
Norme internationale



5963

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Documentation — Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation

Documentation — Methods for examining documents, determining their subjects, and selecting indexing terms

Première édition — 1985-12-01

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO 5963:1985](https://standards.iteh.ai/catalog/standards/sist/e5dcd476-0a03-4ed6-a7ab-e82646c3a1ac/iso-5963-1985)

<https://standards.iteh.ai/catalog/standards/sist/e5dcd476-0a03-4ed6-a7ab-e82646c3a1ac/iso-5963-1985>

CDU 001.815

Réf. n° : ISO 5963-1985 (F)

Descripteurs : documentation, indexation.

Avant-propos

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique créé à cet effet. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO, participent également aux travaux.

Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour approbation, avant leur acceptation comme Normes internationales par le Conseil de l'ISO. Les Normes internationales sont approuvées conformément aux procédures de l'ISO qui requièrent l'approbation de 75 % au moins des comités membres votants.

La Norme internationale ISO 5963 a été élaborée par le comité technique ISO/TC 46, *Documentation*.

[ISO 5963:1985](https://standards.iteh.ai/catalog/standards/sist/e5ded176-0a03-4ed6-a7ab-cf24442a1e0c/iso-5963-1985)

L'attention des utilisateurs est attirée sur le fait que toutes les Normes internationales sont de temps en temps soumises à révision et que toute référence faite à une autre Norme internationale dans le présent document implique qu'il s'agit, sauf indication contraire, de la dernière édition.

Sommaire

Page

1	Objet et domaine d'application	1
2	Références	1
3	Définitions	1
4	Processus de l'indexation	2
5	Examen du document	2
6	Identification et sélection des notions	2
7	Choix des termes d'indexation	3
8	Contrôle de la qualité	4
	https://standards.iteh.ai/catalog/standards/sist/e5dcd476-0a03-4ed6-a7ab-16c9a0b00549 ISO 5963:1985	
	Annexe Organigramme du processus d'indexation utilisant un thésaurus	5

Page blanche

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 5963:1985

<https://standards.iteh.ai/catalog/standards/sist/e5dcd476-0a03-4ed6-a7ab-e82646c3a1ac/iso-5963-1985>

Documentation — Méthodes pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation

1 Objet et domaine d'application

1.1 La présente Norme internationale recommande des procédés pour l'analyse des documents, la détermination de leur contenu et la sélection des termes d'indexation. Son objet se limite à ces premières étapes de l'indexation et se veut indépendant des pratiques de tout système d'indexation, qu'il soit pré-coordonné ou postcoordonné. Elle décrit des méthodes générales pour l'analyse des documents qui doivent s'appliquer quelle que soit la situation. Cependant, ces méthodes sont particulièrement destinées aux systèmes d'indexation dans lesquels on exprime le sujet des documents sous forme abrégée, à l'aide des termes d'un langage d'indexation contrôlé représentant les notions voulues. Dans ce contexte, le langage contrôlé est généralement constitué d'un sous-ensemble de termes tirés de la langue naturelle et, par exemple, organisés en thésaurus. On peut cependant appliquer ces méthodes à des systèmes dans lesquels les notions sont représentées à des fins de recherche documentaire, par des symboles choisis dans un plan de classement.

1.2 Les techniques décrites dans la présente Norme internationale peuvent être employées dans tout organisme où l'analyse des documents et l'expression du contenu en termes d'indexation sont faites par des indexeurs. Elles ne s'appliquent pas aux organismes qui utilisent des techniques d'indexation automatique dans lesquelles les termes d'un texte sont organisés en ensembles ou classes selon des critères qui peuvent être appliqués par ordinateur, par exemple la fréquence d'apparition et/ou l'adjacence dans le texte, bien que la finalité de ces systèmes soit la même.

