International Standard

ISO 24620-5

First edition
2024-06

Language resource management — Controlled human communication (CHC) —

Part 5: Lexico-morpho-syntactic principles and methodology for personal data recognition and protection in text

Gestion des ressources linguistiques — Communication humaine contrôlée (CHC) —

Partie 5: Principes lexico-morpho-syntaxiques et méthodologie pour la reconnaissance et la protection des données à caractère personnel dans du texte

Reference number
ISO 24620-5:2024(en)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

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**

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 4, *Language resource management.*

A list of all parts in the ISO 24620 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The exchange of personal data between public and private actors, including natural persons, associations and undertakings, is continually increasing. Rapid technological developments and globalization have brought new challenges for the protection of personal data. The scale of the collection and sharing of personal data has increased significantly. Technology allows both private companies and public authorities to make use of personal data on an unprecedented scale in order to pursue their activities. Natural persons increasingly make personal information available publicly and globally. Nevertheless, technology has transformed both the economy and social life, and should further facilitate the free flow of personal data within a country as well as the transfer to and between other countries and international organizations, while ensuring a high level of protection of personal data. These developments require a robust and coherent data protection framework. For example, ISO/IEC 27701 defines processes and provides guidance for protecting personally identifiable information (PII) on an ongoing, ever-evolving basis.

Effective protection of personal data requires the strengthening and setting out in detail of the rights of natural persons as data subjects, and the obligations of those who process and determine the processing of personal data.

EXAMPLE The European Union's (EU) General Data Protection Regulation (GDPR).[6][15]

The principles of data protection apply to any information concerning an identified or identifiable natural person.

In this context, numerous industries, governmental bodies, and private and public companies or organizations need to variously hide (mask)[16], remove, anonymize or pseudonymize personal data before text containing such data is processed.[4][8]

This document provides principles and a methodology to detect and identify personal data so that it can be hidden or suppressed, i.e. protected before transmitting and/or processing a text containing such data. The problem is not so much the suppression or hiding of data, but rather the recognition of personal data in a written text. Unlike personal data in text, personal data in structured data (e.g. as presented in tables) does not represent a real problem as such data are easily recognizable.[5]

This document is aimed at national and international micro, small, medium and large enterprises, as well as private/public bodies processing text which can contain personal data in all domains (e.g. law, finance, health) and languages and from different countries.[14] The principles and methodology are already in use in industry and government bodies.

Due to regulations such as the EU's GDPR, personal data protection presents a major challenge for micro, small, medium and large enterprises, as well as private and public bodies. For example, the GDPR forbids the transfer of the personal data of EU data subjects to "third countries" (countries outside of the European Economic Area (EEA)) unless appropriate safeguards are imposed, or the third country's data protection regulations are formally considered adequate by the European Commission. In addition, the state of California in the United States passed the California Consumer Privacy Act on 28 June 2018, taking effect 1 January 2020, granting rights to transparency and control over the collection of personal information by companies in a similar manner to the GDPR (see Reference [2] and ISO/IEC 27701).

All the examples in this document are fictitious but could exist if real data were to be substituted for the fictitious data.

Language resource management — Controlled human communication (CHC) —

Part 5: Lexico-morpho-syntactic principles and methodology for personal data recognition and protection in text

1 Scope

This document establishes basic principles and a methodology to recognize personal data written in free text, in different languages (whether agglutinating, inflectional or isolating) and countries.

This document is applicable to protecting human data circulating in national and international industries, and private and public organizations.

This document is applicable to processing by human beings and/or automated processing, and to various domains (e.g. law, finance, health).

It does not apply to automated image processing.

This document uses formal methods only, as statistical methods are very different in nature.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the terms and definitions given in the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
intension
internal content of a term or concept that constitutes its formal definition

Note 1 to entry: Extension is the range of applicability of a concept by naming the particular objects that it denotes.

