	n.a.	n.a.
	No change	Modified	No change	Modified	Modified	Modified
	Modified	Modified	Modified	Modified	Modified	Modified
	Replace	Replace	Replace	Replace	Replace	Replace
	Retired	n.a.	n.a.	n.a.	n.a.	Retired

Table 8 – Derivation of status indicator for DOCUMENT

The table has to be used as follows:

- Determine the value of the content status indicator of the “Main” FILE instance. This results in the row to look at.
- Determine the value(s) of the content status indicators of the “attachment” FILE instance. This results in the column to look at.
- The value in the intersection cell returns the value for the content status of the DOCUMENT instances.

Please note the following constraints:

- Some combinations are not possible and marked with “n.a” (not applicable): As FILE instances cannot be referenced by multiple DOCUMENT instances an attachment FILE instances cannot possess any other value than “New” when the “Main” file has the content status indicator set to “New”. A similar thought applies to “Retired”
- The above “rule of thumb” is only violated for the row with content status indicator of “Main” = “No change”. Here, if any attachment FILE instance has any value other than “No change” then the content status indicator of the DOCUMENT instance is set to “Modified”.

The following table summarizes the usage of the content status indicators for DOCUMENT:

Value	Description
New	This DOCUMENT instance is new in the dossier, as “Main” FILE instance is new in the dossier
No change	All referenced FILE instances (“Main” and “attachment”) have the content status indicator “No Change”
Modified	The content status indicator is set to “Modified” in one of the following scenarios: “Main” FILE instance is “Modified” “Main” FILE instance is “No change”, but at least one attachment file has one value other than “No Change” – so “New”, ”Modified”, ”Replace”, ”Retired”. The unique identifier of the modified DOCUMENT instance does not change. The result of the evaluation of the modified document by the regulatory authority should not change compared to the evaluation of its predecessor document.
Replace	This DOCUMENT instance is set to “Replace”, as the “Main” FILE instance is replacing its predecessor. The replacing DOCUMENT instance has a different unique identifier and is considered a different instance than its predecessor.
Retired	This DOCUMENT instance is retired, as the “Main” FILE instance is retired.

Table 9 – content status indicator for DOCUMENT

In the following sections the content status indicators are explained in more detail. In the figures the right column stands for a submission in focus where the content status indicator in the element CONTENT_STATUS has to be set. The left column represents its predecessor submission.

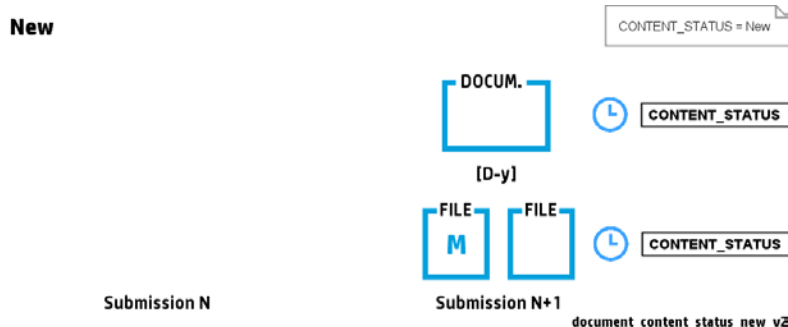


Figure 19 – DOCUMENT - content status indicator “New”

The scenario “New” is straightforward and does not require additional explanation.

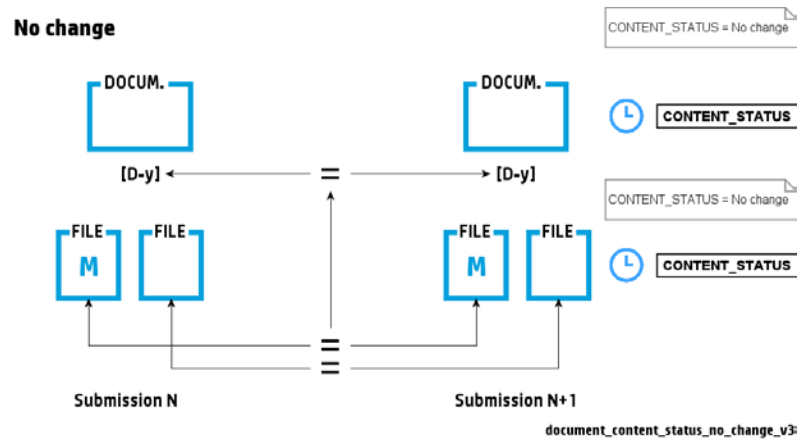


Figure 20 – DOCUMENT - content status indicator “No Change”

The content status “No change” is set for the DOCUMENT instance when all FILE instances have content status indicator “No change”.

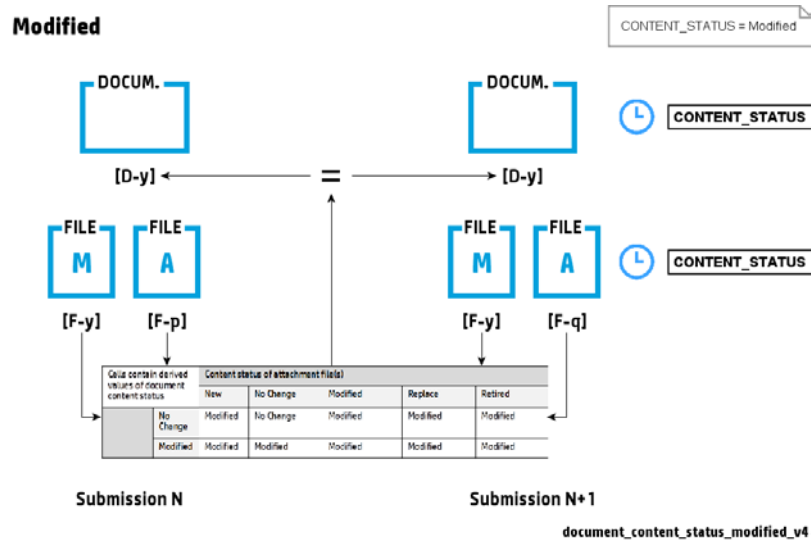


Figure 21 – DOCUMENT - content status indicator “Modified”

The content status “Modified” is set according to the “Table 8 – Derivation of status indicator for DOCUMENT”. Please note that the identifier of the DOCUMENT instance does not change.

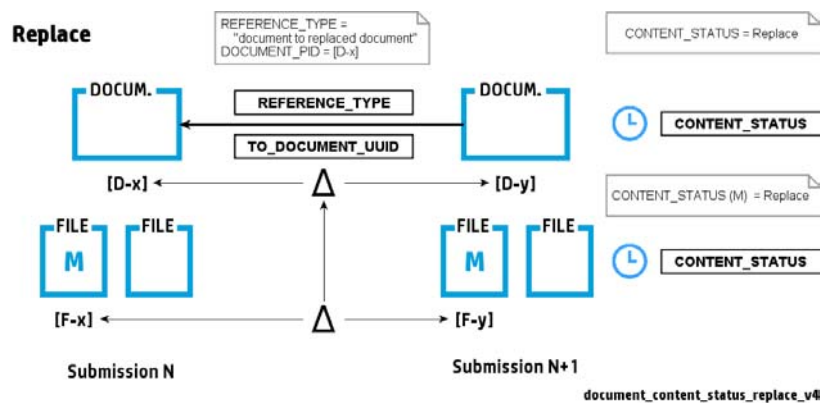


Figure 22 – DOCUMENT - content status indicator “Replace”

The content status “Replace” is set in the DOCUMENT instance in submission N+1 when the “Main” FILE instances has content status “Replace”. To be able to locate the predecessor document in submission N a document reference from [D-y] to [D-x] is established by setting the elements REFERENCE_TYPE to the value “document to replaced document” and the element DOCUMENT_PID to the unique identifier (PID) of [D-y].

The replacing DOCUMENT instance [D-y] must satisfy the constraint that it is part of the same document family as the replaced document [D-x] (see chapter 3.2.13). Hence the elements DOCUMENT_FAMILY and DOCUMENT_FAMILY_PID have to be identical for [D-x] and [D-y].

If a document [D-x] must be “replaced” with a document [D-y] from a different document family, because, for example, a document has been part of a dossier in error, then this has to be modelled by a parallel retirement of [D-y] and the addition of a new document [D-x] in the submission N+1.

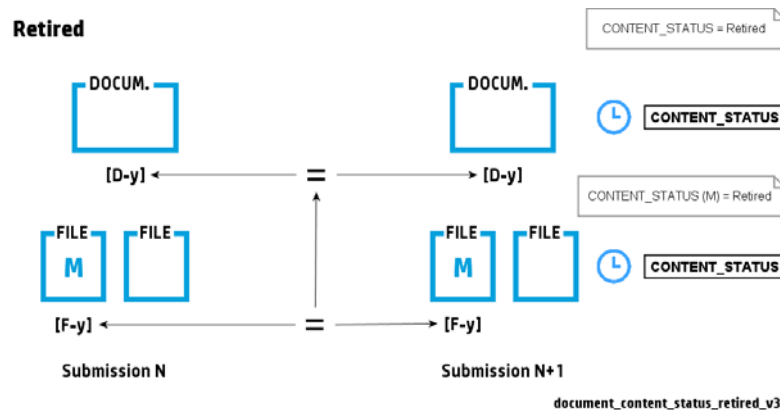


Figure 23 – DOCUMENT - content status indicator “Retired”

If the registrant wants to retire a DOCUMENT instance the content status of the DOCUMENT instance as well as of FILE instances have to be set to “Retired”.

3.2.18 Dependencies between content status changes and metadata changes

As shown in “Figure 17 – Lifecycle information for FILE and DOCUMENT” the concepts of DOCUMENT and FILE possess both elements for metadata changes and content changes.

There are the following dependencies between content status changes and metadata changes for DOCUMENT and FILE:

- If the content status indicator of the concept is “New”, then its metadata status value is also set to “New”.
- If the content status indicator of the concept is “Replace”, and hence a new instance is part of the dossier, then the metadata status value is set to “New” for the new instance.

3.2.19 *Full and incremental submissions*

GHSTS supports full and incremental submissions. The following table defines two those terms. They relate to the set of computer files that are referenced by the set of FILE instances.

Submission	Description
Full	The submission package contains all computer files of FILE instances that are part of the GHSTS XML backbone independent on content status indicator. The submission package also contains the computer files that are set to “Retired”.
Incremental	The submission package contains only computer files of FILE instances that are part of the GHSTS XML backbone and that have the content status indicator set to either “New”, “Modified” or “Replace”. The file system references of each FILE instance with “No change” and “Retired” point to the submission with the last change of the FILE instance (“New”, “Modified”, “Replace”) or the latest full submission, dependent on what submission was the latest in the submission history.

Table 10 – full and incremental submission

The initial submission of a regulatory action is always full. For each subsequent submission it is up to the communicating parties to decide whether a full or incremental submission should be created.

Please note that the term “incremental submission” only refers to the referenced file content, not to the GHSTS XML backbone. The XML backbone always contains the complete information of the dossier in its current state. The set of FILE and DOCUMENT instances are identical in both cases, only the file system references would differ: In case of a full submission they point to file within the current submission package, for incremental submission they point to a previous submission package.

The only instances that may technically vanish in the GHSTS XML backbone in the course of a regulatory action are FILE or DOCUMENT instances that are replaced by successor instances with new identifiers. However, as usually the replacement is done with similar information (e.g. document from the same document family) the new information is not completely unrelated to the old one.

Computer files already submitted under a different regulatory action should be resubmitted in a full submission; however it is incumbent upon the regulatory authority to decide on the base of the supplied unique identifiers (see chapter 3.2.20) how to consume it. Therefore the careful assignment of unique identifiers is crucial.

3.2.20 *Persistent identifiers*

GHSTS enforces the usage of Persistent Identifiers - or short PID – for a number of concepts in GHSTS. In the pictures above the identifiers like [D-x] and [F-x] are all PID. The main driver for PID is to avoid duplicate information and to be able to uniquely identify an instance of an information concept throughout its lifetime independent on, for example, variations in spelling.

Usually the creation and management of PID can be done within the scope of the registrant. The registrant is responsible to carefully decide when a new instance is required that requires a new PID (see chapter 3.2.16 about what is only a modification and what is a new instance). This decision has to be done in the scope of the registrant's internal dossier management system, not in the scope of one individual regulatory action: for example, the same PID for a document has to be used across different regulatory actions.

Changes in metadata of concepts do usually not change the PID. For example, the change of the name of a product keeping its formulation must not trigger a new PID assignment.

For some concepts that need be to uniquely identified across registrants and authorities there are centralized repositories for the management of PID. The following table lists the GHSTS concepts using PID and their type of management:

- Local – self-generated PID at the registrant side and stored within the in-house dossier management system for usage in the scope of all regulatory actions
- Central – managed on the global level by the GMB (see chapter 2.4.9), involving registrants and authorities. Details about the centrally managed PID can be found in chapter 4.10.7.

Attention: Central management of persistent identifiers is currently under consideration.

The table also supplies the element names for the concept that serve as user-readable identifier. For centrally managed PID also those user-readable identifiers are managed centrally.

Concept	Management	Element name / description for machine-readable identifier	Element name(s) serving as user-readable identifier
File	Local	FILE_PID (mandatory) Create new value when a new computer file with changed content is stored in the dossier management system of the registrant. No PID must be created for “modifications” (see chapter 3.2.14).	FILENAME (mandatory)
Document	Local	DOCUMENT_PID (mandatory) Create new value when new document is created in the dossier management system of the registrant (see chapter 3.2.13). A new PID is required exactly when the PID of the “Main” content file is new or has changed.	DOCUMENT_NUMBER (0..n)
Document family	Local	DOCUMENT_FAMILY_PID (mandatory) Create new value when initial document of new document family is created in the dossier management system of the registrant.	DOCUMENT_FAMILY (mandatory)
Dossier	Local	DOSSIER_PID (mandatory) Create new value when new regulatory action is created in the dossier management system of the registrant.	DOSSIER_DESCRIPTION_TITLE (mandatory), PROJECT_ID_NUMBER (0..n)

Concept	Management	Element name / description for machine-readable identifier	Element name(s) serving as user-readable identifier
Product	Local	PRODUCT_PID (mandatory) A “product” is a Product or a Technical Grade Active Ingredient (U.S.)/Active Substance (EU). A new Product PID has to be generated when the composition of active ingredients is changed.	GENERIC_PRODUCT_NAME (optional) PRODUCT_NAME (optional), ADMIN_NUMBER (optional)
Legal entity	Central	LEGAENTITY_PID (mandatory) Unique identifiers for registrant and regulatory authority. Use values from centralized repository. If no PID value is available in central repository, then the usage of the “nil PID” is permitted (see chapter 4.10.7).	LEGAENTITY_NAME (mandatory), OTHER_NAME (0..n)
Substance	Central	SUBSTANCE_PID (optional) Unique identifier for pure reference substance. Use value from centralized repository. Currently a similar PID is created by IUCLID using the UUID format (see chapter 4.10.8).	SUBSTANCE_NAME (mandatory), SUBSTANCE_IDENTIFIER (0..n)
ToC	Central	TOC_PID (mandatory) Identification of a standard ToC. The management of standard ToCs is managed centrally. If values are provided the registrant should use them in the regulatory action.	TOC_SHORT_NAME (mandatory), TOC_FULL_NAME (mandatory)
ToC node	Central	TOC_NODE_PID (mandatory) Identification of a node within a standard ToC. The management of standard ToCs is managed centrally. If values are provided the registrant should use them in the regulatory action.	TOC_NODE_NAME (mandatory)

Table 11 – Overview of concepts identified by PID

For the description how a PID is technically represented please see chapter 4.10.7.

3.2.21 Further identifiers on submitter and regulatory authority level

GHSTS supports also the transport of existing identification schemes to ease the integration into existing identification systems. The following table provides an overview:

Concept	Stakeholder	Element name / description
Regulatory Action	Regulatory authority	PROJECT_ID_NUMBER Each participating regulatory authority can assign a Project ID number to a dossier.
Product	Regulatory authority	ADMIN_NUMBER Each participating regulatory authority can assign a regulatory action specific admin number to a product. The type can be specified (BVL Kenn-Nr., EPA Decision No, ...).

Concept	Stakeholder	Element name / description
Document	Regulatory authority	RA_DOCUMENT_NUMBER Each participating regulatory authority can assign a regulatory authority specific Document number to the Document. The type can be specified (MRID, PRMA document number, ...).
Document	Registrant	DOCUMENT_NUMBER The registrant can assign one or many different document numbers and specify their individual type (Report number, Study number, Project number, Plot number, Old report number, other).

Table 12 – Legal entity specific identifier

3.2.22 Impact of document content status on evaluation price

The amount of “New” and “Replaced” documents should be a parameter for the effort (costs) of an evaluation of a submission. Documents with status “No Change”, “Modified” or an RA_DOCUMENT_NUMBER should have no impact on the costs.

Therefore industry should use the content status indicators very carefully, for example not declare every document as “New” when it is not, in order to decrease the work of the Regulatory authorities. Each document that is assigned as “New” has to be registered and research has to be done to confirm the document really is new and it has never been reviewed under a different regulatory action with a RA-specific document number.

3.2.23 GHSTS Picklists

Most metadata information in GHSTS can be provided uncontrolled, but the possible values of some elements are restricted by so-called picklists. The GHSTS submission schema definition uses a mechanism of picklists compatible to mechanism used by the OECD Harmonised Templates (OHT) for Reporting Chemical Test Summaries (see <http://www.oecd.org/ehs/templates>). The content of the picklists is independent of the OHTs.

There are two types of picklists:

- A fixed picklist consists of a static enumeration list of strings.
- An open picklist is an extendable picklist. The registrant may add a new element to the picklist if needed.

Each picklist entry consists of a tuple (coded value, decoded value) that allows the interpretation of the assigned value both for human and machine interpretation.