1.3 La présente Norme internationale doit, en premier lieu, servir de guide aux indexeurs pour les étapes de l'analyse des documents et de l'identification des notions. On peut aussi l'appliquer à la recherche documentaire en transformant les questions des utilisateurs en termes d'indexation contrôlés. Ce peut être aussi un guide pour l'élaboration de résumés analytiques mais, cependant, on doit attirer l'attention sur le fait que ces tâches, quoique analogues, ne sont pas identiques.

1.4 La présente Norme internationale est destinée à promouvoir l'utilisation d'une pratique normalisée

- a) dans un organisme ou un réseau d'organismes;
- b) entre différents services d'indexation, en particulier ceux qui échangent des informations bibliographiques.

2 Références

ISO 2788, *Documentation — Principes directeurs pour l'établissement et le développement de thésaurus monolingues.*

ISO 5964, *Documentation — Principes directeurs pour l'établissement et le développement de thésaurus multilingues.*

3 Définitions

Dans le cadre de la présente Norme internationale, les définitions suivantes sont applicables:

3.1 document: Tout article, imprimé ou non, qu'il est possible de cataloguer ou d'indexer.

NOTE — Cette définition désigne non seulement des matériaux écrits et imprimés sur papier ou microforme (par exemple les livres, les revues spécialisées, les cartes, diagrammes, etc.), mais aussi des médias non imprimés (par exemple les enregistrements lisibles par machine, films et enregistrements sonores) et les objets de collection.

3.2 notion: Toute unité de pensée. Le contenu sémantique d'une notion peut être ré-exprimé en combinant d'autres notions qui peuvent être différentes d'une langue à l'autre.

3.3 sujet: Toute notion ou combinaison de notions représentant un thème dans un document.

3.4 terme d'indexation: La représentation d'une notion sous forme

- soit d'un terme dérivé de la langue naturelle, de préférence substantif simple ou composé;
- soit un code de classement.

NOTE — Un terme d'indexation peut comprendre plus d'un mot. Dans un langage d'indexation contrôlé, un terme est appelé soit *descripteur*, soit *non descripteur*.

3.5 descripteur: Terme toujours employé à l'indexation pour représenter une notion donnée; parfois appelé terme préférentiel.

3.6 non descripteur: Synonyme ou quasi synonyme d'un descripteur. On ne peut l'attribuer à un document mais il sert de point d'entrée dans un thésaurus, l'utilisateur étant alors dirigé par une instruction (par exemple VOIR ou EMPLOYER) vers le descripteur approprié; parfois appelé terme non préférentiel.

3.7 index: Liste alphabétique et systématique des sujets, qui renvoie à la place de chaque sujet dans un document ou une collection de documents.

3.8 indexation: Action de décrire ou d'identifier un document en ce qui concerne son contenu.

4 Processus de l'indexation

4.1 L'indexation ne concerne pas la description d'un document en tant qu'entité physique (par exemple, elle ne donne pas la forme, l'éditeur, la date, etc.). Ces éléments peuvent quand même être inclus dans un index matières si ces informations permettent à un utilisateur de déterminer plus précisément si un document donné est ou non une réponse à sa question.

4.2 Lors de l'indexation, on extrait les notions des documents par un processus d'analyse intellectuelle, puis on les transcrit à l'aide d'éléments du langage d'indexation. L'analyse et la transcription doivent toutes deux être faites avec l'aide d'outils d'indexation tels que thésaurus et plans de classement.

4.3 L'indexation se décompose essentiellement en trois étapes qui ont tendance à se chevaucher dans la pratique:

- a) examen du document et définition de son contenu;
- b) identification et sélection des notions principales du contenu;
- c) choix des termes d'indexation.

Chacune de ces étapes, ainsi qu'un chapitre sur le contrôle de la qualité, sont traités ci-après dans les chapitres 5 à 8.

5 Examen du document

5.1 La précision avec laquelle un document peut être examiné dépend dans une large mesure de sa forme physique. On peut distinguer deux cas: les documents imprimés et les documents non imprimés.