3.2
personal data
any information relating to an identified or *identifiable natural person* (3.6)

[SOURCE: Regulation (EU) 2016/679[6], Article 4 (1)]

3.3
pseudonymization
processing (3.4) of *personal data* (3.2) in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or *identifiable natural person* (3.6)

[SOURCE: Regulation (EU) 2016/679[6], Article 4 (5)]

3.4
processing
any operation or set of operations which is performed on *personal data* (3.2) or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction

[SOURCE: Regulation (EU) 2016/679[6], Article 4 (2)]

3.5
seme
Saussure's signified with its different signifiers (instantiations) in text

Note 1 to entry: Saussure was the first person to use the terminology "signified" and "signifier". Saussure offered a "dyadic" or two-part model of the sign. He defined a sign as being composed of a "signifier" (signifiant) and a "signified" (signifié) (see References [17] and [18]).

3.6
identifiable natural person
data subject
person who can be identified, directly or indirectly, in particular by reference to an identifier

Note 1 to entry: An identifier can be a name, an identification number, location data or an online identifier of a natural person. Further examples which are excluded from the examples in this document are references to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of the natural person.

[SOURCE: Regulation (EU) 2016/679[6], Article 4 (1)]

3.7
indicant
significant occurrence of interaction between lexical, morphological and syntactic phenomena or of one of these phenomena across a wide spectrum of languages or in few languages or in just one language that is suited to identify *personal data* (3.2)

4 Motivation for controlled human communication

The first step in protecting personal data is being able to recognize such data automatically, especially when they are not structured but rather occur in free text, as shown in Example 1 in Clause A.1.

Once data are detected or recognized as personal data, different ways can be used to hide them in the text: they can be hidden (masked), removed, anonymized (see References [9] and [10]) or pseudonymized (see Reference [7]), as shown in Example 2 in Clause A.2.

Examples 3 and 4 in Clauses A.3 and A.4 show a similar example in French.

5 Basic principles and methodology

5.1 General

For the basic principles, various lexical, morphological and syntactic linguistic phenomena shall be used, in particular concerning the way in which personal data are represented in free text. For example, addresses

not respecting the English format in the UK as seen in Example 1 in Clause A.1, i.e. “Stoneham-le-Willows at 24 Britanny Park, F2 7AN (GB29 NWBK 6016 1331 9268 19)”.

The methodology specifies formal representations designed in intension (see References [11], [12] and [13]) based on lexical, morphological and syntactic phenomena that shall apply in a sequential order at each of the levels of linguistic analysis which have an impact on the recognition of personal data.

The basic principles and methodology are specifically formulated to provide an explanatory power to show how and when each of the linguistic phenomena (lexical, morphological and/or syntactic) and/or their combinations and interactions embedded in context shall be used and applied according to different semes recognition in the analysis of text. In consequence, the methodology uses linguistics phenomena conforming to basic principles for the recognition of personal data (instead of a lexicon), and specifies a system of constraint rules completed with an algorithm, which, when applied, results in extracting personal data.

5.2 Specific issues

The basic problem is the recognition of personal data in free text in different languages, from different countries and from different domains (e.g. law, finance, health).

The problem also concerns the use of the same language within different countries as, for example, addresses are not written the same way in France, Switzerland, Belgium and Canada, or in Austria and Germany (see Examples 5 and 6 in Clauses A.5 and A.6). Text in one language can also include personal data in other languages and from different countries.

The problem is thus fourfold:

a) personal data in free text;

b) personal data in different countries;

c) personal data in different languages;

d) personal data in different domains.

The methodology described in this document uses indicants to detect personal data such as a telephone number, identification or bank account number, an address, etc.

The methodology works in intension. For this reason, it is rule-based: it establishes a system of ordered constraint rules based on linguistic phenomena, which conform to the basic principles that shall be followed (see 5.3). Indicants shall be established, and they shall be lexical, morphological or/and syntactic.

5.3 Principles

5.3.1 Overview

The principles addressing the four specific issues listed in 5.2 formulate rules for an explanatory linguistic formal representation with its own meta-language and grammar accompanied with examples so that new semes and new languages can be processed.

The user shall establish lexico-morpho-syntactic indicants. The indicants, if they do not already exist in the grammar of the meta-language, shall be added.

5.3.2 Lexical, morphological and syntactic indicants

5.3.2.1 General

Some indicants shall be sets valid for different semes in different languages and different countries such as N or A which very often appear in the text but can also be optional. They shall be formally represented as follows in the meta-language grammar (extract):

— N = digit (0...9)

— L = A,B...Z (without diacritic signs)

— A = 0,1...9,L (L signifies an element of L above)

— listA = 01,02,03,04,05,06,07,08,09,10,11,12

— ListB = ListA, 13...100

— () = element(s) between parentheses optional

— M = any word starting with an uppercase or lowercase letter followed by zero or more uppercase or lowercase letters as well as ‘’°-

— UM = M starting with a majuscule letter

— n = new line followed or not by space(s) and/or tab(s)

— listC = 1...9999

— l = a,b...z (without diacritic signs)

— |any character(s) with character | coded as || terminated by| = literal.

