

---

---

**Gestion des ressources  
linguistiques — Cadre global  
d'annotation (ComAF) —**

**Partie 3:  
Création sémantique  
diagrammatique (DSA)**

iTeh STANDARD PREVIEW  
(standards.iteh.ai)

*Language resource management — Comprehensive Annotation  
Framework (ComAF) —*

*Part 3: Diagrammatic semantic authoring (DSA)*

<https://standards.iteh.ai/catalog/standards/sist/20450947-c072-4892-bf12-39daca6e861e/iso-24627-3-2021>



**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

ISO 24627-3:2021

<https://standards.iteh.ai/catalog/standards/sist/20450947-c072-4892-bf12-39daca6e861e/iso-24627-3-2021>



**DOCUMENT PROTÉGÉ PAR COPYRIGHT**

© ISO 2021

Tous droits réservés. Sauf prescription différente ou nécessité dans le contexte de sa mise en œuvre, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie, ou la diffusion sur l'internet ou sur un intranet, sans autorisation écrite préalable. Une autorisation peut être demandée à l'ISO à l'adresse ci-après ou au comité membre de l'ISO dans le pays du demandeur.

ISO copyright office  
Case postale 401 • Ch. de Blandonnet 8  
CH-1214 Vernier, Genève  
Tél.: +41 22 749 01 11  
E-mail: [copyright@iso.org](mailto:copyright@iso.org)  
Web: [www.iso.org](http://www.iso.org)

Publié en Suisse

# Sommaire

	Page
<b>Avant-propos</b> .....	<b>iv</b>
<b>Introduction</b> .....	<b>v</b>
<b>1</b> <b>Domaine d'application</b> .....	<b>1</b>
<b>2</b> <b>Références normatives</b> .....	<b>1</b>
<b>3</b> <b>Termes et définitions</b> .....	<b>1</b>
<b>4</b> <b>Spécification</b> .....	<b>1</b>
<b>Annexe A (informative) Exemples</b> .....	<b>3</b>
<b>Bibliographie</b> .....	<b>9</b>

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO 24627-3:2021](https://standards.iteh.ai/catalog/standards/sist/20450947-c072-4892-bf12-39daca6e861e/iso-24627-3-2021)

<https://standards.iteh.ai/catalog/standards/sist/20450947-c072-4892-bf12-39daca6e861e/iso-24627-3-2021>

## 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 en général 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. L'ISO collabore étroitement avec la Commission électrotechnique internationale (IEC) en ce qui concerne la normalisation électrotechnique.

Les procédures utilisées pour élaborer le présent document et celles destinées à sa mise à jour sont décrites dans les Directives ISO/IEC, Partie 1. Il convient, en particulier, de prendre note des différents critères d'approbation requis pour les différents types de documents ISO. Le présent document a été rédigé conformément aux règles de rédaction données dans les Directives ISO/IEC, Partie 2 (voir [www.iso.org/directives](http://www.iso.org/directives)).

L'attention est attirée sur le fait que certains des éléments du présent document peuvent faire l'objet de droits de propriété intellectuelle ou de droits analogues. L'ISO ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de propriété et averti de leur existence. Les détails concernant les références aux droits de propriété intellectuelle ou autres droits analogues identifiés lors de l'élaboration du document sont indiqués dans l'Introduction et/ou dans la liste des déclarations de brevets reçues par l'ISO (voir [www.iso.org/brevets](http://www.iso.org/brevets)).

Les appellations commerciales éventuellement mentionnées dans le présent document sont données pour information, par souci de commodité, à l'intention des utilisateurs et ne sauraient constituer un engagement.

Pour une explication de la nature volontaire des normes, la signification des termes et expressions spécifiques de l'ISO liés à l'évaluation de la conformité, ou pour toute information au sujet de l'adhésion de l'ISO aux principes de l'Organisation mondiale du commerce (OMC) concernant les obstacles techniques au commerce (OTC), voir le lien suivant : [www.iso.org/iso/fr/avant-propos.html](http://www.iso.org/iso/fr/avant-propos.html).