5.2 Les documents imprimés constituent le fonds habituel des bibliothèques et des centres de documentation avec monographies, périodiques, rapports, comptes rendus de congrès, etc. La compréhension totale de ces documents passe en principe par une lecture minutieuse des textes. Une lecture complète n'est pas toujours nécessaire, cependant l'indexeur doit s'assurer qu'aucune information utile n'a été négligée. Les parties importantes du texte doivent être étudiées avec soin et une attention particulière doit être accordée aux éléments suivants:

- a) le titre;
- b) le résumé, s'il existe;
- c) le sommaire;
- d) l'introduction, les débuts de chapitres et de paragraphes, et la conclusion;
- e) les illustrations, diagrammes, tableaux et leur légende;
- f) les mots ou groupes de mots soulignés ou mis en valeur par une typographie différente.

Tous les éléments doivent être examinés soigneusement et recensés par l'indexeur lors de l'étude du document. L'indexation à partir du seul titre n'est pas recommandée et un résumé, s'il existe, ne doit pas être considéré comme un substitut à l'examen du texte. En effet, certains titres peuvent induire en erreur; certains résumés sont insuffisants, ni les uns ni les autres ne constituent systématiquement une source sûre pour le type d'information dont un indexeur a besoin.

5.3 Les documents non imprimés, tels que les documents audio-visuels, visuels ou sonores y compris les objets, requièrent des procédures différentes. Il n'est pas toujours possible, en pratique, d'analyser un enregistrement dans sa totalité (par exemple en projetant un film). L'indexation est alors habituellement tirée du titre ou du résumé, tout en conservant à l'indexeur la possibilité de regarder ou d'écouter le document si la description écrite est insuffisante ou apparaît inexacte.

6 Identification et sélection des notions

6.1 Après l'examen du document, l'indexeur doit identifier les notions, éléments essentiels de la description de son contenu. Les organismes doivent établir une grille d'identification à l'aide de critères reconnus comme importants dans le domaine couvert par l'indexation.

Les questions suivantes donnent des exemples de critères à retenir:

- a) le document traite-t-il d'un objet affecté par une activité?
- b) Le sujet contient-il une notion active (par exemple une action, une opération, un processus, etc.)?
- c) L'objet est-il affecté par l'activité identifiée?
- d) Le document traite-t-il de l'agent de cette action?
- e) A-t-il recours à des moyens particuliers pour accomplir l'action (par exemple instruments spéciaux, techniques ou méthodes spéciales)?
- f) Ces facteurs étaient-ils placés dans le contexte d'un lieu ou d'un environnement particulier?

g) Des variables dépendantes ou indépendantes sont-elles identifiées?

h) Le sujet a-t-il été traité d'un point de vue particulier que l'on n'associe pas normalement à ce domaine d'étude (par exemple une étude sociologique de la religion)?

Ce sont des exemples des critères susceptibles de s'appliquer à beaucoup de domaines. Il peut être nécessaire de formuler d'autres questions pour établir une grille d'indexation dans des disciplines particulières.

6.2 L'indexeur ne doit pas nécessairement retenir, comme termes d'indexation, tous les concepts recensés lors de l'examen du document. Le choix de ces concepts dépend en partie de l'utilisation des termes d'indexation qui peut varier de la production d'index alphabétiques imprimés à la constitution d'une banque de données bibliographiques informatisée. L'identification des notions peut aussi être affectée (comme noté ci-dessus) par le document indexé. Par exemple, l'indexation de livres est susceptible de différer de celle d'articles de périodiques. Les deux caractéristiques d'une indexation les plus susceptibles d'être ainsi modifiées sont l'exhaustivité et la spécificité.

6.3 L'exhaustivité est liée au nombre de notions prises en compte et qui caractérisent le contenu intégral du document.

6.3.1 Un indexeur qui suit les procédures indiquées ci-dessus doit pouvoir identifier toutes les notions qui ont une valeur potentielle pour les utilisateurs d'un système d'information dans un document. Dans certains cas, deux thèmes ou plus dans le domaine couvert par l'indexation surviennent indépendamment l'un de l'autre dans le même document. Ils doivent être traités séparément et, si nécessaire, par différents spécialistes du sujet.