NOTE Sequence of literals are illegal (ambiguity of ||).

Some indicants shall be added according to the seme to be recognized, or to the country or language used (see the examples in Annex A).

Example 3 in Clause A.3 contains the following addresses written in different formats:

— 1, rue des échelles

25620 Besançon

— 43 Bd du 11 novembre 1918 à Dijon dans la Côte d’Or 21280

— Saint-Germain-en-Laye, au n° 25 rue du Château d’If 78100

The algorithmic formal representation (diagram representing the algorithm) to recognize any addresses shall be:

semeAddress = (UM( )(UM)(,)( )(M)( )(M)( ))(listC)( )(listE)(,)( )listL( )(M)( )(M)( )(n)(listNC)( )(listC)( )(listmois)( )(listC)( )(M)( )(M)( )(M)( )(M)( )(M)( )(M)( )(M)( )(M)(,)(n)( )listB( )NNN( )(UM)( )(UM)

Using elements (indicants) of the general meta-language grammar and the following complementary indicants:

— listE = bis,ter,Bis,Ter

— listL = rue,avenue,allée,impasse,place,chemin,square,boulevard,Bd,ruelle

— listNC = 1ER,1er,1ERE,1ère,2ème,2EME,3...999|ème|

— listmois = mai,juillet,novembre

5.3.2.2 Lexical indicants

In Example 3 in Clause A.3, the complementary indicants for “semeAddress” are lexical.

5.3.2.3 Morphological indicants

See Example 7 in Clause A.7.

A suffix 都(morphological indicant) shall be used; it indicates a department, as in the two following examples of Japanese addresses:

— 212-0008 神奈川県真田郡ニセコ町43番地 一丁目2番3-405

— 252-1234東京都渋谷区三軒茶屋211-2朝日ハイツ244

5.3.2.4 Syntactic indicants

In Example 1 in Clause A.1, the following English (UK) addresses are syntactically represented in different manners:

— 7b Brians Close,

Upper Avon FY7L 2PQ

— Stoneham-le-Willows at 24 Britanny Park, F2 7AN (GB29 NWBK 6016 1331 9268 19)

In Example 3 in Clause A.3, the following French (France) telephone numbers are syntactically represented in different manners:

— 0671425324

— +33 67 142 53 24

A syntactic algorithmic formal representation shall include all syntactic possibilities for writing a seme such as, for example:

— English (UK) address:

semeAddress = (UM( )(UM)(,)( )(M)( )(M)( ))(listC)(l)(,)( )(UM)( )(UM)( )(UM)
( )(listLUK)(,)( )(n)(UM)( )(UM)( )(UM)(,)( )(n)(UM)( )(UM)(,)( )(n)listUKpostcode

grammar (complementary indicants)
listLUK =
Alley,Avenue,Bay,Circus,Close,Common,Court,Crescent,Drive,Gardens,Gate,Green,Grove,
Hill,Lane,Mall,Mead,Mews,Orchard,Parade,Park,Pass,Piece,Place,Rise,Road,Row,Square,
Strand,Street,Terrace,Vale,Wall,Way,Wharf
listUKpostcode = LN| |NLL,LNN| |NLL,LLN| |NLL,LLNN| |NLL,LNL| |NLL,LLNL| |NLL

— French (France) telephone number:

semeTelephone = <|listT( )N( )NN( )NN( )NN( )NN|>

grammar (complementary indicants)
<| = N followed or not by space illegal before telephone
|> = space or not followed by N illegal after telephone
listT = +NN|(0)|,+NN,00NN|(0)|,00NN,0
+ = used in writing international phone numbers

6 Applications

6.1 General

The principles and methodology are valid for different languages (see 6.2) in different countries (see 6.3) and for different semes (see 6.4).

6.2 Different language families

The principles and methodology are valid for different languages belonging to different families:

— agglutinating, such as Finnish;

— inflectional, such as French;

— isolating, such as Chinese.

