



HAL
open science

Description sémantique des stades de développement phénologique des plantes, cas d'étude de la vigne

Catherine Roussey, Xavier Delpuech, Marc Raynal, Florence Amardeilh,
Stéphan Bernard, Clement Jonquet, Camille Noûs

► To cite this version:

Catherine Roussey, Xavier Delpuech, Marc Raynal, Florence Amardeilh, Stéphan Bernard, et al..
Description sémantique des stades de développement phénologique des plantes, cas d'étude de la
vigne. 32èmes Journées Francophones d'Ingénierie des Connaissances (IC 2021), Jun 2021, Bordeaux,
France. pp.30-38. emse-03260085

HAL Id: emse-03260085

<https://hal-emse.ccsd.cnrs.fr/emse-03260085>

Submitted on 14 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Description sémantique des stades de développement phénologique des plantes, cas d'étude de la vigne

Catherine Roussey¹, Xavier Delpuech², Marc Raynal³,
Florence Amardeilh⁴, Stephan Bernard¹, Clement Jonquet^{5,6},
Camille Noûs⁷

¹ Université Clermont Auvergne, INRAE, UR TSCF, Aubière

² Institut français de la vigne et du vin, Pôle Rhône-Méditerranée, Montpellier

³ Institut français de la vigne et du vin, Pôle Nouvelle Aquitaine / UMT SEVEN, Bordeaux

⁴ Elzeard R&D, Pessac

⁵ MISTEA, Université de Montpellier, INRAE, Institut Agro, Montpellier

⁶ LIRMM, Université de Montpellier, CNRS, Montpellier

⁷ Laboratoire Cogitamus

catherine.roussey@inrae.fr, xavier.delpuech@vignevin.com, marc.raynal@vignevin.com,
florence.amardeilh@elzeard.co, stephan.bernard@inrae.fr, jonquet@lirmm.fr,
camille.nous@cogitamus.fr

Résumé

Le projet ANR "Des Données aux Connaissances en Agronomie et Biodiversité" (D2KAB) met à disposition une archive de bulletins agricoles publiée sur le Web. Pour annoter les bulletins à l'aide des stades de développement des plantes, nous avons besoin d'une nouvelle ressource sémantique. Plusieurs échelles phénologiques existent, chacune de ces échelles définit un ensemble spécifique de stades. L'Institut Français de la Vigne et du vin (IFV) a aligné plusieurs échelles existantes associées à la vigne. Nous présentons ici la création d'un cadre méthodologique pour la description sémantique des stades de développement des plantes fondé sur l'ontologie PPDO (BBCH-based Plant Phenological Description Ontology).¹

Mots-clés

Ontologie, Web de données, Données liées, Stades phénologiques /de développement, Agriculture, Vigne, BBCH.

Abstract

The French ANR project "Data to Knowledge in Agronomy and Biodiversity" (D2KAB) builds an archive of French agricultural alert newsletters. In order to annotate plant development stages, we need a new semantic resource. Several phenological scales already exist to describe plant development stages. The French Wine and Vine Institute (IFV) has aligned several existing scales related to grapevine. In this paper, we present our work of creating an ontological framework for semantic description of plant development stages; we introduce the BBCH-based Plant Phenological

1. Cet article est une traduction et extension d'un article publié en anglais en Décembre 2020 à la 14ème conférence internationale *Metadata and Semantics Research (MSTR)*.

Description Ontology.²

Keywords

Ontology, Web of Data, Linked Open Data, Phenological / Development stages, agriculture, Grapevine, BBCH

1 Introduction

L'agronomie et l'agriculture sont confrontées à plusieurs défis sociétaux, économiques et environnementaux majeurs, qu'une approche sémantique de la science des données aidera à relever. Le projet ANR *Des Données aux Connaissances en Agronomie et Biodiversité* (D2KAB)³ illustre comment la science des données sémantiques contribue au développement d'applications agricoles innovantes. L'objectif de D2KAB est de créer un cadre pour transformer les données d'agronomie et de biodiversité en connaissances décrites sémantiquement, interopérables, exploitables et ouvertes. Pour construire un tel cadre, nous nous appuyons sur des ressources sémantiques (par exemple, des terminologies, des vocabulaires ou des ontologies) pour décrire nos données et les publier en tant que données ouvertes liées [6]. Nous utilisons le portail web AgroPortal⁴ [13] pour trouver, publier et partager des ressources sémantiques puis nous exploitons ces ressources sémantiques dans des applications dédiées à l'agriculture ou l'environnement.

L'un des scénarios agricoles de D2KAB consiste à construire un navigateur web augmenté pour les bulletins officiels d'alertes agricoles français, appelés *Bulletins*

2. This article is a translation and an extension of an article published in English, in December 2020 at the 14th International Conference on Metadata and Semantics Research (MSTR).