The following table provides an overview of the elements in GHSTS that are controlled by picklists:

Concept	Element	Name of picklist	Type	Examples
RECEIVER PRODUCT DOCUMENT FILE TOC LEGAL_ENTITY SUBSTANCES	METADATA_STATUS	TYPE_METADATA_STATUS	fixed	Information on whether the metadata for this concept has changed. Values are: <ul style="list-style-type: none"> • New • Modified (at least one metadata item has changed) • No Change

Concept	Element	Name of picklist	Type	Examples
PRODUCT	ADMIN_NUMBER_TYPE	TYPE_ADMIN_NUMBER_TYPE	open	E.g. Application Number, BVL Kenn-Nr., EPA Decision No, Experimental Use Permit, Inert Ingredient clearance request, ...
PRODUCT	FORMULATION_TYPE	TYPE_FORMULATION_TYPE	open	E.g. AE, AL, BB, ...
INGREDIENTS	UNIT	TYPE_UNIT	open	Units for the ingredients
DOSSIER	REGULATORY_TYPE	TYPE_REGULATORY_TYPE	open	E.g. 21 CFR Part 11, 91/414/EEC, 1107/2009/EG, ...
DOSSIER	APPLICATION_TYPE	TYPE_APPLICATION_TYPE	open	E.g. EU OECD dRR, EPA OPPTS, OECD Annex II Joint Review, ...
DOCUMENT	DATA_PROTECTION	TYPE_DATA_PROTECTION	fixed	Yes, No, Open
DOCUMENT	DATA_REQUIREMENT	TYPE_DATA_REQUIREMENT	fixed	Yes, No, Open
DOCUMENT	RA_DOCUMENT_NUMBER_TYPE	TYPE_RA_DOCUMENT_NUMBER_TYPE	open	MRID, PMRA document number, other
DOCUMENT FILE	CONTENT_STATUS	TYPE_CONTENT_STATUS	fixed	New, Retired, Modified, No Change, Replace (see chapter 3.2.17)
DOCUMENT	REFERENCE_TYPE	TYPE_REFERENCE_TYPE	fixed	Semantics of reference E.g. document to replaced document, document to submitted document, translation to main document, ...
DOCUMENT	REFERENCE_TARGET	TYPE_REFERENCE_TARGET	fixed	Target of referenced document submission, dossier, external
DOCUMENT	DOCUMENT_NUMBER_TYPE	TYPE_DOCUMENT_NUMBER_TYPE	open	Study Number, Report Number, Project Number, Plot Number
FILE	CONTENT_STATUS	TYPE_CONTENT_STATUS	fixed	New, Retired, Modified, No Change, Replace (see chapter 3.2.16)
FILE	FILE_TYPE	TYPE_FILE_TYPE	Fixed	Main, Source, Supplemental, XML Summary (see chapter 3.2.12)
TOC	TOC_OWNER	TYPE_TOC_OWNER	open	E.g. OECD, EPA, BVL (Germany), PMRA (Canada), ...
LEGAL_ENTITY	LEGALENTITY_TYPE	TYPE_LEGALENTITY_TYPE	open	E.g. Company, Regulatory authority, Consultant, Test House, Third Party, ...
LEGAL_ENTITY	LEGALENTITY_IDENTIFIER_TYPE	TYPE_LEGALENTITY_IDENTIFIER_TYPE	open	E.g. DUNS-number, SAP, REACH, VAT-number, ...
LEGAL_ENTITY	COUNTRY	TYPE_COUNTRY	open	List of alpha-2 country codes according to ISO 3166-1
SUBSTANCE	SUBSTANCE_IDENTIFIER_TYPE	TYPE_SUBSTANCE_IDENTIFIER_TYPE	open	E.g. CASNO, ECNO, IUBMB, ...

Table 13 – Overview of picklists

The picklists are supplied to the registrant in the form of a separate XSD file called GHSTS picklist XSD. This XSD file should be used to populate the dossier management system of the registrant so that the correct metadata can be assigned.

If a suitable value is missing in the GHSTS picklist XSD for an open picklist the registrant can add new values in the GHSTS XML backbone. The technical mechanism is described in chapter 5.2.4. However this should be an exception mechanism only and the registrant should request an update of the GHSTS picklist XSD. The same applies to regulatory authorities when they notice that items are missing.

Registrants and authorities can request new values to be added to the picklists at the GMB.

The individual values of the picklists are not part of this specification as they possess a different lifecycle than this specification. There is no separate document about the GHSTS picklist XSD; the XSD itself serves as core information. See chapter 2.5 2.4.11 for information on where to find the most current version of the GHSTS picklist XSD.

3.2.24 Information for and targeting at individual receivers

GHSTS allows the addition of metadata for individual regulatory authorities (RA) acting as receivers. The following table summarizes the RA-specific elements and the information that can be delivered with the element and its sub-elements. The column “Man” lists, whether the information is mandatory to be provided for *all* receivers of the submission.

Concept	Element	Man	Description of RA-specific information
PRODUCT	PRODUCT_RA	N	Allows to specify an RA-specific product name and an ADMIN_NUMBER (see table “Table 12 – Legal entity specific identifier”)
DOSSIER	DOSSIER_RA	Y	Specification of the regulatory type, the application type (for both see table “Table 13 – Overview of picklists”) and a RA-specific Project ID number (see Table 12 – Legal entity specific identifier).
DOCUMENT	DOCUMENT_RA	N	If element is present, then this document is targeted at this RA. Specification of additional metadata: data protection, the data requirement (for both see “Table 13 – Overview of picklists”), a document comment, other national guideline number and guideline systems and an RA-specific document number (see table “Table 12 – Legal entity specific identifier”)
FILE	FILE_RA	N	If element is present, then this file is targeted at this RA. Specification of additional metadata: CBI designation (CBI = Confidential Business Information) and a file comment

Table 14 – Overview of RA-specific information

The elements DOCUMENT_RA and FILE_RA are used both for targeting information at RA as well as providing additional RA-specific information. If there is no element DOCUMENT_RA for a specific RA, then this document and all referenced FILE instances are not relevant and can be ignored by this RA. If there is an element DOCUMENT_RA for a specific RA, but for parts of the FILE instances of this document there is no element FILE_RA for the same RA, then those referenced files can be ignored by the RA.

Each RA-specific element has an XML attribute @To_Specific_for_RA_Id to make references to the corresponding attribute “Id” of the receiver element (RECEIVER@Id).

The following figure depicts the references in a graphical manner:

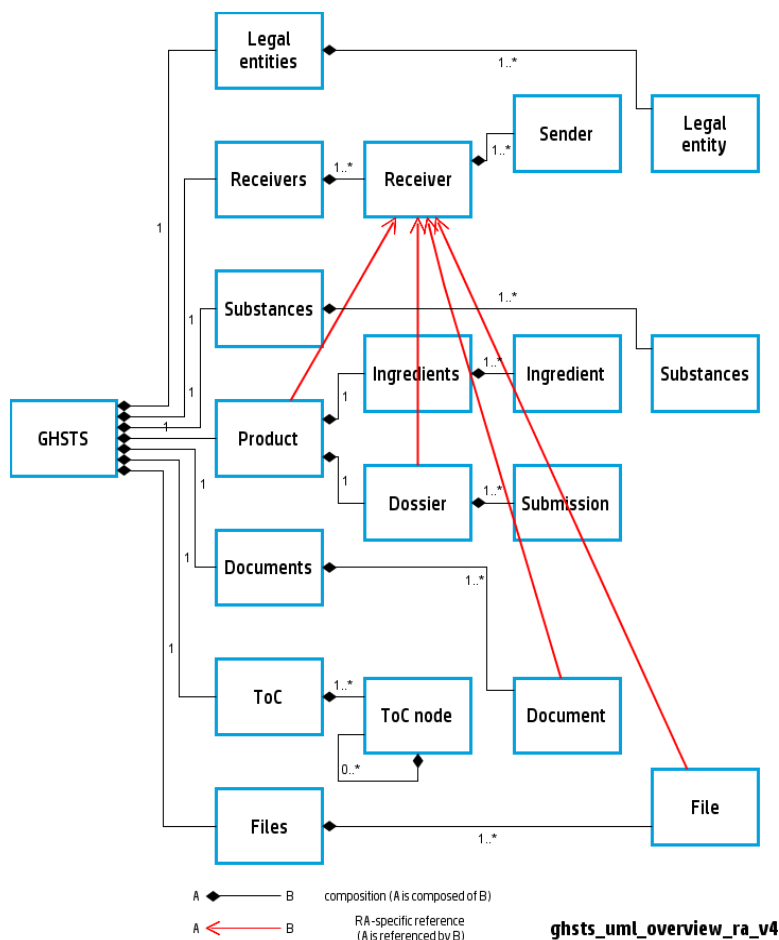


Figure 24 – overview of receiver-specific references

Please note that the information in the RA-elements serves as *additional* RA-specific information for the respective instances of the concept. GHSTS does not support assigning complete instances of concepts (e.g. products, files, documents) in an exclusive manner to specific receivers.

3.2.25 Internal and external references

The following table shows all references that can explicitly be set in the GHSTS backbone. Internal references have a target concept within the same submission package; external references have the target concept outside of the submission package.

Element to represent reference (incl. cardinality)	From concept	To concept	int ext	Description
REFERENCED_DOSSIER_NUMBER (0..n)	Dossier	Dossier	ext	Reference from a current dossier to another referenced dossier. The referenced dossier is identified with its PID or a previous dossier numbering system for example the Caddy Dossier Number (e.g. LANDE005 01.00). A reference reason can be provided. The referenced dossier can be outside of the current regulatory action.

Element to represent reference (incl. cardinality)	From concept	To concept	int ext	Description
DOSSIER_RA (1..n)	Dossier	Receiver	int	Reference from Dossier to Receiver to specify for which RA the additional metadata below DOSSIER_RA is assigned. The target Receiver is identified with unique key in attribute To_Specific_for_RA_Id.
REFERENCED_DOCUMENT (0..n)	Document	Document	int ext	<p>Reference between documents, both to model content changes during document evolution in a document family as well as other types of references.</p> <p>The referenced document can be internal (within the current dossier and its submissions) or external (outside of the current dossier).</p> <p>Note: Internal references to documents only part of preceding submissions can only occur in the rare case of referencing replaced documents, as the XML backbone is “complete” in all other cases. Hence, internal references will most likely reference documents in the current GHSTS XML backbone.</p> <p>The referenced document is identified by its PID, in the case of external references also existing document numbers can be used, if no document PID has been assigned and is known to the receivers.</p> <p>External references can be inside or outside of the current regulatory action.</p>
RELATED_TO_SUBSTANCE (0..n)	Document	Substance	int	Reference from Document to Substance to specify relationship of document with substance. The target Substance is identified with unique key in attribute To_Substance_Id.
DOCUMENT_RA (1..n)	Document	Receiver	int	Reference from Document to Receiver to specify for which RA the document is targeted at and to provide additional metadata. The target Receiver is identified with unique key in attribute To_Specific_for_RA_Id.
REFERENCED_TO_FILE (1..n)	Document	File	int	Reference from document to all file instances of which the document is composed of. File is identified with unique key stored in attribute To_File_Id.
REPLACED_FILE_PID (0..1)	File	File	ext	<p>Reference from file to replaced file in predecessor submission package of current regulatory action.</p> <p>The referenced file is identified by its PID.</p>
FILENAME (1..1)	File	File system	int	Reference from file to file system. The element stores a relative file path (for details see chapter 4.10.1)
FILE_RA (1..n)	File	Receiver	int	Reference from File to Receiver to specify for which RA the file is targeted and to provide additional metadata. The target Receiver is identified with unique key in attribute To_Specific_for_RA_Id.
TOC2DOC (0..n)	ToC Node	Document	int	Reference from ToC node to document. The document is identified with unique key stored in attribute To_Document_Id.
STANDARD_TOC_REFERENCE (0..1)	ToC	Standard ToC	int	Reference to standard ToC XML file with PID and filename (the standard ToC is an optional separate XML file in the submission package)

Element to represent reference (incl. cardinality)	From concept	To concept	int ext	Description
RECEIVER (1..n)	Receiver	Legal Entity	int	Reference from Receiver to Legal Entity (Legal Entity has standard contact information, Receiver adds specific information). Legal Entity is identified via unique key in attribute To_Legal_Entity_Id.
SENDER (1..n)	Sender	Legal Entity	int	Reference from Sender to Legal Entity (Legal Entity has standard contact information, Receiver adds specific information). Legal Entity is identified via unique key in attribute To_Legal_Entity_Id.
PRODUCT_RA (0..n)	Product	Receiver	int	Reference from Product to Receiver to specify for which RA the additional metadata below PRODUCT_RA is assigned. The target Receiver is identified with unique key in attribute To_Specific_for_RA_Id.
INGREDIENT (1..n)	Ingredient	Substance	Int	Reference from Ingredient to Substance to specify for which substance the quantity in the product composition is provided. The target Substance is identified with unique key in attribute To_Substance_Id.

Table 15 – Overview of references

All references that are exclusively internal within the same GHSTS XML backbone file use a unique key to identify the target object using XML methods: The identifier is of type XS:ID and the reference to the identifier of type XS:IDREF. This unique key is deliberately different from the PID concept.

In addition see **Table 11** for a list of PIDs that refer to information stored in centrally managed repositories.

The internal references are depicted in “Figure 4 – Overview of GHSTS ”, the external references are depicted in the following figure:

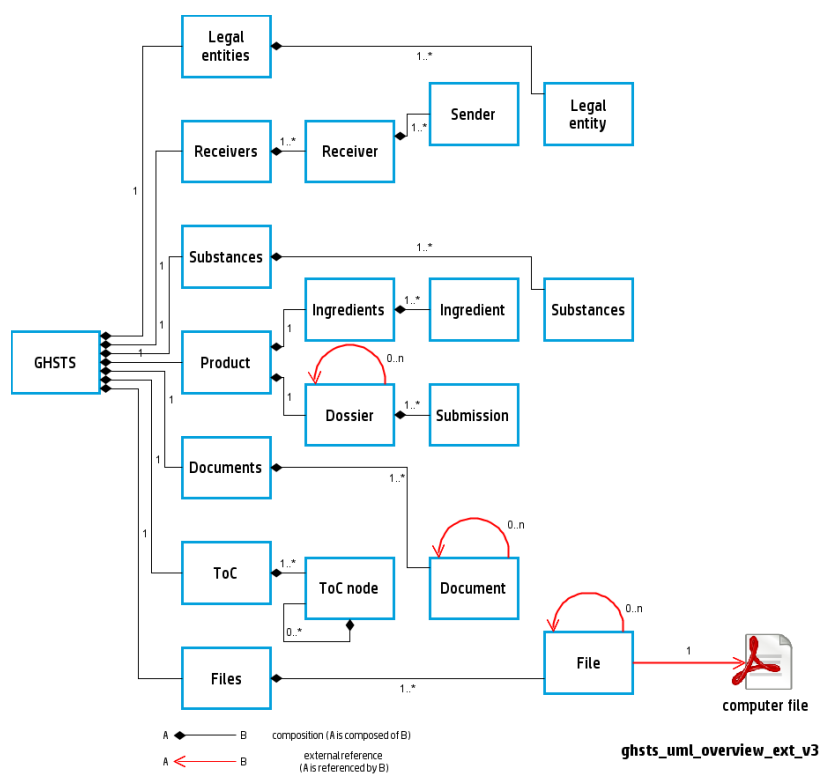


Figure 25 – overview of external references

The figure shows with red arrows all references that can point from the element of one submission to the corresponding element of a predecessor GHSTS submission (for Document and File) or to an arbitrary GHSTS or non-GHSTS submission (for Dossier and Document).

4 GHSTS SUBMISSION SCHEMA DEFINITION

4.1 Overview

This chapter provides a technical reference for the GHSTS XML schema definition, as defined in the GHSTS XSD schema file. This chapter contains the most relevant aspects of the XSD, especially an explanation on how elements should be used in the business context. The complete definition is contained in the XSD itself.

This chapter is intended for

- IT project leads and IT consultants that need to understand how to generate a valid GHSTS XML backbone file. For this audience knowledge of XML, XSD and related technologies (e.g. XPath) is recommended.
- Business domain experts that need to know the semantics of individual elements of the GHSTS metadata. This audience can skip the technical aspects of this chapter and directly address the description of the individual elements.

The XSD for the Table of Contents is described in chapter 5.1. The XSD for the GHSTS picklists is described in chapter 5.2.

4.2 Filename and version

The GHSTS XSD is versioned using a version numbering scheme of type XX.XX.XX. It is composed of a two digit major, a two digit minor and a two digit patch level number separated by single dots. Major versions are used for global functional changes, minor versions for local changes and patch level versions for error corrections. Patch level versions do not require new versions of other GHSTS components (see chapter 2.4).

This version is contained in the version attribute of the root schema element in the XSD file.

The filename uses the same versioning scheme with the hyphen instead of the dot. The file naming convention of the GHSTS XSD is therefore

ghsts_XX-XX-XX.xsd

for example

ghsts_01-00-00.xsd

This chapter describes the version 01.00.00 of the GHSTS XML schema, published on **XXXX**.

4.3 Scope

The GHSTS XML schema definition describes the structure of the XML backbone file of *one* single submission that is part of a regulatory action. The XML backbone file has the following references to the outside world:

- For internal and external references from the XML backbone see chapter 3.2.24.
- For the folder structure of the submission package of which the XML backbone is part of, and the folder structure of the parent regulatory action see chapter 6.

4.4 Namespace

The GHSTS submission schema uses the following namespace definitions:

- `xmlns:xs="http://www.w3.org/2001/XMLSchema"`
- `xmlns="http://www.oecd.org/GHSTS"`
- `targetNamespace="http://www.oecd.org/GHSTS"`

4.5 Unique keys

As described in chapter 3.2.25, the internal references within a submission are made using unique keys to identify the target object. The unique key of the target object is represented as attribute "Id" and is of type XS:ID, the reference to the target object is represented as attribute "To_<xxx>" with xxx being a variable string. The reference is of type XS:IDREF. Please note that values of the attribute "Id" must be valid XML names, which means that they must start with a character. For details please consult the definition of the XML type XS:ID.

The following table lists the XPath expression of the attributes with Id values as well as the XPath expression of the attributes that reference this unique key:

XPath of element with Id attribute	Xpath of element attribute with reference to element in first column
GHSTS/LEGAL_ENTITIES/LEGAL_ENTITY/@Id	GHSTS/PRODUCT/DOSSIER/DOSSIER_RA/@To_Specific_for_RA_Id GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_RA/ @To_Specific_for_RA_Id GHSTS/FILES/FILE/FILE_RA/@To_Specific_for_RA_Id GHSTS/PRODUCT/PRODUCT_RA/@To_Specific_for_RA_Id GHSTS/RECEIVERS/RECEIVER/@To_Legal_Entity_Id GHSTS/RECEIVERS/RECEIVER/SENDER/@To_Legal_Entity_Id
GHSTS/SUBSTANCES/SUBSTANCE/@Id	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/RELATED_TO_SUBSTANCE/@To_Substance_Id GHSTS/PRODUCT/INGREDIENTS/INGREDIENT/@To_Substance_Id
GHSTS/FILES/FILE/@Id	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/REFERENCED_TO_FILE/@To_File_Id
GHSTS/DOCUMENTS/DOCUMENT/@Id	GHSTS/TOC/STRUCTURE/TOC_NODE/TOC2DOC/@To_Document_Id

Table 16 – internal references with unique keys

The value for the unique key can be freely assigned by the registrant during the publication process. The only constraints are the uniqueness for each concept and for each submission package. The GHSTS XSD contains checks to validate the uniqueness constraint and the referential integrity of internal references by checking the existence of the target element with the unique key.

4.6 Combination of identifier and identifier type

The elements ADMIN_NUMBER, DOCUMENT_NUMBER, LEGALENTITY_IDENTIFIER, RA_DOCUMENT_NUMBER and SUBSTANCE_IDENTIFIER possess each a combination of identifier and identifier type. The element IDENTIFIER contains the actual value, whereas the type of identification is provided using a picklist.

The following picture shows this structure for the element DOCUMENT_NUMBER:

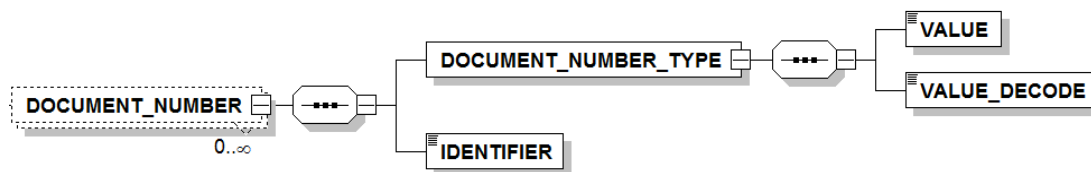


Figure 26 – Identifier and identifier type

For example, a DOCUMENT_NUMBER (e.g. X120976) will be submitted using the IDENTIFIER element in combination with the corresponding DOCUMENT_NUMBER_TYPE, which characterizes the type of the number using a picklist (e.g. “Report Number”). Like all picklists each picklist item contains a value and a decoded value. If an IDENTIFIER is supplied then the corresponding type element has to be supplied as well.