6.3 Languages and countries

The following language examples are given in this document (see Annex A):

— English (UK);

— French (France);

— German (Austria and Germany);

— Japanese (Japan);

— Chinese (China);

— Finnish (Finland).

The methodology is not limited to these languages, which have been chosen as they represent different families.

6.4 Semes in text

The following examples of processed personal data detected and hidden are given in this document (see Annex B):

— addresses represented by ADD;

— telephone numbers represented by TEL;

— bank account numbers represented by BAN;

— identification numbers represented by IDN.

The semes are not limited to these examples.

6.5 Applications for personal data recognition

The principles and methodology are usable for data protection, as shown in Annexes A and B.

This document has other different applications, such as structuring personal data found in free text (see Annex C).

Annex A
(informative)

Examples of text in different languages and different semes

A.1 Example 1: English (UK) text containing personal data

27, Milton Road,
Fredlington,
North Yorkshire PJ7 4HT
12 December 2019

John Alfred St John Esq.
7b Brians Close,
Upper Avon FY7L 2PQ

Dear John,
We are pleased to invite you to our daughter's wedding which will take place at Wexham Register Office at 3 pm on 15 July 2020. If you wish to give a present to the future happy couple, either send your gift to my brother, now living in Stoneham-le-Willows at 24 Britanny Park, F2 7AN (GB29 NWBK 6016 1331 9268 19), or directly to NO93 8601 1117 947. In case you need more information, you can contact us at 0121 496 0998.
On another matter, as you may know, my sister is at the hospital. As a doctor you may access her medical data with her NHS number 485 777 3456. In respect of any emoluments, these will be sourced from her old age pension (her NINO being ZP 24 47 29 B), to be debited from LC55 HEMM 0001 0001 0012 0012 0002 3015.
Very sincerely yours,
Anon Nona Other

A.2 Example 2: English (UK) text with personal data hidden

****

12 December 2019

John Alfred St John Esq.
****

Dear John,
We are pleased to invite you to our daughter's wedding which will take place at Wexham Register Office at 3 pm on 15 July 2020. If you wish to give a present to the future happy couple, either send your gift to my brother, now living in ********** (**********), or directly to **********. In case you need more information, you can contact us at**********.
On another matter, as you may know, my sister is at the hospital. As a doctor you may access her medical data with her NHS number **********. In respect of any emoluments, these will be sourced from her old age pension (her NINO being **********), to be debited from **********.
Very sincerely yours,
Anon Nona Other

A.3 Example 3: French (France) text containing personal data

Jean-François Dupont d'Alembert de la Pauserie
1, rue des échelles
25620 Besançon

Cher Monsieur Duconte,
Je vous écris de la part de Madame la Maire de Fougerolles suite à l'accident de mon fils Hervé.
Orange a déjà été contacté par Mme Le Guevel qui habite au 43 Bd du 11 novembre 1918 à Dijon dans la Côte d'Or 21280.
Je vous précise le n° de ss de mon fils: 183102582314552.
Afin que vous puissiez verser l'indemnité et les frais d'hôpitaux, veuillez noter mon compte: FR14 2004 1010 0505 0001 3M02 606.
Je vous laisse aussi un numéro pour me contacter: 0671425324 ou bien vous pouvez me joindre, en cas d'urgence ici +33 67 142 53 24. Le médecin traitant est à Saint-Germain-en-Laye, au n° 25 rue du Château d'If 78100. Mon fils est contactable au hervedupond@orange.fr
Besançon, le 12 décembre 2019

A.4 Example 4: French (France) with personal data hidden

Jean-François Dupont d'Alembert de la Pauserie
****

Cher Monsieur Duconte,
Je vous écris de la part de Madame la Maire de Fougerolles suite à l'accident de mon fils Hervé.
Orange a déjà été contacté par Mme Le Guevel qui habite au **********.
Je vous précise le n° de ss de mon fils: **********.
Afin que vous puissiez verser l'indemnité et les frais d'hôpitaux, veuillez noter mon compte: **********.
Je vous laisse aussi un numéro pour me contacter:**********ou bien vous pouvez me joindre, en cas d'urgence ici**********Le médecin traitant est à **********.
Mon fils est contactable au hervedupond@orange.fr
Besançon, le 12 décembre 2019

A.5 Example 5: German (Austria) text containing personal data

This example shows why traceability is important. In German (Austria), the complementary indicants are deliberately not provided, e.g. for the seme telephone for Austria (see Table C.1). The trace indicates that one telephone number is not recognized as some () are missing in the telephone formal representation for this country.