3. www.d2kab.org

4. <http://agroportal.lirmm.fr>

de Santé du Végétal (BSV). Le prototype interrogera une archive de BSV disponible sous forme de fichiers PDF. Chaque bulletin sera annoté sémantiquement, et les annotations seront publiées sur le Web de données liées. Les annotations seront produites, entre autre, à partir de techniques de traitement automatique de la langue appliquées sur les contenus textuels des BSV. Les annotations seront décrites en utilisant les ontologies Semantic Sensor Network combinées au Web Annotation Data Model comme présenté dans de précédents travaux [3].

Des mentions de stades phénologiques sont toujours présentes dans les BSV, afin d'identifier les périodes de sensibilité d'une culture aux facteurs abiotiques (par exemple, le gel printanier) ou à des facteurs biotiques (par exemple, la maladie du mildiou). Afin d'annoter les stades phénologiques des cultures, nous avons besoin d'une ressource sémantique spécifique pour représenter les stades, l'échelle phénologique à laquelle ils appartiennent et les cultures auxquelles ils s'appliquent. Comme il existe plusieurs échelles phénologiques par culture, nous avons aussi besoin d'aligner les stades des échelles différentes concernant la même culture. A noter que la plupart des échelles phénologiques sont dédiées à une seule culture. Une des échelles les plus connues et qui s'applique à plusieurs cultures est l'échelle BBCH [17]. Malheureusement, ces échelles ne sont pas publiées sur le Web de données liées, ni disponibles dans un format du Web sémantique (e.g., SKOS, RDF-S ou OWL), empêchant leur utilisation dans différentes applications agricoles. Notre corpus de BSV montre que différentes échelles ont été utilisées pour observer les stades phénologiques. Par exemple, dans un BSV lié à la vigne, on trouve la phrase : *bourgeon dans le coton (stade 03 ou B ou BBCH 05) dans les secteurs tardifs*, qui est à la fois une référence à un 'label' ("bourgeon dans le coton ") et la codification de plusieurs échelles : (i) 03 est un code de l'échelle Eichhorn-Lorenz [14]; (ii) B est un code de l'échelle de Baggiolini [5]; (iii) et BBCH 05 est un code de l'échelle BBCH [17]. Ces codes devront être identifiés automatiquement dans un texte pour être associés à un identifiant pérenne représentant le stade d'une échelle phénologique. De plus, il semble que les labels les plus fréquemment utilisés soient ceux d'une échelle produite par l'Institut Français de la Vigne et du vin (IFV). Malheureusement cette échelle propose uniquement des labels. D'après notre expérience sur l'usage des BSV, il est nécessaire de construire un ensemble de jeux de données alignés pour décrire sémantiquement les stades des différentes échelles.

Cet article est une traduction et extension d'un article précédent [22]; nous présentons la finalisation de notre travail de transformation en RDF de six échelles (une de plus que l'article précédent) : l'échelle générale BBCH qui s'applique à toutes les cultures [17], l'échelle individuelle BBCH de la vigne [17], l'échelle Eichhorn-Lorenz [14] qui ne concerne que la vigne, l'échelle Baggiolini de la vigne [5], l'échelle produite par IFV qui est une sélection de l'échelle individuelle BBCH de la vigne (intitulée IFV-labels), l'échelle produite par IFV qui est un enrichissement de l'échelle

Eichhorn-Lorenz (intitulée IFV-Epicure).

Nous montrerons plus particulièrement notre travail d'alignement entre ces échelles. Nous présenterons la première version de l'ontologie intitulée *BBCH-based Plant Phenological Description Ontology (PPDO)*. Cette ontologie est le support d'un cadre méthodologique pour encoder sémantiquement toute échelle phénologique. Cette ontologie décrit l'échelle générale BBCH sous forme de classes OWL. Puis, chaque stade d'une échelle BBCH spécifique à une culture est représentée par une instance des classes précédentes.

Le reste de l'article est organisé comme suit : La section 2 introduit la phénologie et présente les échelles que nous avons transformées; La section 3 passe en revue les thésaurus/ontologies existants pour décrire les stades phénologiques; La section 4 illustre la méthodologie utilisée pour construire l'ontologie PPDO et la peupler avec les stades des échelles de la vigne; La section 5 discute des avantages et des inconvénients de la publication des échelles phénologiques sous forme de données ouvertes liées.

2 Echelles phénologiques

La phénologie des plantes est l'étude du développement saisonnier des végétaux déterminé par l'influence des variations du climat. Les événements périodiques sont par exemple l'apparition des fleurs, le changement de couleurs des feuilles, etc. Un stade phénologique ou stade de développement des plantes caractérise une phase de développement de la plante pendant son cycle de vie. Par exemple, le stade floraison est observé quand au moins 50% des fleurs sont épanouies. L'étude phénologique d'une plante consiste à observer à quelle date les stades apparaissent. Les variations temporelles sur l'apparition d'un stade peuvent dépendre du climat ou d'autres facteurs comme la variété de la plante.