Le présent document a été élaboré par le comité technique ISO/TC 37, *Langage et terminologie*, sous-comité SC 4, *Gestion des ressources linguistiques*.

Une liste de toutes les parties de la série ISO 24627 se trouve sur le site web de l'ISO.

Il convient que l'utilisateur adresse tout retour d'information ou toute question concernant le présent document à l'organisme national de normalisation de son pays. Une liste exhaustive desdits organismes se trouve à l'adresse [www.iso.org/fr/members.html](http://www.iso.org/fr/members.html).

## Introduction

Les graphes (diagrammes composés de nœuds et de liens) sont utilisés depuis des décennies pour représenter et visualiser à la fois des documents (données d'instance) et des schémas de données. Le présent document concerne la représentation (et non la visualisation) de documents (et non de schémas de données) basée sur des graphes.

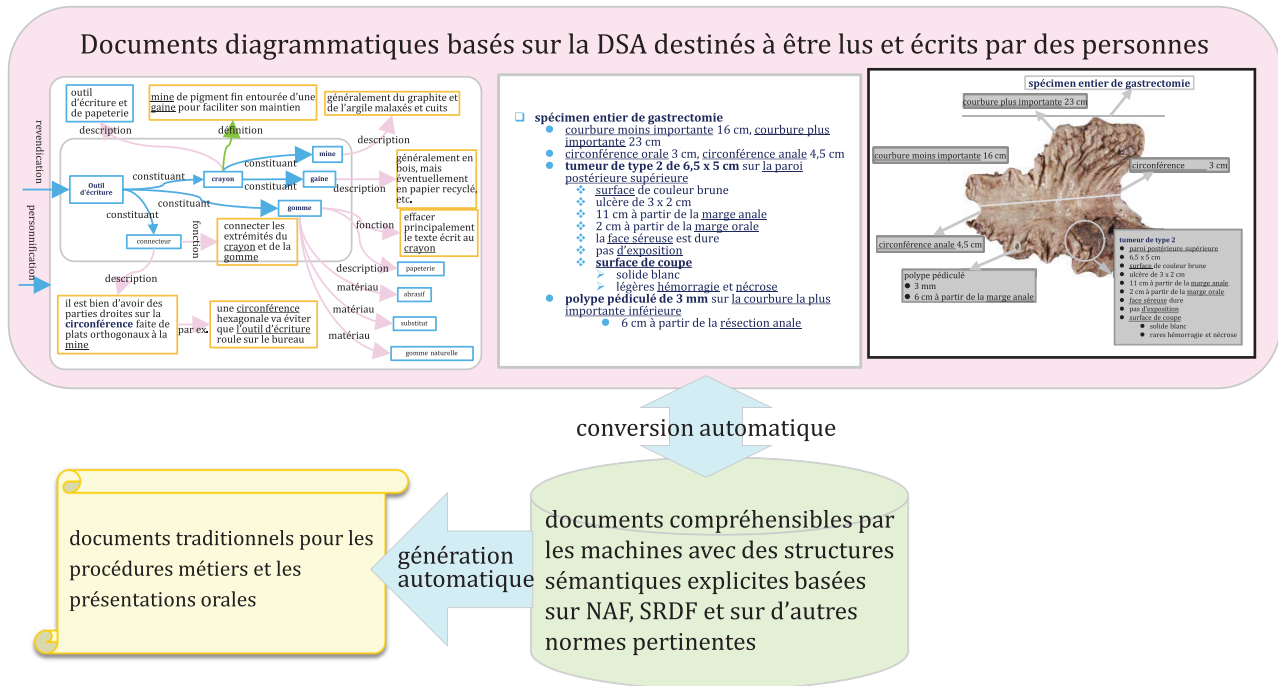
La représentation et la visualisation de documents basées sur des graphes sont traitées par des cartes conceptuelles<sup>[15]</sup>, des cartes mentales, des cartes d'arguments, etc. La linguistique théorique et l'intelligence artificielle utilisent aussi une visualisation du contenu basée sur des graphes, associée à un réseau sémantique, un espace mental<sup>[10]</sup>, une structure de représentation du discours<sup>[13]</sup>, etc.