This structure is the same for the other elements. The name of the element for the type is derived from the root element name:

Root element	Name of type element
ADMIN_NUMBER	ADMIN_NUMBER_TYPE
DOCUMENT_NUMBER	DOCUMENT_NUMBER_TYPE
LEGAENTITY_IDENTIFIER	LEGAENTITY_IDENTIFIER_TYPE
RA_DOCUMENT_NUMBER	RA_DOCUMENT_NUMBER_TYPE
SUBSTANCE_IDENTIFIER	SUBSTANCE_IDENTIFIER_TYPE

Table 17

Using this technique a new picklist item can be added into the GHSTS XML backbone file without any changes to the GHSTS submission schema.

4.7 Notation conventions

This chapter provides a brief overview about the notation conventions that are used to explain the schema definition as well as conventions used for the creation of the schema. It serves to support readers from the business audience to understand the technical core characteristics of the schema.

4.7.1 Elements

The element names of the Schema Description File are uppercase strings. If an element name consists of multiple words, they are concatenated by an underscore.

```
<xs:element name="DOSSIER_DESCRIPTION_TITLE">
```

The principle of the GHSTS submission schema is to transport element values in the GHSTS XML File as text content of the elements.

```
<DOSSIER_DESCRIPTION_TITLE>The text content</DOSSIER_DESCRIPTION_TITLE>
```

4.7.2 Attributes

The Attribute names are mixed uppercase and lower case strings e.g. “To_Legal_Entity_Id“. Almost all attribute values are used for technical aspects in a GHSTS XML file to characterize nodes with unique identifiers or to make references to these identifiers.

```
...<FILE Id="The attribute value"> ...</FILE>
```

4.7.3 Occurrence indicators

The minOccurs / maxOccurs indicator specifies the minimum / maximum number of times an element can occur. For mandatory values the value for minOccurs must be always one, as here for the definition of the number of files that a document must possess, it is a mandatory multi-value element:

```
<xs:element name="REFERENCED_TO_FILE" minOccurs="1" maxOccurs="unbounded">
```

The following table shows how the occurrence indicators are displayed in the graphical schema representation, using elements from the DOCUMENT concept.

Occurrence	Occurrence indicators	Graphical representation
mandatory / single value (1/1)	minOccurs="1" maxOccurs="1"	
optional / single value (0/1)	minOccurs="0" maxOccurs="1"	
mandatory / multi-value (1/*)	minOccurs="1" maxOccurs="unbounded"	
optional / multi-value (0/*)	minOccurs="0" maxOccurs="unbounded"	

Table 18 – Notation and graphical representation of occurrence indicators

Please note that if not stated otherwise, the order of multiple instances within the XML file is of no business significance. Two apparent exceptions are the order of ToC nodes and the order of document references in a ToC node.

4.8 Boilerplate for element description

This chapter contains the boilerplate for the subsequent documentation of the XSD elements including an explanation of the description template.

Please note that this specification contains only main aspects of the actual XSD definition. The full XSD definition can be found in the XSD file itself

4.8.1 *Child elements*

This chapter defines the child elements of the element in focus in the form of a table:

Element name				
Xpath				
Diagram				
Child element	T	Occ	Definition	Description

Table 19 – boilerplate sample: child elements definition

- The row Xpath contains the Xpath expression of the element in an instance.
- The cell next to “Diagram” shows a graphical representation of the element and its direct child elements.
- The rows below “child elements” contain the child elements of the element in focus. The order of child elements in the parent element is of no significance.
- The column “T” stands for “Type”. It contains the character “T” if the child element contains text content, and the character “E” if the element itself contains element content.
- The column “Occ” contains the occurrence indicator as explained in chapter 4.7.3, e.g. “(0/1)”.
- The column “Definition” contains the definition of the element including a reference where the definition can be found (except for standard XML schema “xs” types, see chapter 4.10.10).
- For GHSTS-specific types the column “definition” contains a reference to the chapter with the definition of the element.

4.8.2 *Attributes*

For elements with attributes, this chapter lists the attributes of the element:

Attribute name	Man	Type	Definition	Description

Table 20 - boilerplate sample: attributes definition

- The column “Man” specifies whether the attribute is mandatory or not (Y = Yes, N = No).

The chapter is omitted when an element does not contain any attributes.

4.8.3 *Constraints*

The chapter “Constraints” includes additional constraints and guidance on how to use the element and its child elements and attributes.

Constraint Name	Description

Table 21 - boilerplate sample: constraints definition

This chapter may serve as input for the specification of the GHSTS validator. Please note that the GHSTS validator may perform additional checks on the schema not described here.

The chapter is omitted when an element does not contain any additional constraints.

4.8.4 Example

This chapter contains a short XML example of an element instance and its direct child elements, omitting the further nested XML elements.

Attention: This chapter does not appear for all element instances and child elements.

4.9 Elements

4.9.1 Overview

An overview diagram of the complete XSD is provided as separate image file in PNG format, with the filename

ghsts_XX-XX-XX.png

with XX-XX-XX being the version string of the current GHSTS specification version (see chapter 1.4).

4.9.2 GHSTS

The GHSTS element is the root element of the complete XML backbone file.

4.9.2.1 Child elements

Element name	GHSTS			
Xpath	GHSTS			
Diagram				
Child element	T	Occ	Definition	Description
RECEIVERS	E	(1,1)	chapter 4.9.3	List of receivers
PRODUCT	E	(1,1)	chapter 4.9.6	Product
DOCUMENTS	E	(1,1)	chapter 4.9.14	List of documents
FILES	E	(1,1)	chapter 4.9.25	List of files
TOC	E	(1,1)	chapter 0	ToC
LEGAL_ENTITIES	E	(1,1)	chapter 4.9.32	List of legal entities
SUBSTANCES	E	(1,1)	chapter 4.9.36	List of substances
USED_TEMPLATES	T	(0,1)	chapter 4.9.38	Placeholder for future use

Table 22 –child elements definition for element GHSTS

4.9.2.2 Attributes

Attribute name	Man	Type	Definition	Description
specificationversion	Y	versionnumberType	chapter 4.10.5	Version of the GHSTS specification (this document) to which this submission adheres. Please note that the version of the GHSTS specification is not identical to the version of the GHSTS submission XSD. The GHSTS XSD is references in the standard attribute schemaLocation (see example below)
xsi:schemaLocation	Y		chapter 8.2 XML schema	reference to the GHSTS submission XSD file, see example below. The location of the XSD and the path has to be compliant with the folder structure, see chapter 6.1.

xmlns	Y		chapter 8.2 XML schema	definition of the standard GHSTS namespace, see example below. Do not change this value.
xmlns:xsi	Y		chapter 8.2 XML schema	definition of the standard XSI namespace (XML schema instance). Do not change this value.

Table 23 - attributes definition for element GHSTS**4.9.2.3 Example**

```
<GHSTS specificationversion="00.00.00" xsi:schemaLocation="http://www.oecd.org/GHSTS
../02/utls/ghsts_01-00-00.xsd" xmlns="http://www.oecd.org/GHSTS"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
```

4.9.3 RECEIVERS

The RECEIVERS element is the definition block for all Regulatory authorities who will be a receiver of the GHSTS Submission Package and the relationship to the corresponding SENDER (Submitting Companies). This multi valued relationship is needed for Joint Review.

4.9.3.1 Child elements

Element name	RECEIVERS			
Xpath	GHSTS/RECEIVERS			
Diagram				
Child element	T	Occ	Definition	Description
RECEIVER	E	(1,*)	chapter 4.9.4	Single receiver.

Table 24 –child elements definition for element RECEIVERS**4.9.4 RECEIVER**

The element describes a single receiver of a submission. The main information about the receiver is provided by the element LEGAL_ENTITY which is referenced in an attribute.

For the explanation of the logical concept please see chapter 3.2.8.

4.9.4.1 Child elements

Element name	RECEIVER			
Xpath	GHSTS/RECEIVERS/RECEIVER			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	The value of the element METADATA_STATUS characterizes the status of all metadata contained within the block RECEIVERS and its direct and indirect children elements (including SENDER and DOSSIER!) see chapter 3.2.3
SHORT_NAME	T	(1,1)	string20, chapter 4.10.11	Typical abbreviation of the receiver, e.g. EPA, BVL. Used for display in the GHSTS browser.
ROLE	T	(0,1)	string255, chapter 4.10.11	Each receiver is characterized by his role in the work sharing process. Example: "RMS"
SENDER	E	(1,*)	chapter 4.9.5	For each receiver multiple senders can be supplied.

Table 25 –child elements definition for element REICEIVER

4.9.4.2 Attributes

Attribute name	Man	Type	Definition	Description
Id	Y	XS:ID	chapter 4.5	unique ID, which is used for reference in other elements that are specific for a RA (attribute To_Specific_for_RA_Id)
To_Legal_Entity_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance supplying the details for the receiver. Reference validated in schema using keyref RefIntegrityReceiverToLegalentity

Table 26 - attributes definition for element RECEIVER

4.9.5 SENDER

The element SENDER describes a sender of a submission. Please note that this information is RA-specific: For each RECEIVER different and eventually multiple senders can be submitted.

The main information about the sender is provided by the element LEGAL_ENTITY which is referenced in an attribute. For the explanation of the logical concept please see chapter 3.2.8.

4.9.5.1 Child elements


Element name	SENDER			
Xpath	GHSTS/RECEIVERS/RECEIVER/SENDER			
Diagram				
Child element	T	Occ	Definition	Description
COMPANY_CONTACT_REGULATORY_ROLE	T	(0,1)	string255, chapter 4.10.11	Each sender is characterized by his role in the work sharing process. In case of “joint submissions” there should be defined one company with the role “Lead submitting company”. That means that all additional submissions were built by this company.
REMARK	T	(0,1)	string2000, chapter 4.10.11	Area to communicate additional company contact details.

Table 27 –child elements definition for element SENDER

4.9.5.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Legal_Entity_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance supplying the details for the sender. Reference validated in schema using keyref RefIntegritySenderToLegalentity

Table 28 - attributes definition for element SENDER

4.9.6 PRODUCT

The element PRODUCT represents the product that is the subject of the complete regulatory action. It contains the DOSSIER and SUBMISSION as direct and indirect child elements, because both are dependent on the PRODUCT.

There is exactly one product per regulatory action.

For the explanation of the logical concept please see chapter 3.2.9.

4.9.6.1 Child elements

Element name	PRODUCT			
Xpath	GHSTS/PRODUCT			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	The value of the element METADATA_STATUS characterizes the status of all metadata containing within the block PRODUCT and its direct and indirect children elements see chapter 3.2.3
PRODUCT_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the product, see chapter 3.2.20.
GENERIC_PRODUCT_NAME	T	(1,1)	string255, chapter 4.10.11	generic name (common name or active ingredient)
PRODUCT_RA	E	(0,*)	chapter 4.9.7	optional element to supply additional metadata for each receiver.
FORMULATION_TYPE	E	(0,1)	chapter 3.2.23, 5.2	open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_FORMULATION_TYPE Two character code describing the formulation of the product following GIFAP rules, according to "Catalogue of Pesticide Formulation types and International Coding Systems " GIFAP Technical Monograph No. 2 4th Edition 1999 Examples: "EC", "WP", "GR"
INGREDIENTS	E	(1,1)	chapter 4.9.8	list of ingredients
DOSSIER	E	(1,1)	chapter 4.9.10	dossier element as representation of the metadata for the complete regulatory action = sequence of GHSTS submissions

Table 29 –child elements definition for element PRODUCT

4.9.7 **PRODUCT_RA**

For the explanation of the logical concept please see chapter 3.2.24.

4.9.7.1 *Child elements*

Element name	PRODUCT_RA			
Xpath	GHSTS/PRODUCT/PRODUCT_RA			
Diagram				
Child element	T	Occ	Definition	Description
PRODUCT_NAME	T	(0,1)	string2000, chapter 4.10.11	RA-specific product name associated with the dossier
ADMIN_NUMBER	E	(0,*)	chapter 4.6	<p>identifier and open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_ADMIN_NUMBER_TYPE. product references to the numbering world of a Regulatory authority.</p> <p>Examples: “BVL Kenn-Nr.”, “EPA Decision No”, “PMRA Decision No”, “PRIA Pre-application”.</p>

Table 30 –child elements definition for element PRODUCT_RA

4.9.7.2 *Attributes*

Attribute name	Man	Type	Definition	Description
To_Specific_for_RA_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityProduct_RAToLegalentity

Table 31 - attributes definition for element PRODUCT_RA

4.9.7.3 *Constraints*

Constraint Name	Description
	Elements PRODUCT_NAME and ADMIN_NUMBER must not both be missing, at least one element has to be present.

Table 32 - constraints definition for element PRODUCT_RA

4.9.8 **INGREDIENTS**

This element contains list of active ingredients and their content in the product.

4.9.8.1 Child elements

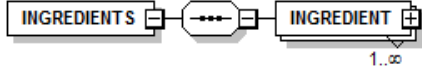
Element name	INGREDIENTS			
Xpath	GHSTS/PRODUCT/INGREDIENTS			
Diagram				
Child element	T	Occ	Definition	Description
INGREDIENT	E	(1,*)	chapter 4.9.9	single ingredient

Table 33 –child elements definition for element INGREDIENTS

4.9.9 INGREDIENT

The ingredient in the GHSTS submission schema is the definition of the quantity of the substance in the product. The substance specific metadata are submitted in the element SUBSTANCE, which is referred by attribute To_Substance_Id.

For the explanation of the logical concept please see chapter 3.2.9.

4.9.9.1 Child elements

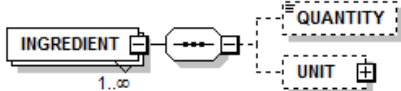
Element name	INGREDIENT			
Xpath	GHSTS/PRODUCT/INGREDIENTS/INGREDIENT			
Diagram				
Child element	T	Occ	Definition	Description
QUANTITY	T	(0,1)	xs:decimal	Quantity, specified as decimal Example: “42.729”
UNIT	E	(0,1)	chapter 3.2.23, 5.2	open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_UNIT Example: “g/kg”

Table 34 –child elements definition for element INGREDIENT

4.9.9.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Substance_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityIngredientToSubstance

Table 35 - attributes definition for element INGREDIENT

4.9.9.3 Constraints

Constraint Name	Description
	For TGAI / substance dossiers both elements can be empty. For formulated product dossiers both values have to be supplied.

Table 36 - constraints definition for element INGREDIENT

4.9.10 DOSSIER

Representation of the metadata for the dossier for one complete regulatory action as ordered sequence of GHSTS submission.

For the explanation of the logical concept please see chapter 3.2.6.

4.9.10.1 Child elements

Element name	DOSSIER			
Xpath	GHSTS/PRODUCT/DOSSIER			
Diagram				
Child element	T	Occ	Definition	Description
DOSSIER_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the dossier, see chapter 3.2.20.
DOSSIER_DESCRIPTION_TITLE	T	(1,1)	string2000, chapter 4.10.11	human-readable identifier of the dossier designated by the registrant. Examples: "OECD Annex II Dossier Spiroamine", "ACME Natural Insecticide"
DOSSIER_COMP_ID	T	(1,1)	string20, chapter 4.10.11	Company internal short code for the current dossier. This value will be used as folder name for the regulatory action folder, see chapter 6.1. Example: "SYNCH127"
REFERENCED_DOSSIER	E	(0,*)	chapter 4.9.11	possibility to reference other dossiers

DOSSIER_RA	E	(1,*)	chapter 4.9.12	mandatory element to specify additional metadata for each receiver
SUBMISSION	E	(1,*)	chapter 4.9.13	Each GHSTS XML backbone file contains the SUBMISSION element for the current and all – if any – preceding submissions. Multiple instances are ordered by ascending submission number in the order of their appearance in the XML file.

Table 37 –child elements definition for element DOSSIER

4.9.11 REFERENCED_DOSSIER

It is possible to make a content reference into one or more other dossiers with the help of the element REFERENCED_DOSSIER_NUMBER, which was already submitted, for example

- under another "Regulatory Action" or
- from a Product dossier to an Active Ingredient dossier

4.9.11.1 Child elements

Element name	REFERENCED_DOSSIER			
Xpath	GHSTS/PRODUCT/DOSSIER/REFERENCED_DOSSIER			
Diagram				
Child element	T	Occ	Definition	Description
REFERENCED_DOSSIER_NUMBER	T	(0,1)	string255, chapter 4.10.11	The referenced dossier is identified with its PID (preferred) or a previous dossier numbering system for example the Caddy Dossier Number (e.g. LANDE005 01.00)
REFERENCED_DOSSIER_REASON	T	(0,1)	string2000, chapter 4.10.11	human-readable reason why this reference is supplied in this submission

Table 38 –child elements definition for element REFERENCED_DOSSIER

4.9.12 DOSSIER_RA

For the explanation of the logical concept please see chapter 3.2.24.

4.9.12.1 Child elements

Element name	DOSSIER_RA			
Xpath	GHSTS/PRODUCT/DOSSIER/DOSSIER_RA			
Diagram				
Child element	T	Occ	Definition	Description
REGULATORY_TYPE	E	(1,1)	chapter 3.2.23, 5.2	open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_REGULATORY_TYPE Legislation, act, regulation or directive under which the submission is made. Examples: “91/414/EWG”, 1107/2009/EG,98/8/EG”
APPLICATION_TYPE	E	(1,1)	chapter 3.2.23, 5.2	open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_APPLICATION_TYPE Example: “6(a)(2) submission”, “OECD Annex II Joint Review”
PROJECT_ID_NUMBER	T	(0,*)	string255, chapter 4.10.11	The PROJECT_ID_NUMBER(s) is / are assigned by the Regulatory authority to the dossier. Example: “2011-1234”

Table 39 –child elements definition for element DOSSIER_RA

4.9.12.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Specific_for_RA_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityDossier_RAToLegalentity

Table 40 - attributes definition for element DOSSIER_RA

4.9.12.3 Constraints

Constraint Name	Description
	The element has to be supplied for all receiver of the submission

Table 41 - constraints definition for element DOSSIER_RA

4.9.13 SUBMISSION

For the explanation of the logical concept please see chapter 3.2.5.