L'échelle phénologique BBCH améliorée (*Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie*) propose une codification homogène des stades de développement communs à différentes espèces végétales cultivées. BBCH décrit plusieurs ensembles de stades de développement : l'échelle générale et des échelles spécifiques par culture (dites « échelles individuelles ») [17]. L'échelle générale forme le cadre dans lequel les échelles individuelles sont développées. Elle peut également être utilisée pour les cultures n'ayant pas fait l'objet de description d'échelle individuelle : « *Les différentes phases du développement d'une plante sont divisées en dix stades principaux numérotés de 0 à 9. (...) Pour définir précisément les stades d'application ou d'évaluation, il n'est pas conseillé d'utiliser seulement les stades principaux, car ceux-ci recouvrent une durée importante dans le développement de la plante. On utilise les stades de développement secondaires pour déterminer un moment exact du développement. Par rapport aux stades principaux, les stades secondaires définissent des stades limités dans le temps. Ils sont donnés par des valeurs de 0 à 9, qui se suivent à l'intérieur d'un stade principal. On obtient ainsi un code à deux chiffres composé par le stade principal et le stade secondaire.* » [17]. BBCH est

considéré comme une référence pour décrire les stades et la monographie est disponible en anglais, français, allemand et espagnol [15, 17, 16, 18].

Pour la vigne, la première échelle phénologique a été proposée par Baggiolini [4] en 1952. Elle comportait initialement 10 stades qui ont été complétés en 1993 par Baillod et Baggiolini [5] pour atteindre 16 stades. En 1977, Eichhorn et Lorenz (EL) ont proposé une échelle plus détaillée de 24 stades pour la vigne. En 1992, une échelle universelle connue sous le nom de BBCH pour les plantes cultivées mono et dicotylédones a été proposée par Hack et al., puis adaptée à la vigne en 1995 par Lorenz et al. [14]. L'échelle individuelle BBCH de la vigne permet la comparaison avec d'autres espèces, aussi bien annuelles que pérennes, et s'est progressivement imposée comme une échelle de référence dans la communauté scientifique.

En 1995, Coombe [7] a proposé un alignement entre les différentes échelles existantes de la vigne tout en proposant des modifications de l'échelle EL. Cependant, dans la communauté technique agricole française, différentes échelles phénologiques sont parfois utilisées selon les habitudes des techniciens. L'IFV a donc produit une fiche technique [11] pour synthétiser les stades de développement de la vigne les plus utiles et les plus pertinents. Les stades sélectionnés ont été définis à l'aide d'expressions françaises spécifiques et illustrés par des images représentatives. Leurs correspondances avec les trois échelles phénologiques (échelle individuelle BBCH, EL et Baggiolini) ont été indiquées. Cette échelle, qui peut être vue comme une sélection des stades BBCH, porte ici le nom de IFV-Labels.

L'échelle IFV-Epicure complète celle d'EL et propose d'objectiver, par des éléments de mesure (pourcentage ou métrique), certaines définitions de stades trop imprécises : Le développement de la base de données Epicure au début des années 2000, a conduit l'IFV de Bordeaux à standardiser l'observation de stades phénologiques recueillis sur le terrain par les viticulteurs et techniciens. L'échelle IFV-Epicure reprend donc la numérotation proposée par EL, car c'est la plus utilisée en pratique sur le vignoble. Elle propose une numérotation continue du stade repos d'hiver (1) à la fin de la chute des feuilles (47) en occupant tous les numéros de stades laissés vacants par EL. Ce faisant, IFV-Epicure propose un échelonnage plus régulier et affine l'observation de certains stades trop imprécis par adjonction de critères mesurables : par exemple, le stade EL 31 (ou K de Baggiolini) défini par "grains de pois" reste le stade 31 d'IFV-Epicure mais est complété par une évaluation du diamètre des baies (4 à 5 mm).

Même s'il existe d'autres échelles pour d'autres cultures, comme Zadoks (1974), il n'y a manifestement pas de liste de stades phénologiques universellement reconnue ; cependant, en raison de sa large adoption, l'échelle BBCH peut être vue comme un langage pivot pour décrire les stades de développement spécifiques aux cultures. En d'autres termes, les jeux de données RDF décrivant les stades spécifiques d'une culture devront être alignés avec l'échelle générale BBCH (soit via une instanciation, soit via des correspondances explicites entre les échelles).