La visualisation des schémas de données (ou ontologies, terminologies, métamodèles, etc.) basée sur des graphes est une pratique plus courante. Les ontologies sont souvent visualisées sous forme de graphes dans lesquels les nœuds sont des classes (et des types de données) et les liens des propriétés (relations). L'ISO 24156-1 spécifie une visualisation de la modélisation conceptuelle basée sur le langage UML. D'autres métamodèles sont généralement aussi représentés sous forme de diagrammes similaires.

Le présent document fournit un schéma de données pour les documents de graphes, afin de faciliter la composition et la compréhension en rendant la structure logique du document explicite. Il ne concerne pas les visualisations ou les manipulations des graphes et il ne définit pas non plus les annotations de documents existants, mais il traite plutôt de la représentation graphique/diagrammatique des documents pour les besoins de la création sémantique, c'est-à-dire pour que les personnes visualisent et manipulent directement des structures syntactiques/sémantiques sur des écrans d'ordinateur ou leurs futures alternatives. La linéarité des documents textuels traditionnels est due à la linéarité des langages parlés, qui contraint l'interaction entre les personnes et les documents, et qui rend difficiles la lecture et l'écriture. La DSA définit des documents graphiques/diagrammatiques avec des structures plus explicites que dans le texte, afin de faciliter leur lecture et écriture. Les documents basés sur la DSA, associés à des interfaces utilisateur impliquant des visualisations appropriées et des opérations simples, peuvent améliorer la collaboration entre les personnes, et entre les personnes et les machines.

La DSA traite principalement des structures syntactiques ou documentaires. Elle traite aussi de certaines structures sémantiques fragmentaires, mais une autre spécification peut fournir une sémantique plus systématique (mappage formel entre documents et leurs significations ou formes logiques), pour que les machines puissent mieux « comprendre » les documents basés sur la DSA et donc mieux assister le partage d'informations et l'obtention d'un consensus entre les personnes.

La [Figure 1](#) illustre un flux de travail impliquant la DSA et d'autres types de documents. Les documents basés sur la DSA dans la moitié supérieure peuvent être convertis automatiquement (tout en préservant le contenu propositionnel) à partir de et vers des documents compréhensibles par les machines, sur la base de normes appropriées concernant les représentations et annotations sémantiques. Il est possible de générer automatiquement des documents textuels traditionnels à partir de ces documents compréhensibles par les machines (tout en préservant également le contenu propositionnel), bien que la conversion inverse ne puisse généralement pas être automatisée. Comme les documents basés sur la DSA (associés à des interfaces utilisateur appropriées) sont plus faciles à composer et à interpréter par les personnes que les documents textuels, les personnes peuvent généralement toucher et voir les documents basés sur la DSA, tandis que les documents traditionnels peuvent être utilisés pour les procédures métiers (telles que les demandes de brevets) et les présentations orales.



**Figure 1 — Flux de travail d'un document impliquant la DSA**

**iTeh STANDARD PREVIEW**

La DSA est un métamodèle minimal pour l'ISO/TS 24617-5 (SemAF-DS), qui repose à son tour sur l'ISO/IEC 15938.5/Amd.1 (MPEG-7 MDS AMD1 – Schéma descriptif linguistique). Les documents compréhensibles par les machines sur la [Figure 1](#) sont supposés utiliser d'autres normes, notamment l'ISO 24615 (SynAF), l'ISO 24612 (LAF) et l'ISO 24617 (SemAF), tout en incorporant des aperçus tirés d'autres littératures pertinentes [1][8][9][10][11][12][13][14][15][16][17][18]