4.9.13.1 Child elements

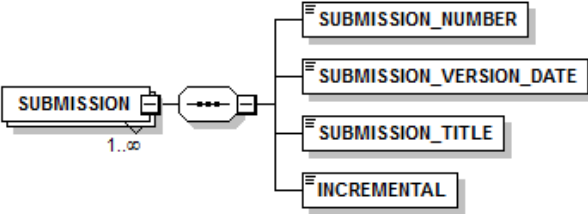
Element name	SUBMISSION			
Xpath	GHSTS/PRODUCT/DOSSIER/SUBMISSION			
Diagram				
Child element	T	Occ	Definition	Description
SUBMISSION_NUMBER	T	(1,1)	submissionnumberType	sequence of the submission represented as two digit number padded with zero, e.g. 01, 02, 03, ...
SUBMISSION_VERSION_DATE	T	(1,1)	xs:date	Date of the submission, in the form YYYY-MM-DD Example: "2009-04-07"
SUBMISSION_TITLE	T	(1,1)	string2000, chapter 4.10.11	Title of the submission Example: "Spiroxamine OECD Annex II Update from March 2009"
INCREMENTAL	T	(1,1)	xs:boolean	Was this submission of type incremental or not (see chapter 3.2.19)

Table 42 –child elements definition for element SUBMISSION

4.9.13.2 Constraints

Constraint Name	Description
	The first submission has have the submission number 01 and has to be of type INCREMENTAL=FALSE

Table 43 - constraints definition for element SUBMISSION

4.9.14 DOCUMENTS

The element DOCUMENTS contains a list of documents.

4.9.14.1 Child elements

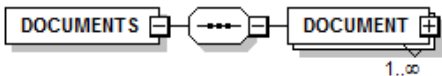
Element name	DOCUMENTS			
Xpath	GHSTS/DOCUMENTS			
Diagram				
Child element	T	Occ	Definition	Description
DOCUMENT	E	(1,*)	chapter 4.9.15	element representing one single document

Table 44 –child elements definition for element DOCUMENTS

4.9.15 DOCUMENT

The DOCUMENT metadata were grouped in a generic branch (DOCUMENT_GENERIC) common for all RA and a branch of the RA-specific document metadata (DOCUMENT_RA). Both branches have an own element METADATA_STATUS to characterize the status of the metadata independently.

For the explanation of the logical concept please see chapter 3.2.12.

4.9.15.1 Child elements

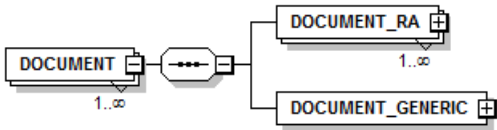
Element name	DOCUMENT			
Xpath	GHSTS/DOCUMENTS/DOCUMENT			
Diagram				
Child element	T	Occ	Definition	Description
DOCUMENT_RA	E	(1,*)	chapter	There is one element for each RA at which the document is targeted.
DOCUMENT_GENERIC	E	(1,1)	chapter	Document metadata that is not dependent on RA.

Table 45 –child elements definition for element DOCUMENT

4.9.15.2 Attributes

Attribute name	Man	Type	Definition	Description
Id	Y	XS:ID	chapter 4.5	unique ID, which is used for reference in other elements

Table 46 - attributes definition for element DOCUMENT

4.9.15.3 Constraints

Constraint Name	Description
	Each DOCUMENT instance must reference at least one ToC node via its ID, using the element TOC2DOC (see chapter 4.10.3.2).

Table 47 - constraints definition for element DOCUMENT

4.9.16 DOCUMENT_RA

The element DOCUMENT_RA is present for each RA to which the registrant wants to assign the document. In addition the element contains RA specific information that may differ between the different RA receiving the same document. The information that is invariant for all RA is represented below the element DOCUMENT_GENERIC (see chapter 4.9.20).

For the explanation of the logical concept please see chapter 3.2.24.

4.9.16.1 Child elements

Element name	DOCUMENT_RA			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_RA			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	Metadata status of all child elements of DOCUMENT_RA
DATA_PROTECTION	E	(1,1)	chapter 3.2.23, 5.2	fixed picklist, defined in GHSTS picklist XSD in type TYPE_DATA_PROTECTION indicates a data protection claim for required or not required documentation Values are Y,N,O (O=Open)
DATA_REQUIREMENT	E	(1,1)	chapter 3.2.23, 5.2	fixed picklist, defined in GHSTS picklist XSD in type TYPE_DATA_REQUIREMENT Values are Y,N,O (O=Open)

DOCUMENT_COMMENT	T	(0,*)	string2000, chapter 4.10.11	regulatory authority specific document comments Example: “Study number: T0035534, Report number 20416”
OTHER_NATIONAL_GUIDELINE	E	(0,*)	chapter 4.9.17	additional Regulatory authority specific references into other national guidelines
RA_DOCUMENT_NUMBER	E	(0,1)	chapter 4.9.18	element to submit document identifiers already assigned by the RA

Table 48 –child elements definition for element DOCUMENT_RA

The correct usage and the correct interpretation of the elements DATA_PROTECTION and DATA_REQUIREMENT depend from the Regulatory authority and the specific regulatory action.

4.9.16.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Specific_for_RA_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityDocument_RA_ToLegalentity

Table 49 - attributes definition for element DOCUMENT_RA

4.9.16.3 Constraints

Constraint Name	Description
	The element has to be supplied for at least one receiver of the submission.
	For every RA referenced by the DOCUMENT_RA there must be at least one referenced FILE instance of this document with an element FILE_RA, where this same RA appears. (Rationale: A document without content for a RA does not make sense).
	Only those RA, that appear (as reference) in DOCUMENT_RA can be used in the FILE_RA elements of the referenced FILE instances of this document. (Rationale: The FILE_RA can only assign content to RA, to which the parent document is targeted.)

Table 50 - constraints definition for element DOCUMENT_RA

4.9.17 OTHER_NATIONAL_GUIDELINE

With the help of the non-mandatory element OTHER_NATIONAL_GUIDELINE and its child elements GUIDELINE_SYSTEM and GUIDELINE_NUMBER it is possible to make additional Regulatory authority specific references into other national guidelines.

4.9.17.1 Child elements

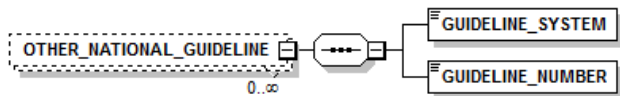
Element name	OTHER_NATIONAL_GUIDELINE			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/ DOCUMENT_RA/OTHER_NATIONAL_GUIDELINE			
Diagram				
Child element	T	Occ	Definition	Description
GUIDELINE_SYSTEM	T	(1,1)	string20, chapter 4.10.11	guideline system e.g. DACO, OPPTS
GUIDELINE_NUMBER	T	(1,1)	string20, chapter 4.10.11	identifier in the specified guideline system

Table 51 –child elements definition for element OTHER_NATIONAL_GUIDELINE

4.9.18 RA_DOCUMENT_NUMBER

4.9.18.1 Child elements

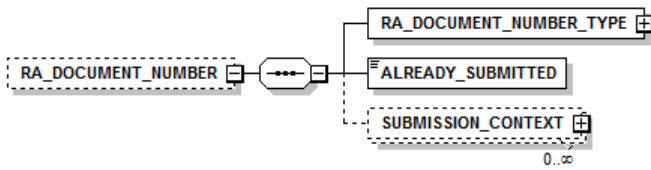
Element name	RA_DOCUMENT_NUMBER			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/ DOCUMENT_RA/RA_DOCUMENT_NUMBER			
Diagram				
Child element	T	Occ	Definition	Description
RA_DOCUMENT_NUMBER_TYPE	E	(1,1)	chapter 4.6	<p>identifier and open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_RA_DOCUMENT_NUMBER_TYPE</p> <p>document identification number assigned by regulatory authority, e.g. the MRID. This element can also be used to identify documents that have been submitted before the usage of GHSTS.</p>
ALREADY_SUBMITTED	T	(1,1)	xs:boolean	true, if this document has already been submitted in one or many different regulatory actions. Details can be provided in the element SUBMISSION_CONTEXT.
SUBMISSION_CONTEXT	E	(0,*)	chapter 4.9.19	element to provide details about previous submission context(s) of this document

Table 52 –child elements definition for element RA_DOCUMENT_NUMBER

4.9.19 SUBMISSION_CONTEXT

This element allows the provision of information about a previous submission of a document, either by a dossier PID for GHSTS dossiers or by textual information for legacy dossiers of other formats, for example the legacy dossier number.

4.9.19.1 Child elements

Element name	SUBMISSION_CONTEXT			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/ DOCUMENT_RA/RA_DOCUMENT_NUMBER/SUBMISSION_CONTEXT			
Diagram				
Child element	T	Occ	Definition	Description
DOSSIER_PID (choice)	T	(1,1)	PIDType, chapter 4.10.7	PID of the dossier (for GHSTS dossiers)
DOSSIER_NUMBER (choice)	T	(1,1)	string255, chapter 4.10.11	dossier number of the dossier (for legacy dossiers)

Table 53 –child elements definition for element SUBMISSION_CONTEXT

4.9.20 DOCUMENT_GENERIC

This element groups the information about a document that is identical for all RA.

4.9.20.1 Child elements

Element name	DOCUMENT_GENERIC
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC
Diagram	<p>The diagram illustrates the structure of the DOCUMENT_GENERIC element. It is a complex type with a sequence of child elements. The children are:</p> <ul style="list-style-type: none"> METADATA_STATUS (required, 1) DOCUMENT_PID (required, 1) DOCUMENT_FAMILY_PID (required, 1) DOCUMENT_FAMILY (required, 1) CONTENT_STATUS_HISTORY (required, 1) REFERENCED_DOCUMENT (optional, 0..∞) RELATED_TO_SUBSTANCE (optional, 0..∞) DOCUMENT_NUMBER (optional, 0..∞) DOCUMENT_TITLE (required, 1) DOCUMENT_AUTHOR (required, 1) DOCUMENT_ISSUE_DATE (required, 1) DOCUMENT_OWNER (optional, 1..∞) PUBLISHED_INDICATOR (required, 1) A choice between COMPLETE_DOCUMENT_SOURCE (required, 1) and an unnamed group (optional, 0..1). <ul style="list-style-type: none"> The unnamed group contains a choice between DOCUMENT_SOURCE (required, 1) and an unnamed group (optional, 0..1). The unnamed group contains a choice between DOCUMENT_YEAR (optional, 1), DOCUMENT_ISSUE (optional, 1), DOCUMENT_VOLUME (optional, 1), and DOCUMENT_PAGES (optional, 1). TEST_LABORATORY (optional, 0..∞) GXP_INDICATOR (required, 1) TESTED_ON_VERTTEBRATE (required, 1) REFERENCED_TO_FILE (optional, 1..∞)

Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	Metadata status of all child elements of DOCUMENT_GENERIC excluding the CONTENT_STATUS_HISTORY (as this element changes with every submission) and REFERENCED_TO_FILE (as references change with full submissions) see chapter 3.2.3
DOCUMENT_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the document, see chapter 3.2.20.
DOCUMENT_FAMILY_PID	T	(1,1)	PIDType, chapter 4.10.7	A machine readable filter to select documents which belongs to one "document family", see chapter 3.2.20.
DOCUMENT_FAMILY	T	(1,1)	string255, chapter 4.10.11	A simple filter for documents which belongs to one "document family"
CONTENT_STATUS_HISTORY	E	(1,1)	chapter 4.10.2	content status history of the document, see chapter 3.2.17
REFERENCED_DOCUMENT	E	(0,*)	chapter 4.9.22	references to other documents
RELATED_TO_SUBSTANCE	E	(0,*)	chapter 4.9.23	references to substances
DOCUMENT_NUMBER	E	(0,*)	chapter 3.2.23, 5.2	open picklist defined in GHSTS picklist XSD in type EXTENSION_TYPE_DOCUMENT_NUMBER_TYPE For usage see comment below this table.
DOCUMENT_TITLE	T	(1,1)	string2000, chapter 4.10.11	Full title of the document provided by the registrant. This value will be used for reference lists.
DOCUMENT_AUTHOR	T	(1,1)	string255, chapter 4.10.11	For usage see comment below this table
DOCUMENT_ISSUE_DATE	T	(1,1)	xs:date	Date of the document
DOCUMENT_OWNER	T	(1,*)	string2000, chapter 4.10.11	Owner of the document at the time of the current submission. If more than one owner exists, repeat the element.
PUBLISHED_INDICATOR	T	(1,1)	xs:boolean	Indicates whether a document has been published and therefore made publicly available. For usage see comment below this table
COMPLETE_DOCUMENT_SOURCE (choice option 1)	T	(1,1)	string2000, chapter 4.10.11	Bibliographic citation source For usage see comment below this table

DOCUMENT_SOURCE (choice option 2)	T	(1,1)	string2000, chapter 4.10.11	Bibliographic citation source For usage see comment below this table
DOCUMENT_YEAR (choice option 2)	T	(0,1)	xs:positiveInteger	For usage see comment below this table
DOCUMENT_ISSUE (choice option 2)	T	(0,1)	string20, chapter 4.10.11	For usage see comment below this table
DOCUMENT_VOLUME (choice option 2)	T	(0,1)	string20, chapter 4.10.11	For usage see comment below this table
DOCUMENT_PAGES (choice option 2)	T	(0,1)	string20, chapter 4.10.11	For usage see comment below this table
TEST_LABORATORY	T	(0,*)	string2000, chapter 4.10.11	Name and address of the test facility / facilities where the study was carried out. Example: "Bayer AG, Wuppertal, Germany"
GPX_INDICATOR	T	(1,1)	xs:boolean	Indicates on the document level whether the test facility was carried out under accredited Good Practices. GXP was used as a general term for GLP, GAP, GBP.
TESTED_ON_VERTEBRATE	T	(1,1)	xs:boolean	Indicates whether a study has been conducted on a vertebrate
REFERENCED_TO_FILE	E	(1,*)	chapter 4.9.24	Connects the DOCUMENT level with the FILE level. It is possible to reference more than one FILE. The reference to the FILE instance with FILE_TYPE="Main" has to appear first in the XML file.

Table 54 –child elements definition for element DOCUMENT_GENERIC

DOCUMENT_NUMBER

The registrant often has a list of historical and additional numbers for each document. The reason for the existence of different numbers on the cover page can be multifaceted. For example, a report of a residue trial can be marked by the following numbers:

- Report number
- Study Code
- Project No.
- Plot No(s).
- Company DocID
- Old document number

All these additional document numbers can be submitted via the multi valued element DOCUMENT_NUMBER.

DOCUMENT_AUTHOR

The element DOCUMENT_AUTHOR contains the author(s) of a document in one string, which is to be used unchanged in the content of document type “reference lists”. It should be possible for Regulatory authorities to split the authors’ string into individual authors while importing the document data. Therefore it is necessary to consider the following business rules:

Each Author is cited by Surname + comma + blank + each first initial of first name + point. In case of multiple first names use a blank between first names. Titles (e.g. Dr.) are not a part of the author name. Authors are separated by semicolon + blank. Don't use concatenation of authors with "and". If there are more than 3 authors, it is possible to end with "et al." without a semicolon.

Examples:

- Roemer, K. G.; Torres Alanis, O.; Garcia de Torres, G.
- Whitford, F. A.; Kronenberg, J.; Lunchick, C. et al.

PUBLISHED_INDICATOR

PUBLISHED_INDICATOR indicates whether a document has been published and therefore made publicly available.

- If the PUBLISHED_INDICATOR is false repeat the owner of the document in the element COMPLETE_DOCUMENT_SOURCE.
- If the document is published (PUBLISHED_INDICATOR = true) submit the publication information according to one of the following two choices:
 7. Use the OECD citation guidance.
Set the value of COMPLETE_DOCUMENT_SOURCE to one string with the full source citation like Journal, [issue], [volume], (year): pages.
 8. Split the bibliographic data in detail fields:
DOCUMENT_SOURCE (contains the title of the journal / book), DOCUMENT_YEAR, DOCUMENT_ISSUE, DOCUMENT_VOLUME and DOCUMENT_PAGES.
It is recommended that reference lists contain as DOCUMENT_SOURCE the concatenation of all non-empty elements of the split bibliographic data detail fields listed above.

4.9.20.2 Constraints

Constraint Name	Description
	The element DOCUMENT_NUMBER is mandatory for unpublished documents (PUBLISHED_INDICATOR = false).
	The value of DOCUMENT_TITLE will be used unchanged in the content of reference lists (The check of this constraint is outside of the scope of GHSTS).
	The value of DOCUMENT_AUTHOR will be used unchanged in the contents of reference lists (The check of this constraint is outside of the scope of GHSTS).
	The value of DOCUMENT_ISSUE_DATE will be used unchanged in the contents reference lists (The check of this constraint is outside of the scope of GHSTS).

	If the PUBLISHED_INDICATOR = true then repeat the value of DOCUMENT_OWNER in the element COMPLETE_DOCUMENT_SOURCE.
	At least one element REFERENCED_TO_FILE must exist. At least one referenced FILE instance must have FILE_TYPE="Main".
	In the values of the element REFERENCED_TO_FILE the reference to the FILE instance with FILE_TYPE="Main" has to appear first in the XML file.

Table 55 - constraints definition for element DOCUMENT_GENERIC**4.9.21 CONTENT_STATUS_HISTORY**

The element CONTENT_STATUS_HISTORY contains the full document content history list from the initial appearance of the document in the dossier to the current submission. Please note that this history is not RA-specific.

For the explanation of the logical concept please see chapter 3.2.17.

4.9.21.1 Child elements

Element name	CONTENT_STATUS_HISTORY			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/CONTENT_STATUS_HISTORY			
Diagram				
Child element	T	Occ	Definition	Description
CONTENT_STATUS	E	(1,*)	chapter 4.10.2	content status of the document, see chapter 3.2.17
SUBMISSION_NUMBER	T	(1,*)	chapter 4.10.6	submission number Examples: "01", "02", "03"

Table 56 –child elements definition for element CONTENT_STATUS_HISTORY**4.9.21.2 Constraints**

Constraint Name	Description
	Multiple entries for CONTENT_STATUS_HISTORY in the XML file are sorted with ascending submission number
	There must be only one entry with CONTENT_STATUS_TYPE = "New" as first entry
	If there is an entry CONTENT_STATUS_TYPE="Retired" then this is the final entry of the history. A retired document cannot be "reanimated".
	The CONTENT_STATUS_HISTORY is complete if there exists a unique record for each SUBMISSION NUMBER between the submission number with CONTENT_STATUS_TYPE="New" and the current submission number.

Table 57 - constraints definition for element CONTENT_STATUS_HISTORY**4.9.22 REFERENCED_DOCUMENT**

Using the element REFERENCED_DOCUMENT the registrant can supply supporting references from the current document to other documents that are part of the current GHSTS submission package or outside.

For the explanation of the logical concept please see chapter 3.2.25.

4.9.22.1 Child elements

Element name	REFERENCED_DOCUMENT			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/REFERENCED_DOCUMENT			
Diagram				
Child element	T	Occ	Definition	Description
REFERENCE_TYPE	E	(1,1)	chapter 3.2.23, 5.2	fixed picklist defined in GHSTS picklist XSD in type TYPE_REFERENCE_TYPE The element specifies the semantics of the reference to the target document. Examples: “document to submitted document”, “document to replaced document”, “translation to document”
INTERNAL	E	(1,1)	xs:boolean	The element specifies where the target document is internal (within the current dossier and its submissions) or external (outside of current dossier). Note: Internal references to documents only part of preceding submissions can only occur in the rare case of referencing replaced documents, as the XML backbone is “complete” in all other cases. Hence, internal references will most likely reference documents in the current GHSTS XML backbone.
DOCUMENT_PID (choice option 1)	T	(1,1)	PIDType, chapter 4.10.7	PID of target document must be supplied if INTERNAL is set to “true”. If “false”, DOCUMENT_PID should be supplied, if document already possesses a PID that is known to the receivers.