3 Ressources sémantiques décrivant les stades phénologiques

Aucune des échelles mentionnées précédemment n'est publiée sur le Web de données liées par leurs producteurs. Cependant, certains vocabulaires RDF ou ontologies OWL existants mentionnent des stades de développement. La recherche du mot clé «BBCH» sur AgroPortal nous retourne 3 ressources sémantiques : la Crop Ontology (CO), la Plant Ontology (PO) et SOY (une ontologie décrivant les traits spécifiques du soja). Nous avons également identifié d'autres ressources sémantiques pouvant contenir des descriptions de stades phénologiques, notamment le thésaurus AGROVOC et la Plant Phenology Ontology (PPO) [25].

3.1 AGROVOC

Le thésaurus AGROVOC est publié par l'Organisation des Nations Unies pour l'Alimentation et l'Agriculture (FAO) [23]. Il est édité par une communauté mondiale d'experts et couvre tous les domaines d'intérêt de la FAO, y compris l'agriculture, la sylviculture, la pêche, l'alimentation et les domaines connexes. Il est disponible en 29 langues, avec une moyenne de 35 000 termes par langue et développé en SKOS-XL. La force de ce thésaurus est sa couverture lexicale multilingue. Il est donc souvent utilisé pour annoter ou indexer des documents ou des images relatifs au domaine de l'agriculture.

AGROVOC contient certains stades phénologiques. L'instance de `skos:Concept` intitulé «stades de développement végétal» est la racine d'une hiérarchie de 38 stades. Par exemple, l'URI http://aims.fao.org/aos/agrovoc/c_2992 identifie le stade «floraison». Notons que la référence à une échelle existante n'est pas indiquée – c'est-à-dire qu'aucun texte associé à l'instance ne fait référence à une échelle en particulier comme BBCH et qu'aucun lien externe ne pointe vers une information de ce type. Idem pour la culture, aucune information ne permet de dire à quelle culture le stade peut s'appliquer. Pour ces raisons, AGROVOC ne correspond pas à notre objectif. Il pourra en revanche être une source externe à laquelle s'aligner pour plus d'interopérabilité.

3.2 La Plant Ontology et la Crop Ontology

La *Plant Ontology* (PO) [8] et la *Crop Ontology* (CO) [24], pour les plantes cultivées, sont deux résultats du projet Planteome [12]. L'objectif était de développer des ontologies de référence en génomique et en phénologie végétale. CO est le regroupement de plusieurs dictionnaires de traits observables spécifiques à une culture, tous liés à PO ainsi qu'à la *Plant Trait Ontology* (TO). L'un des dictionnaires de traits spécifiques à la vigne est l'ontologie *Vitis*,⁵ mais elle ne contient aucune mention de stade phénologique. PO couvre cependant toutes les espèces cultivées ; elle contient des termes normalisés, des définitions et des relations décrivant l'anatomie, la morphologie et les stades [26]. Dans PO, un stade phénologique est représenté par une classe. Un stade correspond à un intervalle de

5. http://agroportal.lirmm.fr/ontologies/CO_356

temps pendant le cycle de croissance d'un élément végétal. Ainsi, un stade bref est défini comme une sous-classe d'un stade plus long englobant ce stade bref. Un stade est lié à une entité anatomique végétale à l'aide de la propriété objet `has_participant`. Certaines références à des échelles phénologiques ont été trouvées dans les valeurs des propriétés d'annotation `has_related_synonym` de PO : 49 pour l'échelle BBCH, 38 pour l'échelle Zadoks. Par exemple, l'URI http://purl.obolibrary.org/obo/PO_0007086 correspond au stade "cinq nœuds visibles". Le code de ce stade dans l'échelle BBCH est BBCH 35. Comme la plupart des ontologies de la communauté OBO, PO a pour base l'ontologie "fondatrice" *Basic Formal Ontology* (BFO) [2] et réutilise des propriétés définies dans *Relation Ontology* (RO). Nous avons trouvé dans PO des propriétés (object properties) intéressantes. Les propriétés `precedes` et `precedes_by` indiquent que la fin d'un stade se produit avant le début d'un autre. Ces propriétés définissent des restrictions de domaine et de co-domaine liées aux classes de BFO. Ainsi, réutiliser ces propriétés signifie s'appuyer sur les classes de BFO. PO a été conçu pour annoter les données génomiques et phénomiques. Cette ontologie est plus détaillée que l'échelle BBCH qui elle a été conçue pour les pratiques agricoles. Par conséquent, nous avons jugé PO non pertinente pour notre cas d'usage.