# Gestion des ressources linguistiques — Cadre global d'annotation (ComAF) —

## Partie 3: Création sémantique diagrammatique (DSA)

### 1 Domaine d'application

Le présent document spécifie comment représenter (et non visualiser) des documents (données d'instance et non schémas de données) sous forme de graphes. Il ne spécifie pas comment visualiser ou utiliser des données documentaires, mais il vise à rendre les documents plus faciles à composer et à comprendre par les personnes en offrant diverses interfaces utilisateur flexibles incorporant éventuellement les pratiques de visualisation de documents (voir l'Introduction). À cet égard, le présent document ne spécifie pas non plus les annotations de documents existants, mais il spécifie plutôt un schéma de documents avec des structures logiques explicites.

### 2 Références normatives

Le présent document ne contient aucune référence normative.

### 3 Termes et définitions

Pour les besoins du présent document, les termes et définitions suivants s'appliquent.

L'ISO et l'IEC tiennent à jour des bases de données terminologiques destinées à être utilisées en normalisation, consultables aux adresses suivantes :

- ISO Online browsing platform : disponible à l'adresse <https://www.iso.org/obp>
- IEC Electropedia : disponible à l'adresse <http://www.electropedia.org/>

#### 3.1

##### **hypernœud**

nœud qui est un segment de graphe

#### 3.2

##### **segment**

partie référençable d'un document basé sur la DSA, qui est soit un segment de graphe, soit un segment de données (texte, image, audio, vidéo, etc.)

#### 3.3

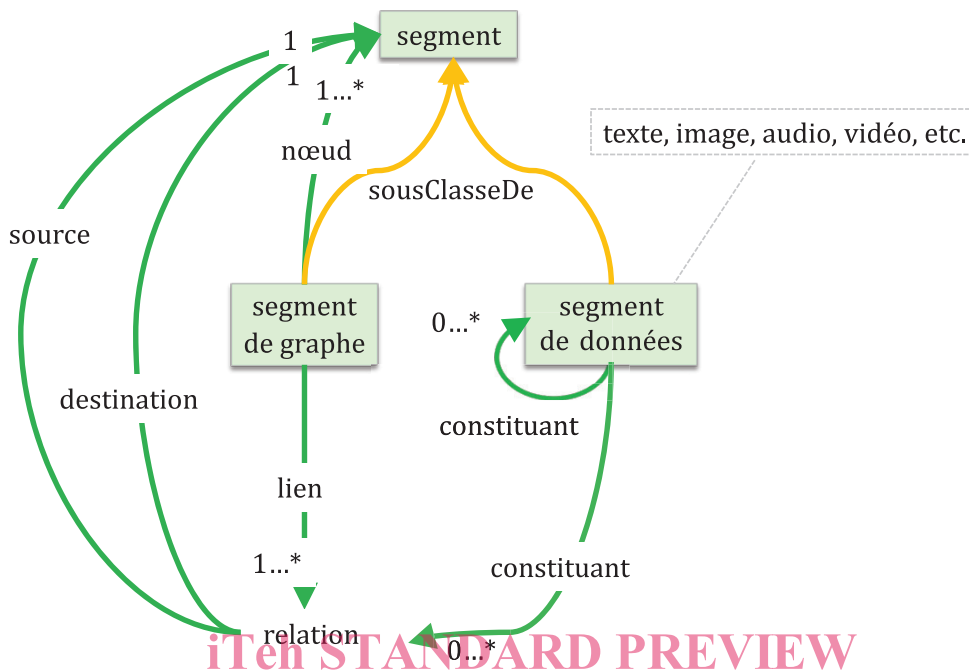
##### **création sémantique**

composition de documents rendant leurs structures logiques explicites

### 4 Spécification

La [Figure 2](#) illustre la DSA, qui est un métamodèle (ontologie) de documents de graphes. Chaque partie marquable (référençable) d'un document basé sur la DSA est appelée un segment. Chaque document basé sur la DSA proprement dit est aussi un segment.