Herr Dominik Hornung
Raxstrasse 32/12
1100 Wien

Sehr geehrter Herr Hornung,
Ich schreibe Ihnen im Auftrag von Herrn Bürgermeister von Grub in Folge des Unfalls meines Sohnes, Ludwig Beethoven.
Die Versicherung wurde durch Frau Mag. Kratochwill, wohnhaft Hauptstrasse 72, in Sulz – Wienerwald, verständigt.
Die Versicherungsnummer meines Sohnes lautet: 4029.
Damit die Krankenversicherung die Kosten für den Spitalsaufenthalt zurückzahlen kann, hier meine Kontodaten:
IBAN: AT25 110067643485
BIC: BKAUAWWT
Meine persönliche Telefonnummer: +43 431 718 2356.
Im Notfall können Sie mich auch unter der Nummer 0043 68187235698 erreichen.
Der behandelnde Arzt hat seine Ordination Badgasse 16, 2412 Mödling, NÖ.
Mein Sohn ist zu kontaktieren an folgende Email Adresse: ludwig.beethoven@gmx.at
Mödling, am 1. März 2020,

A.6 Example 6: German (Germany) text containing personal data

This example also shows why traceability is important. In German (Germany), the fictitious IBAN, not fully recognized (see Table C.1) in the text, is impossible. The trace detects that there is one digit too many (see also Reference [13]).

Frau Dr. Ines Schlaumeier
Clara-Zetkin Straße 23
D - 04092 LEIPZIG

Betr.: Unfallversicherung

Sehr geehrte Frau Kusche,

Ich schreibe Ihnen im Auftrag der Stadtverwaltung Hinterfinsterwalde bzgl. des Unfalls meines Sohnes Otto.

Der Versicherer ist bereits durch Herrn von Bismarck, wohnhaft Königsstr. 17 in 56130 Bad Ems benachrichtigt worden.

Anbei finden Sie die Kopie der elektronischen Gesundheitskarte mit folgender Krankenversichertennummer: A25789135.

Zwecks Auszahlung der Rückerstattungskosten des Krankenhausaufenthalts teile ich Ihnen meine Bankverbindung mit:

IBAN: DE 45 100000000 1234567890.

BIC: GENOVA44DIE

Für Nachfragen stehe ich Ihnen unter folgender Telefonnummer zur Verfügung:

0049 – (0)2605-2565

Meine Email-Adresse lautet elfriede.mustermann@web.de

Der behandelnde Arzt hat seine Praxis in: Nikolaihof 5a, 25578 HINTERTUPFINGEN

Mit freundlichen Grüßen,

Heinrich Schütz

A.7 Example 7: Japanese (Japan) text containing personal data

2019年12月12日

山田 • ボーズリー • 藤佐様

息子雄二の事故につきまして、但野市長に代わりお手紙をさしあげます。

オランジュに対してはすでに、212-0008　神奈川県真田郡ニセコ町43番地　一丁目2番3-405号に居住の四十万一二三様から連絡を入れております。

息子の健康保険証番号は、06 12345678 です。

損害賠償金と治療費を入金していただく私の口座は次のとおりです。

ゆうちょ銀行　五八八　588　普　123451

また私の連絡先は、0123-45-6789 です。緊急時には 080 1234 5678 にご連絡ください。主治医の住所は、787-1000 さいたま市アトサヌプリ原野25です。息子の連絡先は hervedupond@orange.jp です。

四月朔日しおり

252-1234　東京都渋谷区三軒茶屋211-2

朝日ハイツ244

A.8 Example 8: Chinese (China)

陈襄理您好:

非常感谢您对我女儿车祸事故的关心与处理。

车祸肇事地点在台北东区太平洋SOGO百货正对面**********,时间九月28日大约下午5点30分。

小女的保单号码为854772CHEN 8615。

我已于昨天去电与李正浩专员询问理赔相关事宜,麻烦请将退款汇到:华信银行006账号0561254086911

如果您需要了解小女伤势状况,这是她就诊医生的电话:李文祥医师 02-25781491

如果还有任何问题,您可与我**********信huifang33@gmail.com

这是我女儿的联络方**********此

****

A.9 Example 9: Finnish (Finland)

Jaana ja Juho Meikäläinen

****

Olette sydämellisesti tervetulleet meidän häihimme Kiimingin kirkkoon lauantaina 13.6.2020 klo 13.