3.3 Plant Phenology Ontology (PPO)

La *Plant Phenology Ontology* (PPO) pour la phénologie végétale fournit un vocabulaire standardisé utile pour l'intégration de données phénologiques hétérogènes du monde entier [25]. Elle s'appuie sur PO ainsi que sur *Phenotype and Trait Ontology* (PATO) pour décrire les phénotypes et les traits. La phénologie n'est pas l'observation des stades phénologiques à proprement parler ; mais l'observation des traits physiques des plantes, comme la présence de feuilles ou de fruits. Ainsi, PPO définit les traits observables et ne se concentre pas sur l'observation de début ou de la fin d'un stade. PPO permet de décrire précisément la présence ou l'absence de certaines entités anatomiques végétales modélisées comme des traits.

Pour conclure, parmi les ressources sémantiques identifiées, nous n'avons trouvé aucune solution à adopter directement pour notre cas d'usage. Bien que BBCH soit souvent référencé, ce qui confirme son rôle de langage pivot, nous n'avons pas trouvé de ressource sémantique qui s'appuierait sur BBCH et offrirait un moyen d'encoder sémantiquement les stades phénologique de toutes les cultures. Par conséquent, nous avons décidé de proposer un nouveau cadre méthodologique pour décrire les échelles phénologiques : ce cadre contiendra une ontologie OWL déclarant un ensemble de classes spécialisant le modèle SKOS ainsi que plusieurs jeux de données SKOS. Un jeu de données SKOS déclare l'ensemble des stades d'une échelle spécifique à une culture – dans le cas de la vigne nous aurons donc encodé et aligné les échelles suivantes : échelle générique et individuelle BBCH, les échelles EL et Baggiolini, les échelles produites par IFV intitulée IFV-labels et IFV-Epicure.

4 Un cadre méthodologique pour décrire les stades phénologiques

Nous avons suivi la méthode *Linked Open Terms*, une méthode d'ingénierie ontologique inspirée du développement logiciel agile [20]. Cette méthode se concentre sur la réutilisation d'éléments (classes, propriétés objet et de type de données) existants dans des ontologies précédemment publiées sur le Web de données liées. La méthode définit des itérations sur les quatre activités suivantes : (1) spécification d'exigences ontologiques, (2) implémentation d'ontologie, (3) publication d'ontologie et (4) maintenance d'ontologie.

4.1 Spécification

Nos exigences ontologiques ont été spécifiées par plusieurs questions de compétence. Ces questions expriment des besoins relatifs aux échelles phénologiques de la vigne mais elles peuvent être étendues à toute échelle d'une autre culture.

- Quels sont les labels français / anglais préférés et alternatifs pour un stade donné ?
- Quelle est la définition en français / anglais d'un stade donné ?
- Quels sont les stades principaux et secondaires d'une échelle donnée de la vigne (échelle individuelle BBCH, EL, Baggiolini, IFV-labels, IFV-Epicure) ?
- Quel stade principal de l'échelle individuelle BBCH est lié à un stade secondaire S de l'échelle R (EL, Baggiolini, IFV-labels, IFV-Epicure) ?
- Quel stade secondaire de l'échelle individuelle BBCH est équivalent à un stade secondaire S de l'échelle R (EL, Baggiolini, IFV-labels, IFV-Epicure) ?
- Quel stade secondaire de l'échelle individuelle BBCH est aligné avec un stade secondaire S de l'échelle R (EL, Baggiolini, IFV-labels, IFV-Epicure) ?
- Comment sont ordonnés (suivant / précédent) les stades principaux et secondaires d'une échelle donnée ?

Les alignements entre les échelles ont été spécifiés à l'aide d'un schéma organisant les stades par échelles phénologiques et par ordre chronologique. Ce schéma a d'abord été enrichi par un agronome : son but était de représenter les alignements trouvés dans la littérature. Par exemple dans la Figure 1 les lignes bleues claires représentent les alignements proposés par Coombe [7]. Les lignes rouges représentent les alignements proposés par Lorenz [14]. Les lignes vertes représentent les alignements proposés par Bloesh et Viret en 2008. Ensuite les deux agronomes ont proposé des alignements entre les 6 échelles. L'ontologie intervenait lors de leur discussion (1) pour clarifier le type d'alignements (exact, proche, etc...) entre deux stades, (2) pour proposer de nouveaux alignements déduits à partir des alignements précédents. Dans la figure 1, les traits noirs pleins représentent les alignements de type exact. Les traits noirs hachurés représentent les alignements de type proche.