# DSA Metamodel



iteh STANDARD PREVIEW  
(standards.iteh.ai)

Figure 2 — Métamodèle DSA (ontologie)

ISO 24627-3:2021

Un segment de graphe doit être un graphe orienté étiqueté comprenant des nœuds et des liens. Un nœud doit être un segment, et un lien doit être une relation entre deux nœuds terminaux (segments source et destination).

Un nœud dans un segment de graphe peut être un autre segment de graphe. Un nœud qui est un segment de graphe est appelé un hypernœud. Un segment de graphe peut donc intégrer d'autres segments de graphes sous forme d'hypernœuds, ceux-ci étant rarement utilisés dans les cartes conceptuelles. Arbitrairement, les gros documents basés sur la DSA peuvent donc être constitués de petits segments de graphes qui sont des hypernœuds les uns des autres.

Un segment de données peut intégrer de plus petits segments de donnée en tant que constituants. Par exemple, un discours peut être un segment de texte/audio/vidéo intégrant des énoncés de phrases en tant que plus petits segments de textes/audio/vidéo, qui peuvent à leur tour intégrer des segments de textes/audio/vidéo encore plus petits, tels que des expressions, et ainsi de suite.

La DSA est exactement le métamodèle de la [Figure 2](#), qui spécifie formellement des documents de graphes multimodaux impliquant des hypernœuds, mais qui ne spécifie pas comment les visualiser ou les utiliser. Les visualisations diagrammatiques dans l'[Annexe A](#) ne font pas partie de la DSA, elles sont fournies à titre d'exemples informatifs. Les utilisateurs de la DSA peuvent adopter toutes sortes de visualisations diagrammatiques dans la mesure où les structures logiques des documents sont rendues assez explicites.



## Annexe A (informative)

### Exemples

#### A.1 Segments de graphe

Sur la [Figure A.1](#), les segments de graphes sont visualisés sous forme de cadres à bords arrondis. Le segment de graphe plus petit est un hypernœud dans le plus grand. La plupart des nœuds sur la [Figure A.1](#) sont des segments de textes, mais les nœuds peuvent aussi correspondre à d'autres types de segments. Les liens dans cet exemple représentent les relations du discours et les rôles thématiques. Par exemple, une « mine » et une « gaine » sont les constituants d'un « crayon » et la fonction d'une « gomme » est d'« effacer principalement le texte écrit au crayon ».