6.3.2 La couverture de l'indexation ne doit pas être limitée de façon trop étroite. Avec le développement des réseaux d'information, il faut attirer l'attention sur le fait que les termes d'indexation créés initialement pour un groupe d'utilisateurs (par exemple scientifiques ou techniciens) pourraient utilement être étudiés par d'autres groupes d'utilisateurs (par exemple: économistes). Aussi, il est recommandé, par exemple, aux indexeurs de littérature scientifique et technique, de ne pas négliger les autres facettes d'un sujet, par exemple son aspect social ou économique.

6.3.3 Le critère principal, dans la sélection des notions, doit toujours être la valeur potentielle d'une notion en tant qu'élément dans l'expression du contenu d'un document et dans sa recherche documentaire. En choisissant les notions, l'indexeur doit garder à l'esprit les questions, dans la mesure où celles-ci peuvent être connues, qui peuvent être posées au système d'information. En effet, ce critère constitue la principale fonction d'indexation. Dans ce contexte, l'indexeur doit:

a) choisir les notions les plus appropriées à un groupe d'utilisateurs donné, en ne perdant pas de vue l'objectif de l'indexation;

b) modifier, si nécessaire, à la fois les outils d'indexation et les procédures, à partir du résultat des questions posées au système. Mais ces modifications ne doivent pas entraîner une distorsion de la structure de la logique du langage d'indexation.

6.3.4 Le nombre de termes ou de descripteurs pouvant être assignés à un document ne doit pas être limité arbitrairement. Il doit être déterminé entièrement par la quantité d'information contenue dans le document en liaison avec les besoins supposés des utilisateurs. Imposer une limite arbitraire peut conduire à une perte d'objectivité dans l'indexation et à une déformation de l'information qui aurait pu être utile pour la recherche. S'il est nécessaire de limiter le nombre de termes à l'intérieur d'un centre donné, la sélection des notions doit être guidée par le seul jugement de l'indexeur quant au rôle de chaque notion dans l'expression de la totalité du sujet du document.

6.4 La spécificité est liée à l'exactitude avec laquelle une notion particulière qui apparaît dans un document est représentée par un terme d'indexation. Une perte de spécificité survient lorsqu'une notion particulière est représentée par un terme ayant une signification plus générale.

Les notions doivent être identifiées d'une façon aussi spécifique que possible. Des notions plus générales peuvent être préférées dans les circonstances suivantes:

a) Quand l'indexeur considère qu'un excès de spécificité pourrait agir d'une façon négative sur le fonctionnement du système d'indexation. Un indexeur peut décider, par exemple, que des modèles très spécifiques d'équipement peuvent être représentés par des termes plus généraux tels que le nom du fabricant et peut-être la famille des modèles, essentiellement lorsque ces notions apparaissent seulement dans les zones limites du domaine couvert par l'indexation.

b) Selon le poids attaché à une notion par l'auteur. Si l'indexeur considère qu'une idée n'est pas complètement développée, ou qu'une allusion y est faite occasionnellement par l'auteur, l'indexation à un niveau plus général peut être justifiée.

7 Choix des termes d'indexation

7.1 Lorsque les notions sont traduites en termes d'indexation, l'indexeur doit observer les règles suivantes (voir aussi l'annexe):

a) Les notions qui sont déjà présentes dans le langage d'indexation doivent être retenues sous leur forme préférée.

b) Les termes qui représentent de nouvelles notions doivent être vérifiés à la fois sur leur fond et sur leur forme à l'aide des outils de référence tels que:

— dictionnaires et encyclopédies reconnus comme faisant autorité dans leurs domaines;

- thésaurus, particulièrement ceux construits conformément à l'ISO 2788 ou ISO 5964;
- plans de classement.

On peut aussi consulter des spécialistes du sujet, particulièrement ceux ayant quelques connaissances dans les domaines de l'indexation et de la documentation.