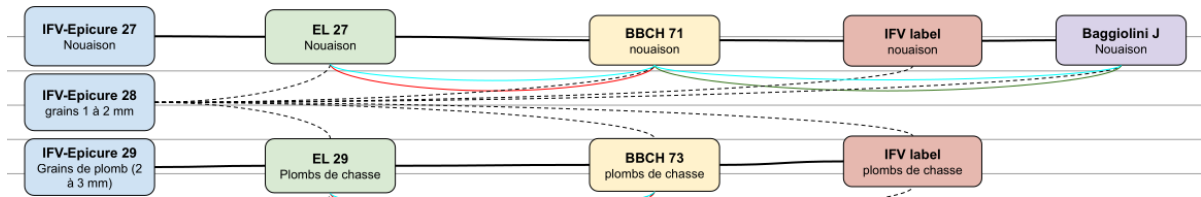


FIGURE 1 – Extrait du schéma de spécification des alignements

4.2 Implémentation

Nous avons choisi de spécialiser le modèle SKOS [1] pour représenter les stades phénologiques; SKOS est une recommandation du W3C. Ce modèle répond à la plupart de nos questions de compétence relatives à l’usage des labels et des alignements. L’ensemble des stades phénologiques sera une sous-classe de `skos:Concept`. Ainsi, un stade sera décrit avec des labels multilingues préférés et alternatifs. Un stade appartiendra à une échelle phénologique qui sera représentée comme une instance de la classe `skos:ConceptScheme`. Un stade secondaire sera lié à son stade principal par la propriété `skos:broader`. L’ontologie sera stockée dans un fichier OWL. L’ensemble des stades d’une échelle phénologique sera représenté par un ensemble d’instances de `skos:Concept` typées par les classes de l’ontologie.

4.2.1 Modèle de l’échelle générale BBCH

Chaque stade de l’échelle générale BBCH est représenté à la fois par une classe et par une instance; Une classe regroupe l’ensemble des stades similaires des échelles BBCH relative à une culture (échelle individuelle BBCH) plus le stade de l’échelle BBCH générale correspondant. Ainsi chaque stade d’une échelle BBCH sera donc une instance d’une de ses classes. Ce découpage en classe devrait permettre de retrouver facilement l’instance représentant le stade d’une culture donnée. Par exemple dans la figure 2, la classe intitulée `ppdo:SecondaryStage01` représente l’ensemble des stades secondaires BBCH 01 de toutes les échelles BBCH (générale ou individuelle). L’instance de cette classe intitulée `ppdo:bbch_secondaryStage_BBCH01` représente le stade de l’échelle BBCH générale. Cet individu est relié à l’instance de `skos:ConceptScheme` intitulée `ppdo:bbch_genericScale` par la propriété `skos:inScheme`. Comme le montre la figure 2, nous avons créé une classe `ppdo:GrowthStage`, sous-classe de `skos:Concept`. Cette classe est ensuite spécialisée en deux sous-classes, une classe pour les stades principaux et une autre pour les stades secondaires. Ces deux classes sont des classes définies pour expliciter le fait que chaque instance doit être liée à une instance de la classe `ppdo:StageDivision` par la propriété objet `ppdo:hasRank`. Ainsi toute instance de la classe `ppdo:GrowthStage` liée à l’individu `ppdo:principal`, sera automatiquement classée comme instance de la classe `ppdo:PrincipalGrowthStage` et toute instance de la classe `ppdo:GrowthStage` liée à l’individu

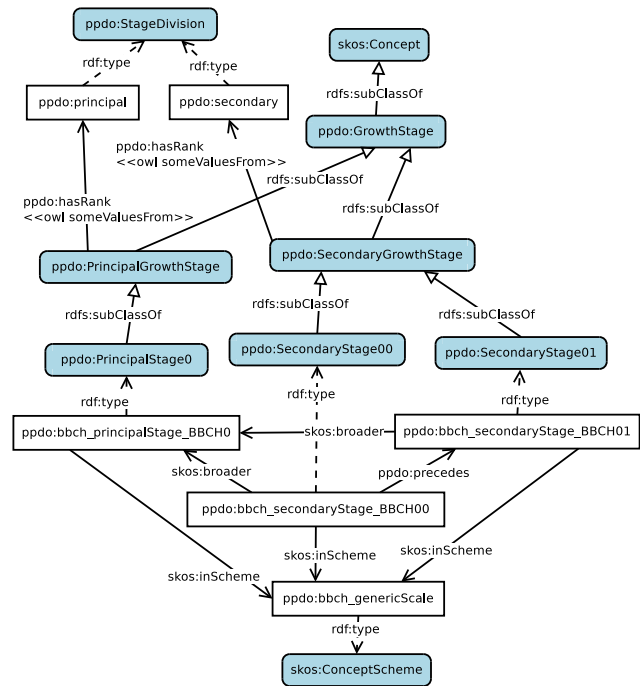


FIGURE 2 – Extrait de l’échelle générale BBCH – le préfixe `ppdo` est : <http://ontology.inrae.fr/ppdo/ontology/>.

`ppdo:secondary` sera automatiquement classée comme instance de la classe `ppdo:SecondaryGrowthStage`. La figure 2 présente des exemples de classes et d’instances représentant des stades principaux et secondaires de l’échelle générale BBCH.

