

ISO 24620-5:2024-06 (E)

Language resource management - Controlled human communication (CHC) - Part 5: Lexico-morpho-syntactic principles and methodology for personal data recognition and protection in text

Contents

Page

- Foreword..... iv
- Introduction..... v
- 1 Scope..... 1
- 2 Normative references..... 1
- 3 Terms and definitions..... 1
- 4 Motivation for controlled human communication..... 2
- 5 Basic principles and methodology..... 2
 - 5.1 General..... 2
 - 5.2 Specific issues..... 3
 - 5.3 Principles..... 3
 - 5.3.1 Overview..... 3
 - 5.3.2 Lexical, morphological and syntactic indicants..... 4
- 6 Applications..... 6
 - 6.1 General..... 6
 - 6.2 Different language families..... 6
 - 6.3 Languages and countries..... 6
 - 6.4 Semes in text..... 6
 - 6.5 Applications for personal data recognition..... 6
- Annex A (informative) Examples of text in different languages and different semes..... 7
- Annex B (informative) Examples of hidden text with seme indications..... 13
- Annex C (informative) Table of semes in context..... 15
- Bibliography..... 18