7.2 L'indexeur doit être familiarisé avec les outils d'indexation. En particulier, il doit être conscient du fait que ces outils peuvent imposer certaines contraintes. Par exemple, une liste d'autorité de vedettes matières ou les rubriques d'un plan de classement peuvent ne pas permettre une représentation exacte de la notion rencontrée dans le document. Si les notions sont représentées par des codes de classement, il faut savoir que ces codes désignent généralement une classe plus large ou plus étroite, qui peut ne pas s'appliquer entièrement au document étudié.

7.3 Si le langage d'indexation est constitué d'un thésaurus, la multiplicité des termes assignés à un document et le nombre des entrées peuvent être réduits sans perte, puisque les relations génériques et autres relations *a priori* peuvent être établies directement à partir du thésaurus lui-même. Lorsqu'on utilise un thésaurus, il faut sélectionner le terme le plus spécifique existant pour représenter une notion donnée.

7.4 Certains systèmes d'indexation utilisent des indicateurs de rôle, des liens, une pondération, etc. L'indexeur doit être familiarisé avec toutes les règles particulières de syntaxe.

7.5 Dans la pratique, l'indexeur rencontrera souvent des notions qui n'existent dans aucun thésaurus ni aucun plan de classement. Selon le système utilisé, ces notions peuvent être traitées de différentes façons, par exemple :

- a) exprimées par des termes ou descripteurs admis immédiatement dans le langage d'indexation ;
- b) représentées temporairement par des termes plus généraux, les nouvelles notions étant proposées comme candidates pour une addition ultérieure.

8 Contrôle de la qualité

8.1 La qualité et la cohérence de l'indexation dépendent de facteurs tels que :

- a) la compétence de l'indexeur ;
- b) la qualité des outils d'indexation.

Dans une situation idéale, les termes d'indexation assignés à un document et le niveau d'exhaustivité atteint lors de l'indexation doivent être identiques quel que soit l'indexeur. Ces facteurs doivent être plus restreints relativement stables dans le temps pour un même système d'indexation. La cohérence est un facteur important dans le fonctionnement d'un système d'indexation performant, surtout lorsque l'information est échangée entre les différents centres d'un réseau documentaire.

8.2 L'impartialité totale de l'indexeur est un facteur nécessaire dans la réalisation de la cohérence d'indexation. Un jugement subjectif dans l'identification des notions et le choix des termes d'indexation affectera inévitablement le fonctionnement du système d'indexation. La cohérence est plus difficile à réaliser dans une grande équipe d'indexation ou lorsque l'indexation est faite par des équipes d'indexeurs travaillant dans des endroits différents, comme dans un système décentralisé. Dans de telles situations, une étape de vérification centralisée, avec rétroaction vers les indexeurs, est recommandée.

8.3 L'indexeur doit avoir une bonne connaissance du domaine couvert par les documents qu'il indexe. Il doit comprendre les termes rencontrés dans les documents aussi bien que les règles et les procédures du langage d'indexation spécifique. Les organismes traitant de documents en langues étrangères doivent avoir recours à des spécialistes de ces langues.

8.4 La qualité d'indexation sera d'autant mieux atteinte que les indexeurs auront un contact direct avec les utilisateurs. Ils pourront alors, par exemple, déterminer si certains termes ou descripteurs sont susceptibles de produire des combinaisons fausses et de créer ainsi des sorties non pertinentes.