4.2.2 Modèle de l’échelle individuelle BBCH

Un stade de la vigne est représenté par une instance d’une des sous-classes de `ppdo:GrowthStage`. Par exemple, la Figure 3 présente une instance de la classe `ppdo:SecondaryStage01`.

Plusieurs propriétés SKOS sont utilisées pour décrire un stade :

Les labels sont déclarés à l’aide des propriétés d’annotation `skos:prefLabel` et `skos:altLabel`. Les labels préférés sont extraits de la documentation (par exemple, la monographie BBCH). Les labels alternatifs sont extraits de la définition textuelle lorsque cela est possible ;

Les définitions sont déclarées à l’aide de la propriété d’annotation `skos:definition`. La monographie BBCH fournit une définition par langue ;

La propriété objet `skos:inScheme` indique l’échelle à la-

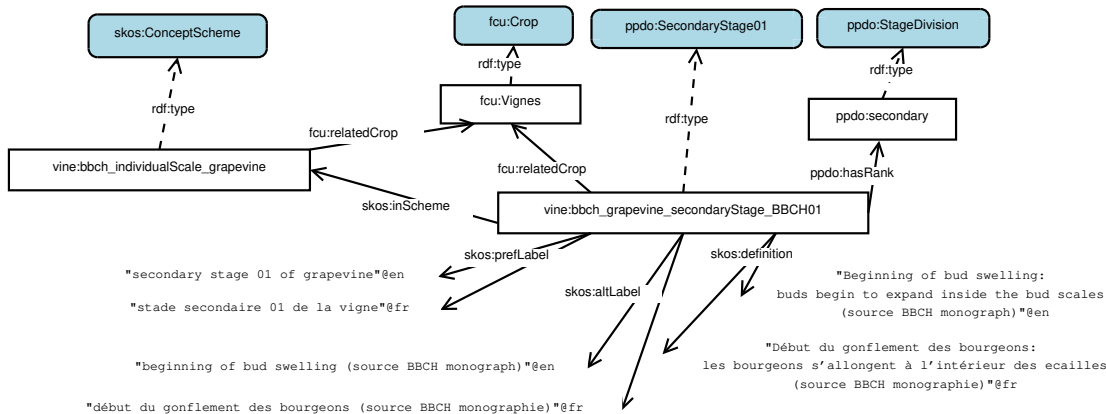


FIGURE 3 – Représentation du stade BBCH 01 de l'échelle individuelle BBCH –le préfixe `vine` signifie <http://ontology.inrae.fr/ppdo/resource/grapevine/>.

quelle appartient le stade. Les échelles individuelles BBCH ou toutes autres échelles sont représentées comme des instances de la classe `skos:ConceptScheme`.

La culture associée à un stade –ou à une échelle– est décrite à l'aide de la propriété objet `fou:relatedCrop` qui provient du thésaurus *FrenchCropUsage*,⁶ ce thésaurus a été créé pour annoter les BSV [21];

Les liens entre les stades principaux et secondaires sont décrits à l'aide des propriétés objet `skos:broader/narrower`. La figure 4 illustre l'utilisation de cette propriété entre le stade principal BBCH 0 et le stade secondaire BBCH 00 de l'échelle individuelle de la vigne;

L'ordre entre les stades, tel que défini dans l'échelle, est représenté avec les propriétés objet `ppdo:precedes` et `ppdo:follows`. La figure 4 présente un exemple de ce lien entre le stade secondaire BBCH 00 et le stade secondaire BBCH 01 de l'échelle individuelle de la vigne.

4.2.3 Modèle des échelles produites par IFV

Nous modélisons les deux ensembles de stades produits par IFV comme deux nouvelles échelles avec la même méthodologie que celle utilisée pour représenter l'échelle individuelle BBCH de la vigne. De plus, le point d'intérêt est l'alignement des échelles entre elles. Pour ce faire, nous avons utilisé les propriétés d'alignements fournies par SKOS :

Etant donné que les stades des échelles de l'IFV sont uniquement des stades secondaires, chaque stade des échelles de l'IFV est aligné à un stade principal de l'échelle individuelle BBCH de la vigne à l'aide de la propriété objet `skos:broadMatch`;

De même, chaque stade des échelles de l'IFV est aligné sur un stade secondaire de l'échelle individuelle BBCH à l'aide des propriétés `skos:exactMatch` ou `skos:closeMatch` ou `skos:broadMatch` ou `skos:narrowMatch`.

La figure 4 présente un exemple d'alignement entre un

stade de l'échelle IFV-labels et deux stades de l'échelle individuelle BBCH. La même approche est utilisée pour aligner les autres échelles (EL, Baggiolini).