DOCUMENT_NUMBER (choice option 2)	E	(1,1)	chapter 4.6	<p>identifier and open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_DOCUMENT_NUMBER_TYPE</p> <p>A document number has to be supplied in cases, where a PID cannot be supplied.</p> <p>For usage comments see chapter 4.9.20.1.</p>
--------------------------------------	---	-------	-------------	--

Table 58 –child elements definition for element REFERENCED_DOCUMENT

For the XSD choice element in REFERENCED_DOCUMENT one of the two elements DOCUMENT_PID or DOCUMENT_NUMBER must be supplied.

4.9.23 RELATED_TO_SUBSTANCE

The element RELATED_TO_SUBSTANCE is useful to characterize, that a study is related only to one of the active substances, a metabolite or a comparable active substance. Example: This is useful for residue trials, where only one of the active substances is analysed.

Every referred substance has to be listed under SUBSTANCES.

For the explanation of the logical concept please see chapter 3.2.25.

4.9.23.1 Child elements

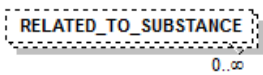
Element name	RELATED_TO_SUBSTANCE			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/RELATED_TO_SUBSTANCE			
Diagram				
Child element	T	Occ	Definition	Description
This element does not possess child elements, only an attribute				

Table 59 –child elements definition for element RELATED_TO_SUBSTANCE

4.9.23.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Substance_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityDocumentToSubstance

Table 60 - attributes definition for element RELATED_TO_SUBSTANCE

4.9.24 REFERENCED_TO_FILE

For the explanation of the logical concept please see chapter 3.2.12.

4.9.24.1 Child elements

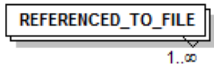
Element name	REFERENCED_TO_FILE			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/REFERENCED_TO_FILE			
Diagram				
Child element	T	Occ	Definition	Description
This element does not possess child elements, only an attribute				

Table 61 –child elements definition for element REFERENCED_TO_FILE

4.9.24.2 Attributes

Attribute name	Man	Type	Definition	Description
To_File_Id	Y	XS:IDREF	chapter 4.5	ID reference to FILE instance. Reference validated in schema using keyref RefIntegrityDocumentToFile

Table 62 - attributes definition for element REFERENCED_TO_FILE

4.9.25 FILES

The element FILES contains a list of files.

4.9.25.1 Child elements

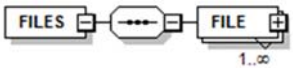
Element name	FILES			
Xpath	GHSTS/FILES			
Diagram				
Child element	T	Occ	Definition	Description
FILE	E	(1,*)	chapter 4.9.26	single FILE instance

Table 63 –child elements definition for element FILES

4.9.25.2 Constraints

Constraint Name	Description
	For each DOCUMENT instance: The sequence of FILEs inside of the element FILES has to be the file order to read the DOCUMENT instance. This means that for each DOCUMENT instance the FILE instance with FILE_TYPE="Main" has to appear first in the XML file.

Table 64 - constraints definition for element FILES

4.9.26 FILE

For the explanation of the logical concept please see chapter 3.2.11.

4.9.26.1 Child elements

Element name	FILE			
Xpath	GHSTS/FILES/FILE			
Diagram				
Child element	T	Occ	Definition	Description
FILE_RA	E	(1,*)	chapter 4.9.27	element to specify at which RA the file is targeted and to supply additional metadata
FILE_GENERIC	E	(1,1)	chapter 4.9.28	element with metadata not dependent on RA

Table 65 –child elements definition for element FILE

4.9.26.2 Attributes

Attribute name	Man	Type	Definition	Description
Id	Y	XS:ID	chapter 4.5	unique ID, which is used for reference in other elements

Table 66 - attributes definition for element FILE

4.9.27 FILE_RA

With the help of the element FILE_RA it is possible to send Regulatory authority specific metadata for each computer file.

The element FILE_RA is present for each RA to which the registrant wants to assign a specific file. In addition the element contains RA specific information that may differ between the different RA receiving the same file. The information that is invariant for all RA is represented below the element FILE_GENERIC (see chapter 4.9.28).

For the explanation of the logical concept please see chapter 3.2.24.

4.9.27.1 Child elements

Element name	FILE_RA			
Xpath	GHSTS/FILES/FILE/FILE_RA			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	Metadata status of all child elements of FILE_RA. see chapter 3.2.3
CBI_DESIGNATION	T	(1,1)	xs:boolean	indicates whether the file contains CBI-protected information (CBI = Confidential Business Information) for this RA
FILE_COMMENT	T	(0,1)	string2000, chapter 4.10.11	allows the registrant to communicate questions, notes, or other remarks to the agency at the FILE level

Table 67 –child elements definition for element FILE_RA

4.9.27.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Specific_for_RA_Id	Y	XS:IDREF	chapter 4.5	ID reference to legal entity instance. Reference validated in schema using keyref RefIntegrityFILE_RAToLegalentity

Table 68 - attributes definition for element FILE_RA

4.9.27.3 Constraints

Constraint Name	Description
	The element has to be supplied for all receivers of the submission

Table 69 - constraints definition for element FILE_RA

4.9.28 FILE_GENERIC

This element groups all metadata on the FILE level that is identical for all receiving RA.

4.9.28.1 Child elements

Element name	FILE			
Xpath	GHSTS/FILES/FILE_GENERIC			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	Any element within FILE_GENERIC except for the block CONTENT_STATUS that contains the information for content changes. see chapter 3.2.3
FILE_PID	T	(1,1)	chapter 4.10.7	see chapter 3.2.20
FILE_COMPANY_ID	T	(0,1)	string255, chapter 4.10.11	registrant internal file identification number Example: “123789456A1b”, “MO-99-009136-001”
CONTENT_STATUS	E	(1,1)	chapter 4.10.2	see chapter 3.2.16
REPLACED_FILE_PID	T	(0,1)	chapter 4.10.7	see chapter 3.2.16
FILE_TYPE	E	(1,1)	chapter 3.2.23, 5.2	fixed picklist defined in GHSTS picklist XSD in type TYPE_FILE_TYPE Example: “Main”
FORMAT_COMMENT	T	(0,1)	string255, chapter 4.10.11	optional textual comment to explain the file format and the originating applications including the application version, to assist the receiver in opening the file in the correct application, also when significant time between initial submission and retrieval has passed.

MD5CHECKSUM	T	(1,1)	chapter 4.10.9	Checksum of the computer file, generated by the registrant. Can be used by receiver to detect data transmission errors between sender – transport medium – receiver. Not relevant for statement about content changes. Example: “79054025255fb1a26e4bc422aef54eb4”
FILENAME	T	(1,1)	chapter 4.10.4	see chapter 6.4

Table 70 –child elements definition for element FILE_GENERIC

4.9.28.2 Constraints

Constraint Name	Description
	Every FILE instance has to be referenced by exactly one DOCUMENT instance.

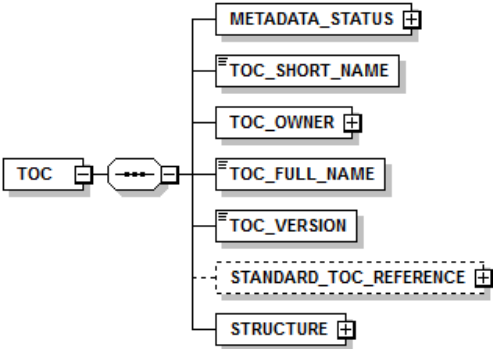
Table 71 - constraints definition for element FILE_GENERIC

4.9.29 TOC

The ToC branch in the GHSTS XML backbone file is a near 1:1 copy of an instance of a standard ToC XML file, enriched with a few additional attributes. It is not permitted to alter element text values for all elements that are direct copies from the standard ToC XML file. For these elements the description below will contain the information “as supplied with standard ToC”.

For the explanation of the logical concept please see chapter 3.2.7.

4.9.29.1 Child elements

Element name	TOC			
Xpath	GHSTS/TOC			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	see chapter 3.2.3
TOC_SHORT_NAME	T	(1,1)	string20, chapter 4.10.11	as supplied with standard ToC

TOC_OWNER	E	(1,1)	chapter 3.2.23, 5.2	open picklist defined in GHSTS picklist XSD in type EXTENSION_TYPE_TOC_OWNER
TOC_FULL_NAME	T	(1,1)	string255, chapter 4.10.11	as supplied with standard ToC
TOC_VERSION	T	(1,1)	versionnumberType	as supplied with standard ToC
STANDARD_TOC_REFERENCE	E	(1,1)	chapter 4.9.30	see chapters 2.4.8, 3.2.7
STRUCTURE	E	(0,1)	chapter 4.9.31	as supplied with standard ToC

Table 72 –child elements definition for element TOC

4.9.30 STANDARD_TOC_REFERENCE

STANDARD_TOC_REFERENCE provides information on what standard ToC is used in the GHSTS XML backbone. The GHSTS validator will use this information to validate the embedded ToC below the TOC element against the standard ToC.

4.9.30.1 Child elements

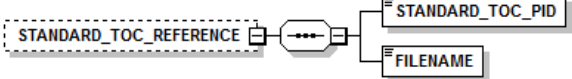
Element name	STANDARD_TOC_REFERENCE			
Xpath	GHSTS/TOC/STANDARD_TOC_REFERENCE			
Diagram				
Child element	T	Occ	Definition	Description
STANDARD_TOC_PID	T	(1,1)	PIDType, chapter 4.10.7	PID of the standard ToC, see chapter 3.2.20.
FILENAME	T	(1,1)	filenamestring, chapter 4.10.4	relative path to Toc file in Toc folder, see chapter 6.1.

Table 73 –child elements definition for element STANDARD_TOC_REFERENCE

4.9.30.2 Constraints

Constraint Name	Description
	If a standard ToC is referenced, validate the TOC_NODES / TOC_PID below the element TOC in the GHSTS backbone XML file against the standard ToC. If an online connection is available, the embedded ToC will be validated against the ToC from the centrally managed platform (see chapter 3.2.20).

Table 74 - constraints definition for element STANDARD_TOC_REFERENCE

4.9.31 STRUCTURE

STRUCTURE is at the top level of the Toc hierarchy.

4.9.31.1 Child elements

Element name	STRUCTURE			
Xpath	GHSTS/TOC/STRUCTURE			
Diagram	<pre> graph LR STRUCTURE[STRUCTURE] --- 1..∞ TOC_NODE[TOC_NODE] </pre>			
Child element	T	Occ	Definition	Description
TOC_NODE	E	(1,*)	chapter 4.10.2.1	as supplied with standard ToC

Table 75 –child elements definition for element STRUCTURE

4.9.32 LEGAL_ENTITIES

The element LEGAL_ENTITIES is a list of all Legal Entities which are used as a RECEIVER (Regulatory authorities) or as a SENDER (registrant)

4.9.32.1 Child elements

Element name	LEGAL_ENTITIES			
Xpath	GHSTS/LEGAL_ENTITIES			
Diagram	<pre> graph LR LEGAL_ENTITIES[LEGAL_ENTITIES] --- 1..∞ LEGAL_ENTITY[LEGAL_ENTITY] </pre>			
Child element	T	Occ	Definition	Description
LEGAL_ENTITY	E	(1,*)	chapter 4.9.33	element for legal entity

Table 76 –child elements definition for element LEGAL_ENTITIES

4.9.33 LEGAL_ENTITY

The element LEGAL_ENTITY contains the metadata of one Legal Entity.

For the explanation of the logical concept please see chapter 3.2.8.

4.9.33.1 Child elements

Element name	LEGAL_ENTITY			
Xpath	GHSTS/LEGAL_ENTITIES/LEGAL_ENTITY			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	see chapter 3.2.3
LEGALENTITY_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the Legal Entity, see chapter 3.2.20. Use of the “nil PID” is permitted.
LEGALENTITY_NAME	T	(1,1)	string255, chapter 4.10.11	standard name of the Legal Entity. Line feeds are allowed in this element.
LEGALENTITY_TYPE	E	(0,1)	chapter 3.2.23, 5.2	open picklist defined in GHSTS picklist XSD in type EXTENSION_TYPE_LEGALENTITY_TYPE.
OTHER_NAME	T	(0,*)	string255, chapter 4.10.11	alternative name of the Legal Entity.
LEGALENTITY_IDENTIFIER	E	(0,*)	chapter 4.6	identifier and open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_LEGALENTITY_IDENTIFIER_TYPE. Characterize Legal Entities by other worldwide established identifier e.g. the DUNS-number or the VAT-number.
CONTACT_ADDRESS	E	(1,1)	chapter 4.9.34	element for the official contact of the legal entity
CONTACT_PERSON	E	(0,*)	chapter 4.9.35	element for content person(s)

Table 77 –child elements definition for element LEGAL_ENTITY

A replacement of a LEGALITY_NAME by a newly assigned common name will not impact the previously assigned LEGALITY_PID.

4.9.33.2 Attributes

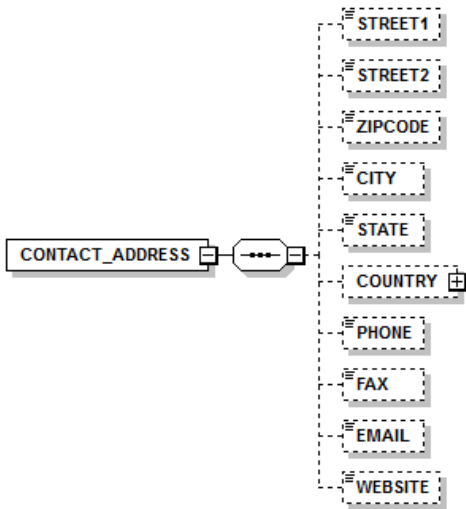
Attribute name	Man	Type	Definition	Description
Id	Y	XS:ID	chapter 4.5	unique ID, which is used for reference in other elements

Table 78 - attributes definition for element LEGAL_ENTITY

4.9.34 CONTACT_ADDRESS

Each Legal Entity is characterized by a mandatory CONTACT_ADDRESS.

4.9.34.1 Child elements

Element name	CONTACT_ADDRESS			
Xpath	GHSTS/LEGAL_ENTITIES/LEGAL_ENTITY/CONTACT_ADDRESS			
Diagram				
Child element	T	Occ	Definition	Description
STREET1	T	(0,1)	string255, chapter 4.10.11	Part of postal address Examples: “Geschäftsbereich Pflanzenschutz”, 608E-2720 Riverside Drive“
STREET2	T	(0,1)	string255, chapter 4.10.11	Part of postal address Examples: “Alfred-Nobel-Straße 50”, Tupper Building, E6607”
ZIPCODE	T	(0,1)	string255, chapter 4.10.11	codes established by postal authorities for purposes of sorting and delivering mail. Example: “D-40789”

CITY	T	(0,1)	string255, chapter 4.10.11	Represents the city, town, village, or hamlet. Named 'town' or 'municipality' in the Rose Model Example: “Monheim am Rhein”
STATE	T	(0,1)	string255, chapter 4.10.11	Represents the State, Province, and/or County.
COUNTRY	E	(0,1)	chapter 3.2.23, 5.2	open picklist defined in GHSTS picklist XSD in type EXTENSION_TYPE_COUNTRY
PHONE	T	(0,1)	string255, chapter 4.10.11	Telephone number including country code Example: “+01 613-736-3666”
FAX	T	(0,1)	string255, chapter 4.10.11	Fax number including country code Example: “+49 2173 38 1234”
EMAIL	T	(0,1)	string255, chapter 4.10.11	Email address of the official contact. Format will not be validated
WEBSITE	T	(0,1)	xs:anyURI	Website URL. Format will not be validated. Example: www.bayercropscience.com

Table 79 –child elements definition for element CONTACT_ADDRESS

4.9.34.2 Constraints

Constraint Name	Description
	The elements of CONTACT_ADDRESS are all optional. The registrant has to take care that a reasonable subset of the elements are supplied. No validation will be done on the elements. The combination of LEGALENTITY_NAME, STREET1, STREET2, ZIPCODE, CITY, STATE and COUNTRY should be usable for postal service.

Table 80 - constraints definition for element CONTACT_ADDRESS

4.9.35 CONTACT_PERSON

It is possible to define one or more contact persons.

4.9.35.1 Child elements

Element name	CONTACT_PERSON			
Xpath	GHSTS/LEGAL_ENTITIES/LEGAL_ENTITY/CONTACT_PERSON			
Diagram	<pre> classDiagram class CONTACT_PERSON { ORGANISATION DEPARTMENT TITLE FIRSTNAME LASTNAME PHONE MOBILE FAX EMAIL } CONTACT_PERSON "0..∞" --> "*" </pre>			
Child element	T	Occ	Definition	Description
ORGANISATION	T	(0,1)	string255, chapter 4.10.11	company name of the contact person if it is different from the parent LEGALENTITY_NAME.
DEPARTMENT	T	(0,1)	string255, chapter 4.10.11	the department of the contact person
TITLE	T	(0,1)	string255, chapter 4.10.11	title / salutation of the individual
FIRSTNAME	T	(0,1)	string255, chapter 4.10.11	firstname of the person
LASTNAME	T	(0,1)	string255, chapter 4.10.11	surname of the person
PHONE	T	(0,1)	string255, chapter 4.10.11	Telephone number including country code Example: "+01 613-736-3666"
MOBILE	T	(0,1)	string255, chapter 4.10.11	Mobile number including country code Example: "+49 172 2345 678"
FAX	T	(0,1)	string255, chapter 4.10.11	Fax number including country code Example: "+49 2173 38 1234"
EMAIL	T	(0,1)	string255, chapter 4.10.11	Email address of the contact person. Format will not be validated

Table 81 –child elements definition for element CONTACT_PERSON

4.9.35.2 Constraints

Constraint Name	Description
	The elements of CONTACT_PERSON are all optional. The registrant has to take care that a reasonable subset of the elements are supplied. No validation will be done on the elements.

Table 82 - constraints definition for element CONTACT_PERSON**4.9.36 SUBSTANCES**

This element represents a list of substances.

4.9.36.1 Child elements

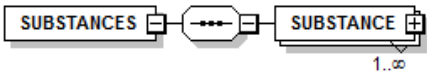
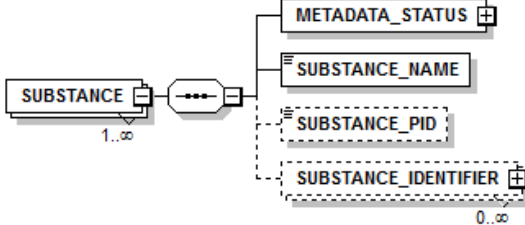
Element name	SUBSTANCES			
Xpath	GHSTS/SUBSTANCES			
Diagram				
Child element	T	Occ	Definition	Description
SUBSTANCE	E	(1,*)	chapter 4.9.37	element for substance

Table 83 –child elements definition for element SUBSTANCES**4.9.37 SUBSTANCE**

For the explanation of the logical concept please see chapter 3.2.10.