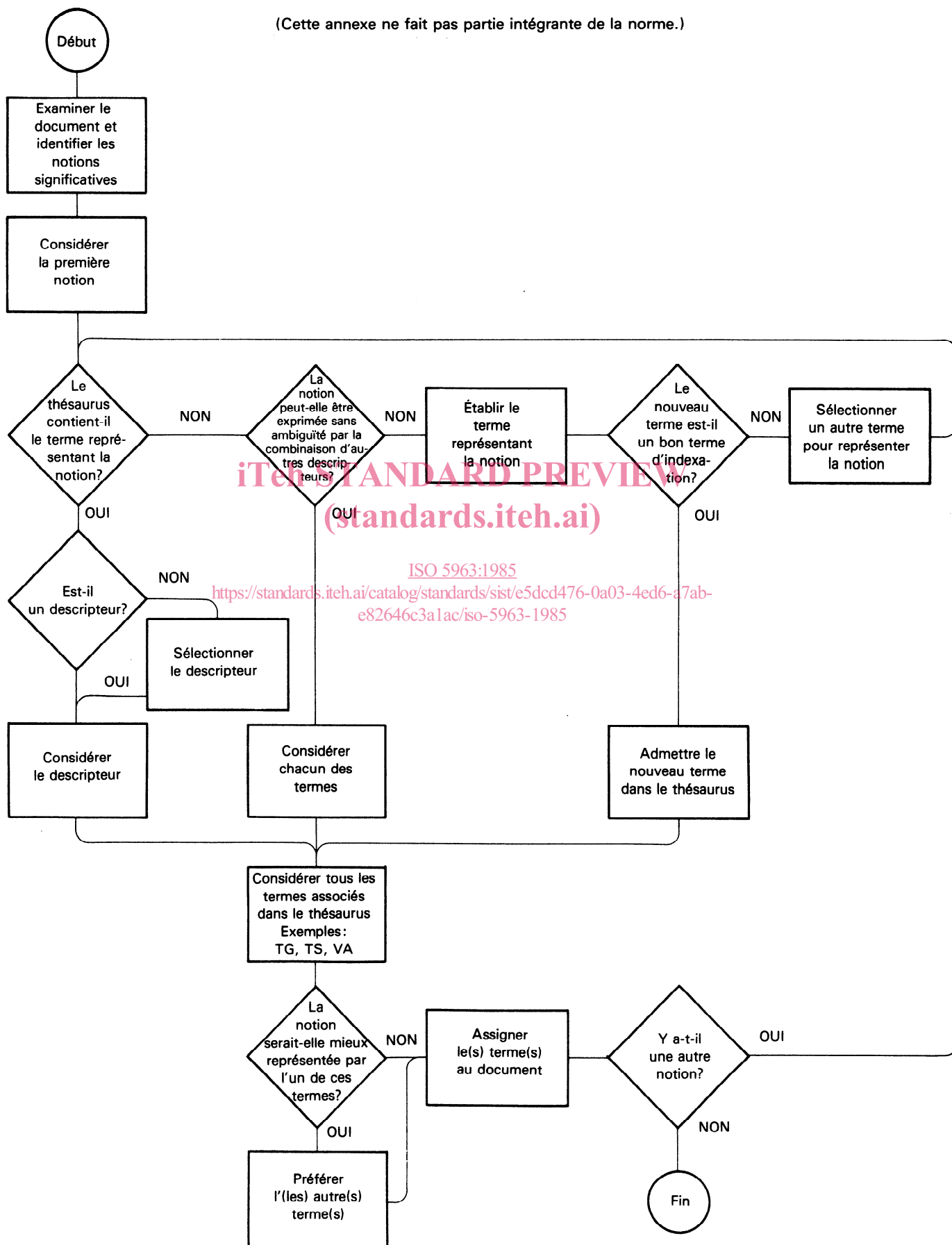
8.5 La qualité d'indexation dépend aussi de la possibilité de mise à jour du langage d'indexation. Il est essentiel que le système permette l'introduction de nouveaux termes dans le langage ou des changements dans la terminologie, répondant ainsi aux nouveaux besoins de ses utilisateurs.

8.6 La qualité d'indexation doit, si possible, être testée en analysant les résultats de la recherche documentaire, par exemple en calculant les taux de rappel et de pertinence.

Annexe

Organigramme du processus d'indexation utilisant un thésaurus

(Cette annexe ne fait pas partie intégrante de la norme.)



ITC STANDARD PREVIEW (standards.iteh.ai) ISO 5963:1985 https://standards.iteh.ai/catalog/standards/sist/e5dcd476-0a03-4ed6-a7ab-e82646c3a1ac/iso-5963-1985