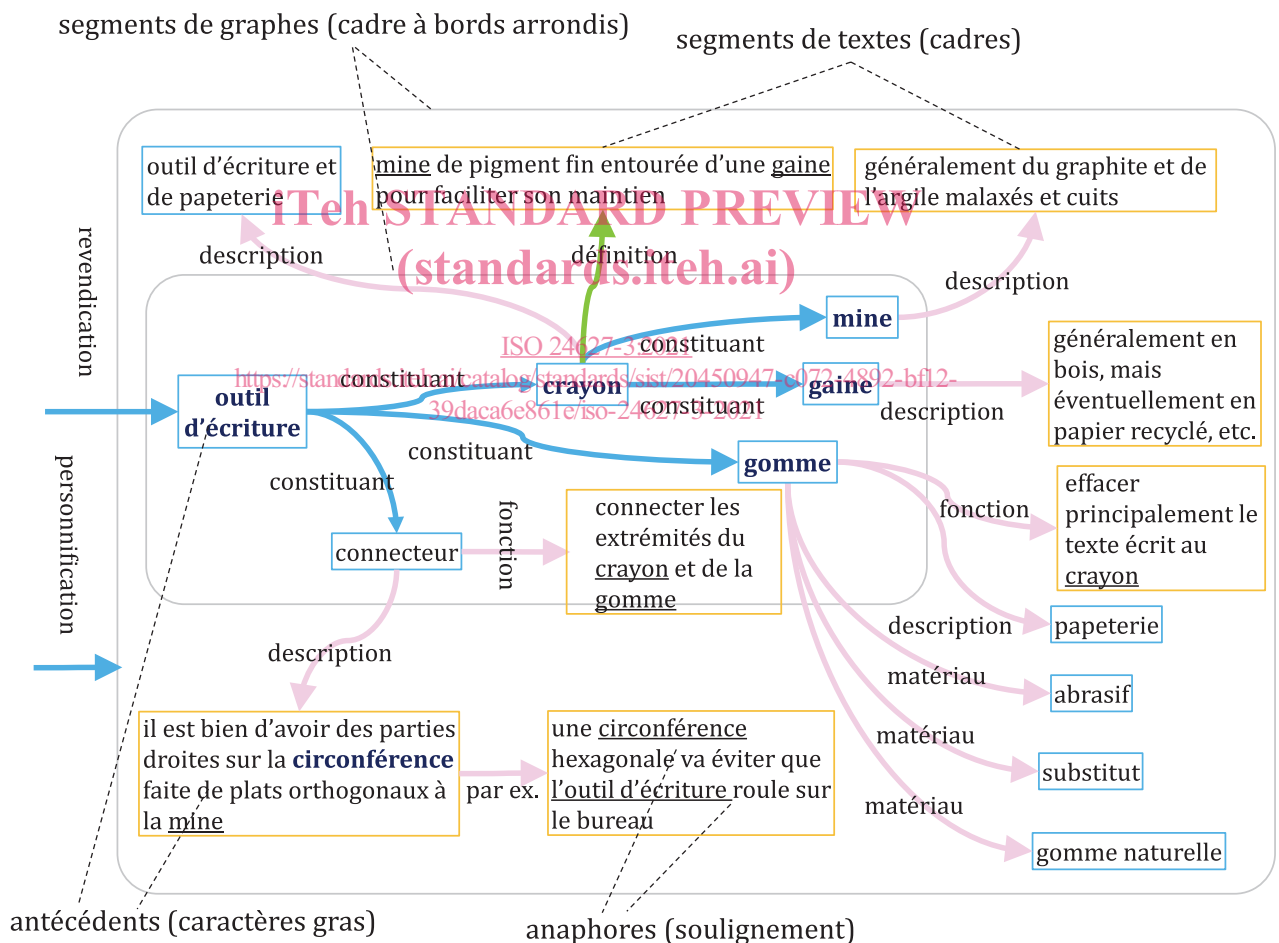


Figure A.1 — Segment de graphe représentant une partie d'un document de brevet

Dans l'ensemble du présent document, les anaphores sont soulignées et leurs antécédents sont en caractères gras. Les relations anaphoriques peuvent être visualisées explicitement sous forme de liens tels que ceux illustrés sur la [Figure A.2](#). Certains nœuds terminaux de ces liens anaphoriques sont des segments de textes intégrés dans des segments de graphes/de textes plus grands, et ces nœuds terminaux n'appartiennent donc pas aux graphes auxquels les relations anaphoriques appartiennent. L'un des problèmes rencontrés est le fait que si un graphe contient un trop grand nombre de ces relations

anaphoriques, il est difficile à comprendre et à composer par les personnes. Ce problème est en partie résolu en visualisant les graphes sous forme d'arbres.