4.9.37.1 Child elements

Element name	SUBSTANCE			
Xpath	GHSTS/SUBSTANCES/SUBSTANCE			
Diagram				
Child element	T	Occ	Definition	Description
METADATA_STATUS	E	(1,1)	chapter 4.10.1	see chapter 3.2.3
SUBSTANCE_NAME	T	(1,1)	string2000, chapter 4.10.11	common name of the substance if available otherwise a registrant internal name Example: "Spiroxamine"
SUBSTANCE_PID	T	(0,1)	PIDType, chapter 4.10.7	PID of the pure reference substance, see chapter 3.2.20.

SUBSTANCE_IDENTIFIER	E	(0,*)	chapter 4.6	<p>identifier and open picklist, defined in GHSTS picklist XSD in type EXTENSION_TYPE_SUBSTANCE_IDENTIFIER_TYPE</p> <p>Characterization by other identifier, e.g. the Chemical Abstracts Number (CASNO) or European Commission Number (ECNO).</p> <p>For CASNO: In normal case: one CAS-Number per chemical ingredient. Only in cases of isomers and mixtures send all correspondent CAS-Numbers.</p> <p>Examples: "118134-30-8", "27012-62-0"</p>
----------------------	---	-------	-------------	--

Table 84 –child elements definition for element SUBSTANCE

4.9.37.2 Attributes

Attribute name	Man	Type	Definition	Description
Id	Y	XS:ID	chapter 4.5	unique ID, which is used for reference in other elements

Table 85 - attributes definition for element SUBSTANCE

4.9.37.3 Constraints

Constraint Name	Description
	<p>If a CAS-Number is supplied in SUBSTANCE_IDENTIFIER: Check the validity with CAS-Rules</p> <p>http://www.cas.org/content/chemical-substances/checkdig</p>

Table 86 - constraints definition for element SUBSTANCE

4.9.38 USED_TEMPLATES

Attention: This element is unused and kept as a placeholder. Currently under review.

4.10 Basic types

GHSTS defines a set of basic types.

4.10.1 Metadata status type

This element provides information about changes in the text content of all sibling elements and all their direct and indirect child elements. The only exception is for the DOCUMENT_GENERIC element (see chapter 4.9.20).

For the explanation of the logical concept please see chapter 3.2.3.

4.10.1.1 Child elements

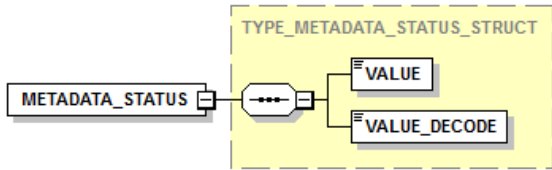
Element name	METADATA_STATUS			
Xpath	GHSTS/RECEIVERS/RECEIVER/METADATA_STATUS GHSTS/PRODUCT/METADATA_STATUS GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_RA/METADATA_STATUS GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/METADATA_STATUS GHSTS/FILES/FILE_RA/METADATA_STATUS GHSTS/FILES/FILE_GENERIC/METADATA_STATUS GHSTS/TOC/METADATA_STATUS GHSTS/LEGAL_ENTITIES/LEGAL_ENTITY/METADATA_STATUS GHSTS/SUBSTANCES/SUBSTANCE/METADATA_STATUS			
Diagram				
Child element	T	Occ	Definition	Description
VALUE	T	(1,1)	chapter 3.2.23, 5.2	fixed picklist defined in GHSTS picklist XSD in type TYPE_METADATA_STATUS Examples: "New", "No Changed", "Modified".
VALUE_DECODE	T	(1,1)	string20, chapter 4.10.11	

Table 87 –child elements definition for element METADATA_STATUS

4.10.2 Content status type

For the explanation of the logical concept please see chapters 3.2.15 and following.

4.10.2.1 Child elements

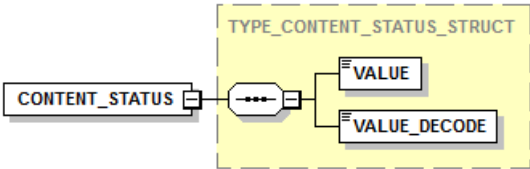
Element name	CONTENT_STATUS			
Xpath	GHSTS/DOCUMENTS/DOCUMENT/DOCUMENT_GENERIC/CONTENT_STATUS_HISTORY/CONTENT_STATUS GHSTS/FILES/FILE/CONTENT_STATUS			
Diagram				
Child element	T	Occ	Definition	Description
VALUE	T	(1,1)	chapter 3.2.23, 5.2	fixed picklist, defined in GHSTS picklist XSD in type TYPE_CONTENT_STATUS Examples: "New", "No Change", "Modified", "Replaced", "Retired".
VALUE_DECODE	T	(1,1)	string20, chapter 4.10.11	

Table 88 –child elements definition for element CONTENT_STATUS

4.10.3 TOC node type

The ToC node type is a near 1:1 copy of the ToC node type of a standard ToC XML file (see chapter 5.1.6.4), enriched with a few additional attributes. The copy in the GHSTS submission must include also the ToC nodes from the standard ToC that are marked as logical deleted.

It is not permitted to alter element text values for all elements that are direct copies from the standard ToC XML file. For these elements the description below will contain the information "as supplied with standard ToC"

4.10.3.1 Child elements

Element name	TOC_NODE			
Xpath	//TOC_NODE GHSTS/TOC/STRUCTURE			
Diagram				
Child element	T	Occ	Definition	Description
NODE_NAME	T	(1,1)	string20, chapter 4.10.11	as supplied with standard ToC
NODE_HEADING	T	(1,1)	string255, chapter 4.10.11	as supplied with standard ToC
LOGICAL_DELETED	T	(1,1)	xs:boolean	as supplied with standard ToC If value is TUE, it is not permitted to have any elements TOC2DOC for this node or any direct or indirect subnode.
TOC_NODE_PID	T	(0,1)	PIDType, chapter 4.10.7	as supplied with standard ToC, see chapter 3.2.20.
EMPTY_NODE	T	(1,1)	xs:boolean	Business information supplied by the registrant. If “true” this indicates that this node and all direct and indirect subnodes were deliberately left empty, if “false”, then at least one document is in current node or any subnode. Please note that this is not technical information, for example, a viewer program or ingestion component should rely on.

TOC2DOC	E	(0,*)		References from ToC node to document instances The sequential order of references to documents in the XML file is identical to the order in which the documents should appear in the ToC node. The order is relevant for referencing documents in reference lists.
TOC_NODE	E	(0,*)	TYPE_TOC_NODE	as supplied with standard ToC

Table 89 –child elements definition for element TOC_NODE

4.10.3.2 TOC2DOC

4.10.3.2.1 Child elements

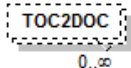
Element name	TOC2DOC			
Xpath	/TOC_NODE/TOC2DOC			
Diagram				
Child element	T	Occ	Definition	Description
This element does not possess child elements, only an attribute				

Table 90 –child elements definition for element TOC2DOC

4.10.3.2.2 Attributes

Attribute name	Man	Type	Definition	Description
To_Document_Id				ID reference to document instance. Reference validated in schema using keyref RefIntegrityTocnodeToDocument

Table 91 - attributes definition for element TOC2DOC

4.10.4 Relative file path type

The type filenamestring defines what a relative path from the GHSTS XML backbone to a computer file location on the file system should be like:

- References to computer files only use relative paths.
- The GHSTS backbone file is located in the submission folder (see chapter 6.1). The relative path always first moves one folder level up to the regulatory action folder using “..” and then descends into the submission folder, consisting of two digits (see chapter 4.10.6). This is true even if the relative path points to a computer file in the current submission.
- The path separator is always “/” independent on the current platform.
- The folder names and file names only support a restricted set of characters.

- The total length of the relative path is limited to 230 characters. This is due to the limitation of some file systems with a total path length of 255 characters.

The type filenamestring is used in the following elements:

- GHSTS/FILES/FILE/FILENAME
- GHSTS/TOC/STANDARD_TOC_REFERENCE/FILENAME

For recommendations on how to choose good folder and file names see chapters 6.3 and 6.4.

4.10.4.1 Definition

```
<xs:simpleType name="filenamestring">
  <xs:annotation>
    <xs:documentation>definition of relative paths for filenames, used in FILE</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:whiteSpace value="preserve"/>
    <xs:minLength value="1"/>
    <xs:maxLength value="230"/>
    <xs:pattern value="(\.\\.)d{2}/((([a-zA-Z0-9_\\-\\.]*)/)*([a-zA-Z0-9_\\-\\.]\\(\\)+)"/>
  </xs:restriction>
</xs:simpleType>
```

4.10.4.1 Examples

```
<FILENAME>../02/attachments/C6313_0016-suppl.bmp</FILENAME>
```

4.10.5 Version number type

The version number type defines the versioning scheme used to specify the GHSTS specification to which a GHSTS submission package adheres.

4.10.5.1 Definition

```
<xs:simpleType name="versionnumberType">
  <xs:annotation>
    <xs:documentation>Version number: six digits, e.g. 01.00.00</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="d{2}.d{2}.d{2}"/>
  </xs:restriction>
</xs:simpleType>
```

4.10.5.2 Example

```
<GHSTS specificationversion="01.00.00" xsi:schemaLocation="http://www.oecd.org/GHSTS GHSTS_V010000.xsd"
xmlns="http://www.oecd.org/GHSTS" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
```

4.10.6 Submission number type

The submission number consists of two digits, starting with 01, 02, 03, The submission number has to be used as the name of the submission folder for the submission package (see chapter 6.1).

4.10.6.1 Definition

```
<xs:simpleType name="submissionnumberType">
  <xs:annotation>
    <xs:documentation>definition of submission number as two digit number, e.g. 01,02,03 ...
  </xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="\d{2}" />
  </xs:restriction>
</xs:simpleType>
```

4.10.7 PID type

The PID type describes how a Persistent Identifier for GHSTS is represented.

The PID is built according to the Uniform Resource Name (URN) definition, see e.g. http://de.wikipedia.org/wiki/Uniform_Resource_Name.

A PID in the form of the URN is built according to the pattern

```
urn:<namespace>:<value>
```

The namespace is an abbreviation for the body that has issued the URN. There is a list of officially registered URN namespaces (e.g. see <http://www.iana.org/assignments/urn-namespaces/urn-namespaces.xml>), however it is possible to use the URN schema without an officially registered namespace. For example, the registrant Syngenta could use “syngenta” as namespace in the PID; Bayer could use “bayer” etc. The namespace “ghsts” is used for the centrally managed identifiers. The GMB (see chapter 2.4.9) can also manage a list of namespaces that can easily be aligned with the list of legal entities.

The value part has to be unique per namespace to make the complete PID unique.

The advantage of the URN format is that it is open to all kinds of namespace and types of values. In this sense GHSTS is open for the UUID identifier format that is generated by its own governing body GMB or for other identifier formats that are taken from a different body. With the namespace part of the URN the sources of the identifiers are easy to distinguish.

For example, the existing IUCLID5 “UUID” examples (please note that IUCLID calls these identifiers UUIDs whereas they are not compliant with the standard UUID definition)

```
ECB5-d966af20-7dea-42ad-a7be-87278757dbab
IUC5-1a16228d-42a1-4ed3-a405-be8728e17f8e
```

can be directly expressed as PID for GHSTS:

```
urn:ECB5:d966af20-7dea-42ad-a7be-87278757dbab
urn:IUC5:1a16228d-42a1-4ed3-a405-be8728e17f8e
```

For the value part of the PID that are either generated locally by registrants or that are issued by the GMB (see chapter 2.4.9), the UUID format is used. UUID stands for Universally Unique Identifier and is usually used in software development. However, as UUIDs can be easily created and all common programming environments support the creation of UUIDs it can also ideally serve as format for the value part of the PID.

The technical definition of the PID type is not restricted to the UUID format as the only value format for the PID, so it can transport values of other namespaces and value types.

For mandatory PID that are managed centralized but where a PID is not yet available (e.g. legal entity), the registrant can use the so-called “nil PID” or “fake PID”

urn:uuid:00000000-0000-0000-0000-000000000000

in the submission instead of a real PID.

Attention: The format, to be used for centrally managed PID, which is currently under consideration.

For details about the UUID format please see chapter 4.10.8.

4.10.7.1 Definition

```
<xs:simpleType name="PIDType">
  <xs:annotation>
    <xs:documentation>definition of Persistent Identifier (PID) according to URN pattern
    as defined in RFC 2141</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:anyURI">
    <xs:pattern value="urn:[a-zA-Z0-9][a-zA-Z0-9-]{1,31}:[a-zA-Z0-9()+,.;:=@;$_!*-]||%[0-9A-Fa-f]{2})+"/>
  </xs:restriction>
</xs:simpleType>
```

4.10.7.2 Constraints

Constraint Name	Description
	The GHSTS validator will check that the PID format is used for a defined set of namespaces.

Table 92 - constraints definition for element PID

4.10.7.3 Example

Please note the different namespace values in the examples:

<SUBSTANCE_PID>urn:ghsts:1a16228d-42a1-4ed3-a405-be8728e17f8e</SUBSTANCE_PID>

<DOCUMENT_PID>urn:basf:d966af20-7dea-42ad-a7be-87278757dbab </DOCUMENT_PID>

4.10.8 UUID type

The UUID type definition is provided here for the registrants that need to create PID with the UUID as value part. The GHSTS XSD itself will not use this definition.

Technically a UUID is a 16-byte (128-bit) number. It is represented in a canonical form by 32 hexadecimal digits, displayed in 5 groups separated by hyphens, in the form 8-4-4-4-12 for a total of 36 characters (32 digits and 4 hyphens). An example for a UUID in the representation is as follows:

761e6400-e26b-56d5-a482-663375330000

There are multiple subtypes of UUID, see http://en.wikipedia.org/wiki/Universally_unique_identifier. For UUIDs generated by the registrant (see “Table 11 – Overview of concepts identified by PID”), GHSTS doesn’t restrict the UUID version to be used but recommends to use one out of version 3 (MD5 hash), version 4 (random) and version 5 (SHA-1 hash).

All common programming languages support the creation of UUIDs. For details please refer to the programming language reference.

4.10.8.1 Definition

```
<xs:simpleType name="uuidType">
  <xs:annotation>
    <xs:documentation>definition of UUID according to canonical representation</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="[a-f0-9]{8}\-[a-f0-9]{4}\-[a-f0-9]{4}\-[a-f0-9]{4}\-[a-f0-9]{12}" />
  </xs:restriction>
</xs:simpleType>
```

4.10.9 MD5 type

The MD5 type is used in GHSTS for the element MD5CHECKSUM of the FILE concept. The MD5 Message-Digest Algorithm is a widely used cryptographic hash function that produces a 128-bit (16-byte) hash value, represented as hexadecimal number of 32 digits. It is specified in RFC 1321. For details see <http://en.wikipedia.org/wiki/MD5>.

The value is supplied by the registrant to enable the receiver to check the successful transmission process of the referenced computer file.

All common programming languages support the creation of MD5 value for a given file. For details please refer to the programming language reference.

4.10.9.1 Definition

```
<xs:simpleType name="md5Type">
  <xs:annotation>
    <xs:documentation>MD5 checksum definition, used in FILE</xs:documentation>
  </xs:annotation>
  <xs:restriction base="xs:string">
    <xs:pattern value="[a-f,A-F,0-9]{32}" />
  </xs:restriction>
</xs:simpleType>
```

4.10.10 Standard XS and XSI types

The following standard types of the xs namespace are used in the GHSTS XSD file:

- xs:anyURI
- xs:boolean
- xs:date
- xs:decimal
- xs:positiveInteger
- xsi:schemaLocation
- xs:ID
- xs:IDREF

4.10.11 Restricted string types

GHSTS defines several string types to restrict the length of a string. The following string types are defined: string20, string255, string2000, where the number is equivalent to the maximum string length.

4.10.11.1 Definition

Here is the definition of the string20 type, the other types are defined similar:

```
<xs:simpleType name="string20">  
  <xs:restriction base="xs:string">  
    <xs:minLength value="1"/>  
    <xs:maxLength value="20"/>  
  </xs:restriction>  
</xs:simpleType>
```

Please note that there is no restriction on whitespace characters, meaning that whitespaces other formatting characters are preserved.

5 REFERENCED SCHEMA DEFINITIONS

5.1 Standard ToC schema definition

5.1.1 Overview

This chapter provides a technical reference for the Table of Contents schema definition. This chapter contains the most relevant aspects of the XSD, especially an explanation on how elements should be used in the business context. The complete definition is contained in the XSD itself.

5.1.2 Filename and version

The ToC XSD is versioned using a version numbering scheme of type XX.XX.XX. It is composed of a two digit major, a two digit minor and a two digit patch level version number separated by single dots. Major versions are used for global functional changes, minor versions for local changes and patch level versions for error corrections. Patch level versions do not require new versions of other GHSTS components (see chapter 2.4).

This version is contained in the version attribute of the root schema element in the XSD file.

The filename uses the same versioning scheme with the hyphen instead of the dot. The file naming convention of the ToC XSD is therefore

`toc_XX-XX-XX.xsd`

for example

`toc_01-00-00.xsd`

Please note that the filename does not contain the string “ghsts” as the ToC definitions next to their usage in the context of GHSTS can be used in other plant protection processes.

For the instances of the GHSTS ToC XSD the following naming convention is proposed

`toc-<TOC_OWNER>-<TOC_SHORT_NAME>_XX-XX-XX.xml`

for example

`toc-oecd-standard_01-00-00.xml`

The values for ToC Owner and ToC short name are taken from the corresponding elements in the XML file (see below). Spaces or hyphens in the values are replaced with underscore.

This chapter describes the version 01.00.00 of the ToC XML schema, published on **XXX**.

5.1.3 Namespace

The ToC schema description file uses the following namespace definitions:

- `xmlns:xs="http://www.w3.org/2001/XMLSchema"`
- `xmlns="http://www.oecd.org/TOCS"`
- `targetNamespace=http://www.oecd.org/TOCS`

The namespace is deliberately different from the GHSTS submission schema definition.

5.1.4 Unique keys

The elements TOC_PID and TOC_NODE_PID are part of the identifiers that are globally managed (see chapter 3.2.20). Their values have to be globally unique.

The value of TOC_NODE_NAME has to be unique within one instance of a standard ToC.

5.1.5 Notation conventions

For notation conventions see chapter 4.7 and chapter 4.8.

5.1.6 Elements

5.1.6.1 Overview

An overview diagram of the complete XSD is provided as separate image file in PNG format, with the filename

toc_XX-XX-XX.png.

5.1.6.2 TOC

TOC is the definition of the ToC for the standard ToC templates.

5.1.6.2.1 Child elements

Element name	TOC			
Xpath	TOC			
Diagram				
Child element	T	Occ	Definition	Description
TOC_SHORT_NAME	T	(1,1)	string20, chapter 4.10.10	short internal identifier Example: “OECD Pest 2005”
TOC_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the ToC, see chapter 3.2.20.
TOC_OWNER	E	(1,1)	chapter 3.2.23, 5.2	fixed picklist defined in ToC XSD in type EXTENSION_TYPE_TOC_OWNER
TOC_FULL_NAME	T	(1,1)	string255	Name of ToC to be displayed in viewer Example: “OECD Plant Protection Products”

TOC_VERSION	T	(1,1)	versionnumberType	version of the ToC
TO_USE_FROM	T	(1,1)	xs:date	ToC valid from this date
TO_USE_UNTIL	T	(0,1)	xs:date	ToC valid until this date
STRUCTURE	E	(1,1)	chapter 4.9.31	Root element of the hierarchical ToC structure. Same element definition as in the GHSTS XSD.