Tilaisuus alkaa vihkimistilaisuudella klo 13 osoitteessa Kirkonniementie 8 ja juhlatilaisuus alkaa klo 14 seurakuntatalolla osoitteessa Kirkkotie 26. Olette lämpimästi tervetulleita molempiin tilaisuuksiin.

Olemme koonneet toivelahjalistan amazonin sivulle amazon.fi. Otamme myös vastaan lahjarahaa, sillä suuntaamme Cannesin lämpöön häämatkalle. Lahjarahan voi lähettää suoraan Jaanan tilille: **********Olette myös tervetulleita jatkoille kotiimme, mutta ilmoittakaa mielellään osallistumisesta etukäteen Tepon numeroon 045-745576. Voitte myös ilmoittaa osallistumisesta ja erityisruokavaliosta osoitteeseen haat2020@jippii.fi.

Tervetuloa häihimme avecin kanssa ! Tuokaa vain hymynne ja jakakaa ilomme !

Jaana ja Juho

****

Annex B
(informative)

Examples of hidden text with seme indications

B.1 Example 1: English (UK)

[ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][

12 December 2019

John Alfred St John Esq.

[ADD][ADD][ADD][ADD][ADD][ADD][ADD][

Dear John,

We are pleased to invite you to our daughter's wedding which will take place at Wexham Register Office at 3 pm 15 July 2020. If you wish to give a present to the future happy couple, either send your gift to my brother, now living in [ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][A ([BAN][BAN][BAN][BAN][BAN][B), or directly to [BAN][BAN][BAN][BA. In case you need more information, you can contact us at[T[IDN][IDN][I][

On another matter, as you may know, my sister is at the hospital. As a doctor you may access her medical data with her NHS number [IDN][IDN][I. In respect of any emoluments, these will be sourced from her old age pension (her NINO being [IDN][IDN][ID), to be debited from [BAN][BAN][BAN][BAN][BAN][BAN][BAN][BAN.

Very sincerely yours,

Anon Nona Other

B.2 Example 2: French (France)

Jean-François Dupont d'Alembert de la Pauserie

[ADD][ADD][ADD][ADD][ADD][ADD][ADD

Cher Monsieur Duconte,

Je vous écris de la part de Madame la Maire de Fougerolles suite à l'accident de mon fils Hervé.

Orange a déjà été contacté par Mme Le Guevel qui habite au [ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][A.

Je vous précise le n° de ss de mon fils: [IDN][IDN][IDN].

Afin que vous puissiez verser l'indemnité et les frais d'hôpitaux, veuillez noter mon compte: [BAN][BAN][BAN][BAN][BAN][BAN][BA.

Je vous laisse aussi un numéro pour me contacter:[TEL][TEL][Tou bien vous pouvez me joindre, en cas d'urgence ici[TEL][TEL][TEL][TELLe médecin traitant est à [ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][ADD][A. Mon fils est contactable au hervedupond@orange.fr

Besançon, le 12 décembre 2019

B.3 Example 3: Japanese (Japan)

2019年12月12日

山田 • ボーズリー • 藤佐様

息子雄二の事故につきまして、但野市長に代わりお手紙をさしあげます。

オランジュに対してはすでに、[ADD][ADD][ADD][ADD][ADD][ADD][ADD][号に居住の四十万一二三様から連絡を入れております。

息子の健康保険証番号[IDN][IDN][IDNです。

損害賠償金と治療費を入金していただく私の口座は次のとおりです。

ゆうちょ銀行 [BAN][BAN][BAN][

また私の連絡先は、0123-45-6789 です。緊急時には 080 1234 5678 にご連絡ください。主治医の住所は、[ADD][ADD][ADD][ADD][ADDです。息子の連絡先は hervedupond@orange.jp です。

四月朔日しおり

[ADD][ADD][ADD][ADD][ADD][ADD][AD