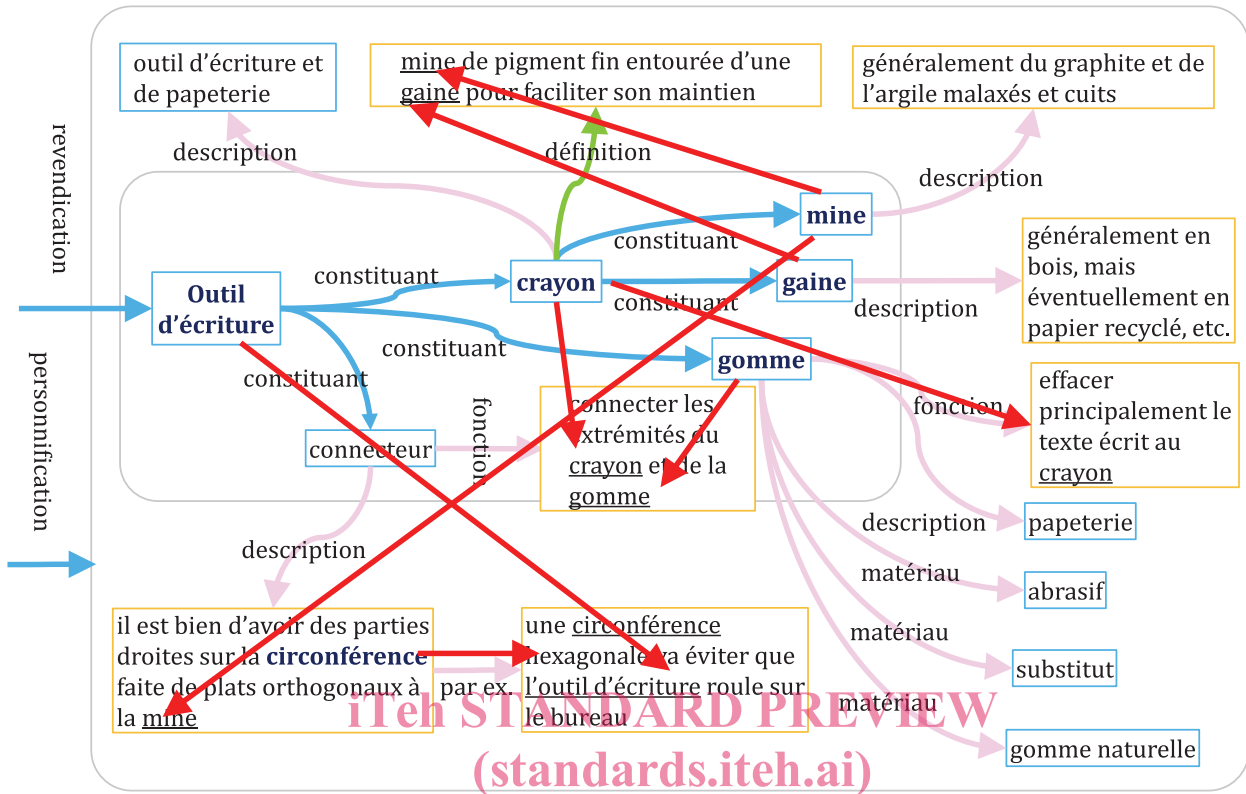


Figure A.2 — Liens anaphoriques dans un segment de graphe

<https://standards.iteh.ai/catalog/standards/sist/20450947-c072-4892-bf12-39daca6e861e/iso-24627-3-2021>

Certains segments de graphes peuvent être présentés visuellement sous forme d'arbres. Dans le présent document, un tel arbre est inclus dans un cadre et contient trois nœuds commençant par des puces, sauf le nœud racine qui n'a pas de puce. Dans l'arborescence illustrée sur la Figure A.3, toutes les puces représentent une sorte d'élaboration, c'est-à-dire que le nœud enfant élabore le nœud parent en décrivant les détails (par exemple les parties) du référent du nœud parent. On doit choisir le bon nœud racine dans l'arborescence afin de bénéficier d'une interprétation uniforme des puces. Il convient d'utiliser une étiquette de relation plus spécifique à la place d'une simple puce si un nœud n'élabore pas son parent. Dans le présent document, les segments de textes sont inclus dans des cadres comme illustré sur la Figure A.2, mais ces cadres sont omis lorsque les segments de textes sont des nœuds d'arbres, comme sur la Figure A.3.