Table 93 –child elements definition for element TOC

5.1.6.3 STRUCTURE

STRUCTURE is the top-level of the ToC hierarchy.

5.1.6.3.1 Child elements


Element name	STRUCTURE			
Xpath	TOC/STRUCTURE			
Diagram				
Child element	T	Occ	Definition	Description
TOC_NODE	E	(1,*)	chapter 5.1.6.4	The top-down sequence of the TOC_NODE instances in the XML file reflects the order of the ToC nodes in the ToC.

Table 94 –child elements definition for element STRUCTURE

5.1.6.4 TOC_NODE

TOC_NODE is the ToC node definition for a standard table of contents. Please note the difference between the definition here and the definition of a ToC node as part of a submission (see chapter 4.10.3).

5.1.6.4.1 Child elements

Element name	TOC_NODE			
Xpath	//TOC_NODE TOC/STRUCTURE			
Diagram				
Child element	T	Occ	Definition	Description
NODE_NAME	T	(1,1)	string20, chapter 4.10.10	code of an annex point, unique within the complete ToC Examples: "0.1.6003", "IIA-5.2.1"
NODE_HEADING	T	(1,1)	string255, chapter 4.10.10	textual context as a complete heading for outputs Example: "Acute oral toxicity"
LOGICAL_DELETED	T	(1,1)	xs:boolean	True, if node has been retired but is kept for information purposes. It is not permitted to reference documents to this node and all subnodes in a submission.
DOCUMENT_REFERENCES	T	(1,1)	xs:boolean	True, if this node can contain document references of type TOC2DOC (see chapter 4.10.3.2) when used in GHSTS XML backbone. If false, no document references are permitted from/to this node.
TOC_NODE_PID	T	(1,1)	PIDType, chapter 4.10.7	unique PID of the ToC node, see chapter 3.2.20.
TOC_NODE	E	(0,*)	TYPE_TOC_NODE, chapter 5.1.6.4	The top-down sequence of the TOC_NODE instances in the XML file reflects the order of the ToC nodes in the ToC

Table 95 –child elements definition for element TOC_NODE

5.1.6.4.2 Constraints

Constraint Name	Description
	The value for NODE_NAME has to be unique within the complete ToC
	The top-down sequence of the TOC_NODE instances in the XML file reflects the order of the ToC nodes in the ToC

Table 96 - constraints definition for element TOC_NODE

5.1.7 Basic types

The ToC XSD contains the definition basic types that are identical to the definition in the GHSTS submission XSD or the GHSTS picklist XSD, but are copied here to make the ToC XSD independent on GHSTS. However, the type definitions will be kept identical for future releases as well:

Type name	Definition identical to
string20	GHSTS submission XSD, see chapter 4.10.11
string255	GHSTS submission XSD, see chapter 4.10.11
PIDType	GHSTS submission XSD, see chapter 4.10.7
TYPE_TOC_NODE	GHSTS submission XSD, see chapter 4.10.3
TYPE_TOC_OWNER / EXTENSION_TYPE_TOC_OWNER	GHSTS picklist XSD, see chapter 3.2.23

Table 97 – definition of basic types in the ToC XSD

5.2 GHSTS picklist schema definition

5.2.1 Overview

The GHSTS picklist schema definition is an XSD that contains the definitions for the picklists as well as the picklist items that are used in the metadata elements in the GHSTS backbone XML file. The GHSTS submission schema definition uses a mechanism of picklists compatible to mechanism used by the OECD Harmonised Templates (OHT) for Reporting Chemical Test Summaries (see <http://www.oecd.org/ehs/templates>). The content of the picklists is independent of the OHT.

GHSTS picklists are implemented to harmonize the list of possible values which could be submitted within an element. The values can be used for decisions and trigger follow-up steps while importing the submitted data into in house systems.

Chapter 3.2.23 contains the details about the picklists used in GHSTS:

- Name of the element in the GHSTS XSD for which a picklist is applied
- Name of the picklist type as defined in the GHSTS picklist XSD
- Type of the picklist - “fixed” when only one of the listed items can be submitted, or “open” when the registrant can submit a new value in the GHSTS backbone XML

The registrant needs to understand the GHSTS picklist XSD because

- the list of picklists and picklist items that are defined in the XSD have to be part of the dossier management system to support the metadata assignment part of the GHSTS submission creation process
- an automated import of the picklists and picklist items requires a knowledge of the XSD

5.2.2 *Filename and version*

The GHSTS picklist XSD is versioned using a version numbering scheme of type XX.XX.XX. It is composed of a two digit major, a two digit minor and a two digit patch level version number separated by single dots. This version is contained in the version attribute of the root schema element in the XSD file.

The filename of the GHSTS picklist XSD is not changed when it is versioned, as it is referenced by the GHSTS submission XSD: If the picklist filename changed with every version, then also the submission XSD would need to be modified.

The constant file name of the GHSTS picklist XSD is

ghsts-picklists.xsd

The GHSTS picklist XSD may be versioned for two reasons:

- Picklist items were changed on the content level (new picklist items, disabling of picklist items, modification of decoded picklist value)
- The definitions of the picklists itself change (e.g. new picklists, change of picklist from type “open” to “fixed”)

The versioning scheme accommodates for this difference by using minor versions for content changes and major versions for definition changes. Patch level versions are used for error corrections (e.g. typos).

This chapter describes the version 00.09.00 of the GHSTS picklists XML schema, published on **XXXX**.

5.2.3 *Picklists*

A picklist contains picklist items. For each picklist item the following information is supplied in the GHSTS picklist XSD:

- Picklist value representing the “coded” value, meaning that this value is expected to be used by programs and will usually not be used for human display
- Decoded picklist value targeted for human display without additional interpretation, for example in the GHSTS viewer
- Status of the picklist item, either “enabled” or “disabled”. By this means picklist items can be deprecated by marking them as disabled, but leaving them in the GHSTS XSD picklist file.
- Origin of the picklist item, for example the regulatory authority /country that supplied this picklist item

Here is the example of the picklist definition for the fixed picklist for the FILE_TYPE element of the FILE. The three dots (...) in the definition are inserted where subsequent similar definitions have been omitted for sake of brevity.

```

<xs:simpleType name="TYPE_FILE_TYPE">
  <xs:restriction base="xs:string">
    <xs:enumeration value="Main">
      <xs:annotation>
        <xs:appinfo>
          <DECODE>Main</DECODE>
          <STATUS>enabled</STATUS>
          <ORIGIN>EGEEPD</ORIGIN>
        </xs:appinfo>
      </xs:annotation>
    </xs:enumeration>
    <xs:enumeration value="Source">
      <xs:annotation>
        <xs:appinfo>
          <DECODE>Source</DECODE>
          <STATUS>enabled</STATUS>
          <ORIGIN>EGEEPD</ORIGIN>
        </xs:appinfo>
      </xs:annotation>
    </xs:enumeration>
    ...
  </xs:restriction>
</xs:simpleType>

```

This type picklist definition is identical for fixed and open picklists.

The definition for the usage of this FILE_TYPE picklist in the GHSTS XSD is as follows:

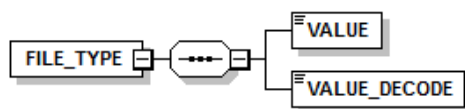


Figure 27 – FILE_TYPE picklist definition (fixed picklist)

The element VALUE_DECODE is of type string255. Following this definition the registrant uses the picklist item in the submission by supplying the coded and decoded value in the following way:

```

<FILE_TYPE>
.....<VALUE>Main</VALUE>
.....<VALUE_DECODE>Core content representation of the document</VALUE_DECODE>
...</FILE_TYPE>

```

For open picklists there is an additional extension type that is required to submit coded values that are not part of the picklist definition.

5.2.4 Open picklists

GHSTS open picklists are implemented as fixed picklists with an additional extension to submit other values.

The following example demonstrates this for the picklist for the element FORMULATION_TYPE which is part of the information of the product:

```

<xs:simpleType name="TYPE_FORMULATION_TYPE">
  <xs:restriction base="xs:string">
    <xs:enumeration value="AE">
      <xs:annotation>
        <xs:appinfo>
          <DECODE>Aerosol dispenser</DECODE>
          <STATUS>enabled</STATUS>
          <ORIGIN>GIFAP</ORIGIN>
        </xs:appinfo>
      </xs:annotation>
    </xs:enumeration>
    ...
  </xs:restriction>
</xs:simpleType>

```

```

</xs:annotation>
</xs:enumeration>
<xs:enumeration value="AL">
<xs:annotation>
  <xs:appinfo>
    <DECODE>Other liquids to be applied undiluted</DECODE>
    <STATUS>enabled</STATUS>
    <ORIGIN>GIFAP</ORIGIN>
  </xs:appinfo>
</xs:annotation>
</xs:enumeration>
...
</xs:restriction>
</xs:simpleType>

```

The definition of the extension type EXTENSION_TYPE_FORMULATION_TYPE allows submitting a new pair of coded and uncoded information:

```

<xs:complexType name="EXTENSION_TYPE_FORMULATION_TYPE">
  <xs:simpleContent>
    <xs:extension base="TYPE_FORMULATION_TYPE">
      <xs:attribute name="Other_Value" type="xs:string">
        <xs:annotation>
          <xs:documentation>Other values can be submitted</xs:documentation>
        </xs:annotation>
      </xs:attribute>
    </xs:extension>
  </xs:simpleContent>
</xs:complexType>

```

The definition for the usage of the FORMULATION_TYPE open picklist in the GHSTS XSD is identical to the definition of the fixed picklist:

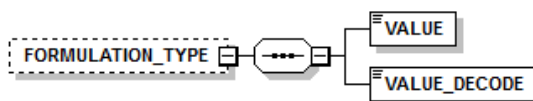


Figure 28 – FORMULATION_TYPE picklist definition (open picklist)

The submission of a known formulation type is also identical to fixed picklists:

```

...<FORMULATION_TYPE>
.....<VALUE>AL</VALUE>
.....<VALUE_DECODE>Other liquids to be applied undiluted</VALUE_DECODE>
...</FORMULATION_TYPE>

```

The submission of a new formulation type uses the new attribute Other_value defined in the extension:

```

...<FORMULATION_TYPE>
.....<VALUE Other_Value="NF">other</VALUE>
.....<VALUE_DECODE>New Formulation Type</VALUE_DECODE>
...</FORMULATION_TYPE>

```

The new coded value is supplied as attribute value of Other_value, the new decoded value is supplied as before.

The usage of a new value for an open picklist should be an exception process only: Missing elements in open picklists should be made part of the existing enumerations in the GHSTS picklist XSD file. Please contact the GMB to request the addition of a new value to a picklist.

6 GHSTS PACKAGE DEFINITION

This chapter describes the GHSTS package, which is the physical representation of a GHSTS submission.

6.1 Overview of Folder structure

The following table shows the folder structure of a GHSTS submission. As an example the displayed submission is the submission with submission number 03 of a series of submissions from 01 to 04.

















Parent hierarchy	Regulatory action level	Submission level	Name	Description (Bold = name of folder, as used in description of submission package)
				Folder hierarchy above dossier level (unknown and out of scope for this specification).
			SYNCH127 (example)	Regulatory action folder Root folder for the regulatory action / dossier. Only submissions of this dossier allowed. The GHSTS viewer will not be able to automatically access content outside of regulatory action folder.
			01	Submission folder Root folder for submission 01
			02	Submission folder Root folder for submission 02
			03	Submission folder Root folder for the submission 03.
			index.html	Start file for the GHSTS viewer running in a web browser.
			ghsts.xml	GHSTS XML backbone file of this submission.
			content	Content folder Folder for all computer files that are referenced by a FILE.
			confidential	Confidential folder Folder for all computer files that are marked confidential for at least one RA in the element CBI DESIGNATION.
			utils	Utils folder Folder for additional files relevant for display or validation.
			ghsts_01-00-00.xsd (example)	XSD (XML schema definition) file of the GHSTS XML backbone in the submission folder.
			ghsts-picklists.xsd	XSD file of the GHSTS picklist file that has been used to generated the values in the GHSTS XML backbone file in the submission folder.
			viewer	Viewer folder Folder for the web-based browser. As viewer component can be versioned separately the number of files and subfolders in this folder may vary.
			toc	ToC folder If the GHSTS backbone file references a standard ToC XML file then this file and its Toc XSD file is placed in this folder.
			resources	Resources folder Folder for additional resources needed for display and validation of individual content files, e.g. style sheets, XSD or other assets to display and eventually validate XML summary files in the GHSTS viewer. The folder may contain subfolders, e.g. to group assets by OHT template. Attention: How OHT files can be displayed is currently under review.
			04	Submission folder Root folder for submission 04

Table 98 – overview of folder structure

There is no separate folder for auxiliary information, for example side letters. All content information supplied by the registrant in the submission package must be put into the content or confidential folder and referenced from a FILE instance in the GHSTS XML backbone file. This means that only content can be submitted in a submission package that is referenced from a ToC node.

6.1.1 *Usage of Content folder*

The actual content of the submission, which is the referenced computer files of the FILE instances in the GHSTS XML backbone, is located in the content folder. For file formats see chapter 3.2.14.

Each referenced computer file is contained once in the submission package, even if the document to which it belongs is assigned to multiple ToC nodes / data points at the same time. Each referenced computer file is either stored in the content folder or the confidential folder, but not in both. It is not allowed to use file system functions like links in Windows or Linux to mimic these multiple assigned documents as multiple computer files on the file system.

The content folder may contain subfolders however this is only recommended when the number of computer files in the content root folder becomes too large to be efficiently displayed in a file system browser. It is not recommended to create a copy of the hierarchical ToC structure in the form of folders – as outlined above this would come to its limits for multiple assigned documents anyway. The source for the ToC node / data point information is the GHSTS XML backbone file, not the folder structure below the content folder.

6.2 **Storage of multiple submissions of the same dossier**

As discussed in chapter 3.2.25 the GHSTS XML backbone file of a submission can contain two types of external references that point to predecessor submissions:

- References with PID.
- For incremental submissions: file references to computer files in predecessor submissions (see chapter 4.10.1)
The reference from the GHSTS XML backbone file to the computer files in the main and attachments folders is a relative path that always first goes up one level to the submission folder and then descends to the correct submission folder (see chapter 4.10.1). By this way the structure of the file path is always identical for full and incremental submissions where the content is located in a previous submission.

If the receiver work with a file system-based dossier management system then the submissions must be stored below the same regulatory action folder as outlined in **Table 98** in order that these references can be resolved programmatically by the GHSTS viewer or another retrieval component working on a file system level.

The registrant has to – at least temporarily – store the submission in the same way to be able to check the consistency of the submission with the GHSTS validator that will need to resolve and verify the external references.

The total length of a relative path is restricted to 230 characters. Due to a restriction on some Windows file systems where the total absolute path must not exceed 255 characters it is advisable to create the regulatory action folder not to “deep” below the drive letter on these systems.

6.3 Folder naming conventions and constraints

The following table lists the folder naming conventions and constraints for the folders listed in the figure above

Folder type	Folder naming convention and constraints	Example
Regulatory action folder	Folder is outside of scope of GHSTS submission package and will only exist if the regulatory authority uses a file system-based dossier management system. Recommendation for folder name: Value of DOSSIER_COMP_ID of the DOSSIER element – this value should be identical in all submissions of the regulatory action.	SYNCH127
Submission folder	Folder name is submission number which is padded to two digits. Folder must only include the files and folders as shown in table “Table 98 – overview of folder structure”.	01 02 03
Content folder	Mandatory folder name is “content” The main folder must contain all computer files of FILE instances that are not marked confidential for any RA in the element CBI_DESIGNATION (see chapter 0). May contain subfolders	content
Confidential folder	Mandatory folder name is “confidential” The main folder must contain all computer files of FILE instances that are marked confidential for at least one RA in the element CBI_DESIGNATION (see chapter 0). May contain subfolders	confidential
Utils folder	Mandatory folder name is “utils” May contain subfolders	utils
Viewer folder	Mandatory folder name is “viewer” May contain subfolders	viewer
ToC folder	Mandatory folder name is “toc” May not contain subfolders	toc
Resources folder	Mandatory folder name is “resources” May contain subfolders	resources

Table 99 – folder naming conventions

6.4 File naming conventions

The following table lists the file naming conventions or recommendations that exist for the computer files that are part of a GHSTS submission package.

Purpose of computer file	Parent folder	Naming convention	Example
Start file for the GHSTS viewer	Submission folder	Filename “index.html” is mandatory.	index.html
GHSTS XML backbone file of this submission	Submission folder	Filename “ghsts.xml” is mandatory.	ghsts.xml
GHSTS XSD that describes the structure of the submitted GHSTS XML backbone file	Utils folder	See chapter 4.2, do not alter supplied filename	ghsts_01-00-00.xsd

XSD file of the GHSTS picklist file that has been used to generated the values in the GHSTS XML backbone file in the submission folder.	Utils folder	See chapter 5.2.2, do not alter supplied filename	ghsts-picklists.xsd
XSD file of the Standard ToC file	Toc folder	See chapter 5.1.2, do not alter supplied filename	toc_01-00-00.xsd
XML file of a Standard ToC	Toc folder	See chapter 5.1.2, do not alter supplied filename	toc-oecd_micro_01-00-00.xml
Computer file in content folder or confidential folder (or subfolder)	Content folder, Confidential folder	<p>There is no binding naming convention for files in the content and confidential folders, as long as the file names are unique per folder. Please note that a filename must not contain all possible characters – see chapter 4.10.4 – and illegal characters must be replaced by a legal character, e.g. “_”.</p> <p>Suggestion for naming convention (non-binding):</p> <p>As every computer file belongs to exactly one document a short human readable document identifier like the company study number in the element DOCUMENT_NUMBER could be used. If multiple languages are submitted simultaneously the language code of the study is added. Alternatively the company can use another internal identifier that identifies the document uniquely across different languages.</p> <p>This file name part is then followed by a suffix indicating the file type of the document (“main”, “source”, “supp” and “xml” for file types “Main”, “Source”, “Supplemental” and “XML Summary”. If there are multiple computer files of the same file type a number could be added.</p>	C6313_0016-EN-main.pdf C6313_0016-EN-source.doc C6313_0016-EN-suppl1.bmp C6313_0016-EN-suppl2.bmp C6313_0016-EN-xml.xml

Table 100 – file naming conventions

6.5 Platform and media support

The definition of a GHSTS submission package is not bound to a specific computer platform, file system type or media. A GHSTS submission package can be created and stored on any standard platform, arbitrary media or transport protocols can be used to transport it from the registrant to the receiver(s).

GHSTS does not contain mechanisms to detect alterations of the submission package itself that may arise during its transport. On the content level the MD5 checksum of the FILE content can help to verify the integrity of the files in the content folder.

7 USING GHSTS

This chapter describes in a brief manner, how GHSTS can be adopted and used by registrants.

7.1 Creation and publication of a GHSTS submission

To get ready to create a first GHSTS submission, the registrant has to do the following preparation steps:

- Enable in-house dossier management system to support equivalents of the main concepts of GHSTS:
 9. Document Families, Documents, Files.
 10. Notion of content changes for files and documents as described in chapters 3.2.16 and 3.2.17.
 11. Notion of metadata changes.
 12. Generation and assignment of local PID.
 13. Submissions, being able to supply incremental submission as comparison to previous submissions.
- Incorporate GHSTS picklists from the GMB into in-house system to use for metadata assignment, allow repeated update of values.
- Incorporate ToC structures from the GMB into in-house system to use for dossier creation, allow repeated update of structures.
- Establish query interfaces to the GMB for the search of centrally managed PID (e.g. for substances).
- Create GHSTS publication component as new output format for dossier management system.

Eventually the use of a converter from an existing standard to GHSTS can serve as intermediate and temporary step. However it is very likely that existing standards do not incorporate all mandatory information in a GHSTS submission package.

7.1.1 *Querying the GHSTS Management Body*

All the existing centrally managed GHSTS components and information on how to contact the GHSTS Management Body are available on the OECD website at: **XXXXXX**.

7.2 Conformity testing

When a GHSTS submission is created, the compliance with the GHSTS standard has to be verified using a GHSTS validator.

Attention: A GHSTS validator does not currently exist. Its development is under consideration.

7.3 Transport of a GHSTS submission

A GHSTS Package can be transferred to the authorities on any kind of electronic media or with any kind of electronic protocol depending on the agreement with the regulatory authority. The regulatory authority is responsible to copy the submission into the file system or into their in-house dossier management system.

7.4 Consumption of a GHSTS submission

There are two scenarios for the consumption of a GHSTS submission by the receiver:

- The receiver can use the GHSTS submission package in a “standalone” manner, using the supplied GHSTS viewer to view the information in the backbone file and the file content.
- The receiver uses an in house dossier management system that possesses comparable concepts as within GHSTS.

For the second scenario the receiver will require a specific GHSTS ingestion component. This component will parse the GHSTS XML backbone and compare the supplied information with the data that is already stored internally, for example for PID of files and documents. The ingestion may then be only partially dependent on the supplied information for content and metadata changes.

8 APPENDIX

8.1 Glossary of Terms

The definitions used in this document are for the purpose of the GHSTS only.

Terms	Definitions
ACVM	See “MPIMAF-ACVM”.
Annex Point	A TOC-Node according to the OECD Guidance for Industry Data Submissions on Plant Protection Products and their Active Substances.
Applicant	Company, “leading company” of a shared dossier of multiple companies, or individual that is asking for approval for use and/or marketing a certain active substance and/or pest control product to a regulatory authority.
Application	See “Dossier”.
Authority	A global, regional or national governance body that establishes, monitors, reforms, and enforces regulations; in this context, the regulations are surrounding pesticides (i.e. Regulatory Agency, Regulatory Authority).
BVL	German Federal Office of Consumer Protection and Food Safety Responsible for the authorisation of plant protection products, veterinary drugs and genetically modified organisms.
CADDY	Computer Aided Dossier and Data supply CADDY is an electronic format for the exchange, archiving and evaluation of dossiers, developed jointly by industry and regulatory authorities. Citation: Electronic Submission Expert Group (2013), Computer Aided Dossier and Data Supply website, http://caddy.ecpa.eu (accessed 17 September 2013).
CAS Registry Number	A unique numeric identifier generated by Chemical Abstracts Service designating a substance. The number is a numeric identifier that can contain up to 10 digits, divided by hyphens into three parts (e.g. XXXXX-XX-X). CAS Registry Number is a Registered Trademark of the American Chemical Society.
CBI	Confidential Business Information - Since the distinction of what is CBI and what is not tends to vary from country to country, it was determined that the harmonized schema would need to allow for multiple CBI fields which would indicate the CBI status of a document across multiple regulatory authorities. Further distinctions between the documents can be made using specific Document IDs. Confidential Business Information can generally be described as information (trade secrets or information of commercial value) to concern the following:- manufacturing or quality control processes relating to an active substance or pest control product; or- methods for determining the composition of an active substance or pest control product; or- the identity and concentration of the formulants and contaminants in a pest control product, other than those considered to be of health; or - environmental concern that are identified on a list to be made available to the public.
CFR	US Code of Federal Regulations; legal text containing the United States Environmental Protection Agency pesticide regulations.

Terms	Definitions
Dossier numbering systems (see also: ToC: - Table of Content):	<p>System of codes used to identify the types of documents included within the dossier. These codes can identify letters, forms, data, reviews, and other document types. Examples are shown below. Other numbering systems exist in different regions and authorities:</p> <p>OECD International: OECD - Organisation For Economic Co-operation and Development (OECD) uses the OECD IIA (active substance) /IIIA (pest control product) numbering system</p> <p>PMRA (DACO): Canada - Pest Management Regulatory Agency (PMRA) uses the Data Codes numbering system (DACO)</p> <p>USEPA (OPPTS): United States - Environmental Protection Agency (EPA) uses the OPP & OPPTS system</p>
Decision	An authoritative determination made after consideration of facts from a submission, document, or report.
DER	Data Evaluation Record; a review of data from a specific document or data point used by PMRA and EPA. The DERs are summarized to provide the final review of the data.
Document	The written content of study reports, labels, forms, cover letters, etc. which are created by the registrant. A Document is a primary submission component and may be subject to amendments. A document contains of at least one file.
Document family	A Document Family is a logical binding of all elements based on the original document, for example, a collection of documents and its files that are tied to specific content (for example a study) and/or a document and all of its life cycle elements.
Dossier	The set of documents prepared by the applicant submitted for a regulatory purpose. It is maintained continuously throughout the lifecycle of a regulatory action. A product may be composed of a number of dossiers, such as the initial dossier to register the product, as well as subsequent dossiers to amend the product. A Dossier may also have numerous 'submissions', which are used to submit documents to the Regulatory Authorities either as part of a new dossier or an existing dossier. One Dossier = one Regulatory Action = including maximum one product
Dossier, lifecycle management	The coordination of all versions of submissions, documents, and files throughout the regulatory submission process from submission to approval of a regulatory action.
EC/European Commission	The European Commission is the executive body of the European Union (EU).
ECHA	European Chemicals Agency manages the technical, scientific and administrative aspects of REACH and is also the owner of IUCLID 5 database.
e-Dossier	The e-Dossier is the electronic representation of the dossier.
e-Index Builder	<p>Produced by the Canadian Pest Management Regulatory Agency (PMRA).</p> <p>The e-Index Builder is a stand-alone Java-based software application which facilitates the creation of an electronic index (in XML format) that fully describes each document submitted in support of applications to register or amend a pest control product registration, re-evaluation, special review, or in response to requests for information.</p> <p>Citation: Health Canada (n.d.), Health Canada website, www.hc-sc.gc.ca/cps-spc/pest/registant-titulaire/eprs-serp/_eib-gie/index-eng.php (accessed 17 September 2013).</p>

Terms	Definitions
EPA (New Zealand)	New Zealand's Environmental Protection Authority (EPA) administers applications for major infrastructure projects of national significance, and regulates new organisms (plants, animals, GM organisms) and hazardous substances and chemicals. The EPA also administers the Emissions Trading Scheme and New Zealand Emission Unit Register, and manages the environmental impact of activities in New Zealand's Exclusive Economic Zone (EEZ), including prospecting for petroleum and minerals, seismic surveying and scientific research. Reference: Environmental Protection Authority of New Zealand (n.d.), http://www.epa.govt.nz , (accessed 17 September 2013).
EPA (US)	Environmental Protection Agency is an agency of the United States Federal Government charged to protect human health and environment.
e-Prism	US EPA multi-faceted system for consolidation and management of EPA pesticide information. This system serves as a gateway to electronic submission of pesticide-related information and includes a portal for XML based electronic submissions.
e-PRS	Canada's Electronic Pesticide Regulatory System for consolidation and management of PMRA information. This system serves as a gateway to electronic submission of pesticide-related information and includes a portal for XML based electronic submissions.
e-Submission	The electronic representation of a submission to an authority according to agreed standards including lifecycle management.
European Union (EU)	The economic and political union of 28 member states located in Europe.
File	Files – the physical manifestation of a document's written content (or part of the content) in the form of a currently accepted formats.
GHSTS	OECD Global Harmonised Submission Transport Standard provided to Companies and Regulatory Authorities as the standard for the submission of e-Dossiers to OECD regulatory authorities for the review of pest control products.
GIFAP	International Group of National Associations of Manufacturers of Agrochemical Products (from French: Groupement International des Associations Nationales de Fabricants de Produits Agrochimiques, GIFAP). In 1996, GIFAP was renamed to Global Crop Protection Federation (GCPF). In 2001, GCPF was renamed to CropLife International.
Global Review	A formal process with negotiated timelines for the review of a dossier, where the workload is split among participating regulatory authorities globally, and the reviews of data are exchanged and peer-reviewed.
IUCLID 5	International Uniform Chemical Information Database A software application (No 5 indicates the actual version) to capture, store, maintain and exchange data on intrinsic and hazard properties of chemical substances. It is the key tool for chemical industry to fulfil data submission obligations inside the REACH legislation. The software is maintained by the European Chemicals Agency. Reference: European Chemicals Agency (n.d.), IUCLID website 5, http://iuclid.eu/ (accessed 17 September 2013).

Terms	Definitions
Joint Review	<p>A registration dossier package to be reviewed, either independently or dependently, by two or more or regulatory authorities from two or more different global regions.</p> <p>The evaluation of a pesticide dossier is shared by two or more countries. The participating regulatory authorities review the work of the primary reviewers for each particular science discipline, and the end product (ideally a complete monograph or key components of the monograph) is used by all participating countries (and others) as the basis for regulatory decisions.</p>
MPIMAF-ACVM	Ministry for Primary Industries Forestry - Agricultural Compounds and Veterinary Medicines Responsible for registration of agricultural compounds (including pesticides) and setting of Maximum Residue Limits, along with the development, implementation, evaluation and reviews of standards relating to agricultural compounds.
Metadata	“Data about data”. The information fields or attributes describing with keywords the content of a document or dossier.
Monograph	An evaluation report issued by the authorities.
MRID	<p>Master Record Identification Number</p> <p>A unique cataloguing number assigned to an individual pesticide study at the time of its submission to the US EPA.</p> <p>Citation: United States Environmental Protection Agency (n.d.), Pesticides: Glossary website, www.epa.gov/pesticides/glossary/m-q.html (accessed 17 September 2013).</p>
Multinational Submission	A registration application and associated content which is submitted for review to two or more countries.
Numbering system	See "Dossier numbering system".
OECD	The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation which includes 34 member countries committed to democracy and the market economy. Its principal aim is to promote policies for sustainable economic growth and employment, a rising standard of living and trade liberalisation. By “sustainable economic growth” the OECD means growth that balances economic, social and environmental considerations.
OECD <i>Ad hoc</i> Expert Group on the Electronic Exchange of Pesticide Data	A group of experts from OECD governments, industry and other stakeholders, with expertise in approaches and/or issues associated with IT systems and pesticide registration, established by the OECD to identify and address issues to facilitate the efficient exchange of pesticide information in the pesticide regulatory process. The work of this group is overseen by the OECD Registration Steering Group.
OECD Expert Group on Harmonising Templates	Group of experts from OECD member countries, industry and NGOs, in charge of the development and review of OECD Harmonised Templates for Reporting Test Summaries.
OECD Harmonised Templates	An OECD harmonised template is a standard data format for reporting studies done on chemicals to determine their properties or effects on human health and the environment. They are aimed at developers of database systems, as they prescribe the formats by which information can be entered into and maintained in database. By using these templates, governments and industry will easily be able to electronically exchange test study summary information. The templates can be used for reporting summary results for testing on any type of a chemical (e.g., pesticides, biocides, industrial chemicals).

Terms	Definitions
OECD Registration Steering Group	The OECD Registration Steering Group (RSG) is a subgroup of the OECD Working Group on Pesticides that drives the regulatory initiatives of the OECD Pesticides Programme, in particular harmonization of regulatory processes and requirements as well as cooperation and work sharing between regulatory authorities. The RSG oversees the work of the OECD <i>Ad hoc</i> Expert Group on the Electronic Exchange of Pesticide Data on electronic exchange of pesticide data and on registration.
OECD XML study templates	The XML format of an OECD harmonized template.
Pesticide Package	The collective contents of the pesticide registration application, documents/studies, label(s), and any other materials needed to initiate a regulatory review process.
Pesticide Regulatory Process	The process handling the registration in accordance with the government's pesticide regulation laws.
Pesticide Review Process	The process of assessing the validity and conclusions within a registrant's application package, including the examination of all studies and summaries, as well as the proposed product label.
PMRA	Health Canada's Pest Management Regulatory Agency (PMRA) registers all pesticides that are sold, used, manufactured or imported into Canada under the federal <i>Pest Control Products Act</i> 2002 c.28 (PCP Act)
PR Notice 86-5	The EPA-issued notification describing the standard format for data submitted under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and certain provisions of the Federal Food, Drug, and Cosmetic Act (FFDCA). This notice has now been superseded by PR Notice 2011-3.
PRIA	US Pesticide Registration Improvement Act which defines the review time and fee for service for pesticide reviews.
Primary Reviewer	AKA Lead Reviewer. The evaluator from the regulatory authority assigned the lead for each science discipline who does the initial data review and drafts the primary review (study evaluation).
Project	In the e-dossier system, project is designated by the registrant and has a defined start and end.
PRZ	The file extension assigned to the output file from the use of the PMRA e-Index Builder.
RA	Regulatory Authority
Rapporteur	In the context of dossiers, the EU member state who is the primary reviewer of a submitted registration application and who shares their findings with the remaining EU member states.
REACH	Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. A European Union Regulation which addresses the production and use of chemical substances. Reference: European Union, 1998-2013, EUR-Lex website, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32006R1907:en:NOT (accessed 17 September 2013).
Reference List	The list of documents and studies submitted by a registrant as part of a registration submission package.

Terms	Definitions
Registrant	Company, company representative, or individual who holds the registration for an active substance or pest control product. Also called “Approval Holder”.
Registration	The process to apply, assess and register a pesticide.
Regulation	Legal act
Regulatory action	See “Dossier”.
Regulatory Agency/ Regulatory Authority (RA)	See “Authority”.
Regulatory Process	The process handling the registration in accordance to the regulation.
Review Process	The process of assessing the validity and conclusions within a registrant's application package, including the examination of all studies and summaries, as well as the proposed product label.
Reviewer	A regulatory authority employee or contractor involved in a review process, having the responsibility to ascertain the validity and conclusions of a registrant's application package.
Specification	A set of standards for constructing an electronic dossier, which if compliant, enables the submission of said compliant package to an authority for consumption into a database or other organized system.
Study Report	Is a document which reports the results of data from laboratory or field tests. In most cases it is conducted under good laboratory practice (GLP) conditions.
Test Summary	A Test or Study Summary provides key information about a study in a condensed form. Study summaries may be presented in text files or in a database representation (e.g. OECD Harmonised Templates).
Style sheet	Style sheets are the way that standards-compliant Web designers define the layout, look-and-feel, and design of their pages. They are called Cascading Style Sheets or CSS.
Submission	A compilation of documents prepared by the applicant for a specific regulatory purpose in a structured form according to the given regulatory requirements. Multiple submissions can be submitted for each dossier.
Submission, Version of -	A dossier, submitted with the goal of achieving a positive regulatory decision, can consist of multiple deliveries of documents or data toward that goal. Each delivery, termed a submission, is assigned a version number for tracking purposes.
Submitting Entity	The name of the registrant company that submits the registration application and accompanying data.
Table of Contents	Represents a hierarchy for documents within the dossier. For different legislations different Table of Contents (ToCs) exist.
Template	A pre-developed layout in electronic or paper media .An electronic file with a predesigned, customized format and structure, as for a fax, letter, or expense report, ready to be filled in.
TOC	See "Table of Contents".
TOC Node	One node in the tree structure of a TOC
Transport Mechanism	Systems designed for handling large numbers of documents and for exchanging such documents electronically in a well-defined and standardized framework.

Terms	Definitions
Transport Mechanism Subgroup	Subgroup of the OECD Expert Group on Electronic Exchange of Pesticides Data (EGEEDP).
UUID	Universally Unique Identifier: This identifier enables distributed systems to uniquely identify entities, such as legal entities, products, dossiers, substances, documents, files. For more details see RFC 4122: A Universally Unique Identifier (UUID) URN Namespace or Wikipedia.
Validation	The establishment of documented evidence that a system does what it purports to do, and the processes by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose.
Working Group on Pesticides	Working group that directs and oversees the work of the OECD Pesticide Programme, a forum for national pesticide regulators from OECD member and selected non-member countries to discuss and find solutions to issues of common interest.
Work sharing	A formal process among regulatory agencies to share the work to review a dossier.

8.2 Standards References

The following table contains a list of references to official standards that are mentioned in this specification:

References	Description
ISO 3166-1-alpha-2	see http://www.iso.org/iso/iso-3166-1_decoding_table.html
XML	see http://www.w3.org/XML/
XML schema	see http://www.w3.org/XML/Schema.html#dev
RFC 1321	MD5 Message Digest Algorithm, see http://www.ietf.org/rfc/rfc1321.txt
RFC 2141	URN syntax, see http://tools.ietf.org/html/rfc2141
UUID	UUID syntax see http://www.itu.int/rec/T-REC-X.667-200808-I/en
ISO-19005-1:2004	Document management – Electronic document file format for long-term preservation - Part 1: Use of PDF 1.4 (PDF/A-1), ISO, 2004

Table 101 – standards references

8.3 Copyright and Trademark Information

- The copyright for this document is with the OECD.
- The figures in the text are created using the Open Source software yED (http://www.yworks.com/en/products_yed_about.html).
- The graphical representations of the XSD schemata are created using Altova XML Spy 2013.
- The icon for the Adobe PDF document is a trademark of Adobe Systems Incorporated.

- The icons for Microsoft Word and Microsoft Excel documents are trademarks of Microsoft Cooperation.
- The icons used in the figures come from various sources:
 14. Hewlett Packard Icon Database from the HP Experience Center. Copyright Hewlett Packard, released under license.
 15. Picol Library by Melih Bilgil, see <http://www.picol.org>, licensed under the Creative Commons-License BY-SA (<http://creativecommons.org/licenses/by-sa/3.0/>).
 16. Lab icon from <http://www.wpzoom.com>, licensed under the Creative Commons Attribution-Share Alike 3.0 unported License (<http://creativecommons.org/licenses/by-sa/3.0/>).
 17. Format icons from the icon set of the Regional Computing Centre of Erlangen, based on the Tango desktop project, licensed under the Creative Commons Attribution Share-Alike 3.0 license (<http://creativecommons.org/licenses/by-sa/3.0/>).
 18. Folder and document icon from Momentum Designlabs, see <http://momentumdesignlab.com> licensed under the Creative Commons Attribution-Share Alike 3.0 license (<http://creativecommons.org/licenses/by-sa/3.0/us/>).

8.4 Version history

The following Change History log contains a record of changes made to this document. The last column also contains information whether a new version of this GHSTS specification also requires a change of other GHSTS components (e.g. the GHSTS submission schema definition).

Version #	Date	Author	Section / Nature of change / Affected GHSTS components
01.00.00	2014-xx-xx	OECD	Initial publication of this document.

Table 102 – version history

---- End of document ----