4.2.4 Encodage RDF

Une fois les modèles conçus, ils ont été encodés sous forme de classes et d'instances. Nous avons utilisé l'éditeur d'ontologie Protégé (v5.1.0) [19] avec le plugin Cellfie.⁷ Nous avons d'abord créé le modèle général comme une `owl:Ontology` définissant les classes (Figure 2) à partir de l'échelle générale BBCH. Nous avons également ajouté des descriptions de métadonnées (par exemple, auteurs, dates, licences) comme recommandé par [10] ou [9]. Deuxièmement, nous avons peuplé l'ontologie avec des instances représentant tous les stades de l'échelle individuelle BBCH de la vigne, de l'échelle Baggiolini, de l'échelle EL et des deux échelles de IFV pour produire une base de connaissances.

Nous avons utilisé différentes sources pour peupler l'ontologie. Les monographies BBCH ont fourni les informations (labels et descriptions) en quatre langues (français, anglais, espagnol et allemand) pour l'échelle générale et l'échelle individuelle de la vigne. La publication scientifique de l'échelle de Baggiolini et de l'échelle Eichhorn-Lorenz a fourni les informations en anglais. Enfin, l'IFV nous a fourni deux fichiers CSV regroupant les informations en français et en anglais. Ensuite, nous avons extrait les données de manière semi-automatique dans différentes feuilles de calcul avant de les charger dans la base de connaissances à l'aide des règles de transformation Cellfie. Cellfie a également été utilisé pour générer les instances des classes; les règles de transformation ont créé les liens des propriétés objet : `ppdo:precedes`, `ppdo:follow`, `skos:broader`, `skos:broadMatch`,..., `skos:exactMatch`, `rdf:type`. Pour améliorer la cohérence de la base de connaissance finale, nous avons utilisé des règles SWRL pour déduire les propriétés inverses. Une

6. http://agroportal.lirmm.fr/ontologies/CROPUSAGE_FR (préfixe `fou` dans les figures)

7. Cellfie (<https://github.com/protegeproject/cellfie-plugin>) permet de transformer le contenu des feuilles de calcul en axiomes pour peupler une ontologie.

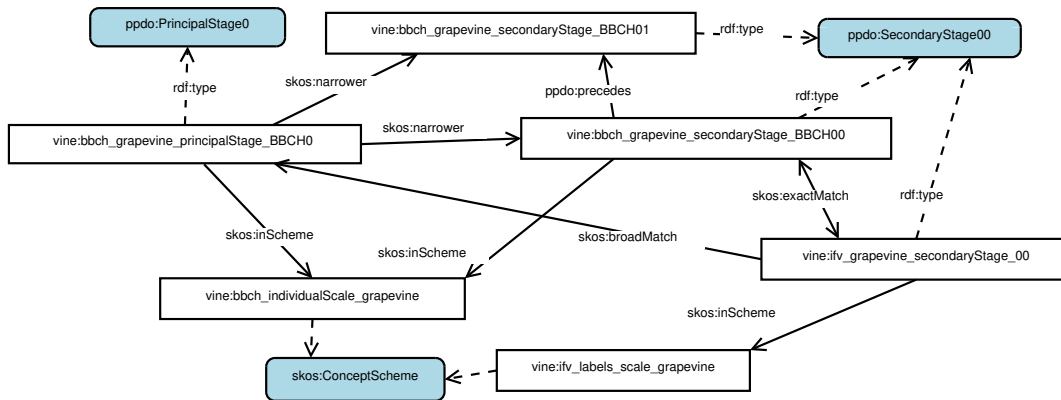


FIGURE 4 – Alignement entre le stade BBCH 01 de la vigne et le stade correspondant dans l'échelle IFV-Labels

dernière vérification a été effectuée à l'aide de l'outil SKOS Play!⁸ : il a permis de visualiser et de contrôler le modèle SKOS et de détecter quelques erreurs.

Le fichier OWL (en syntaxe Turtle) a été utilisé pour peupler un SPARQL endpoint. La conversion a été effectuée à l'aide des outils de la librairie *raptor*⁹. Le serveur SPARQL est jena-fuseki, de la fondation Apache¹⁰.

Nous parlons d'un cadre méthodologique pour décrire le résultat obtenu ; c'est à dire l'ontologie *BBCH based Plant Phenological Development Ontology (PPDO)* ainsi qu'un ensemble de base de connaissances (ou jeux de données) fondées sur cette ontologie, une par culture.¹¹ La version actuelle de l'ontologie PPDO est étiquetée v1.0. L'ontologie importe SKOS et PROV. Elle se compose de 129 classes et 124 propriétés. L'échelle générale de BBCH est représentée par une centaine d'individus. Les 5 échelles de la vigne sont représentées par 125 individus au total et 1208 assertions de propriété objet et 1696 assertions de propriété d'annotation.

4.3 Publication et maintenance de l'ontologie

L'ontologie PPDO est accessible au public sous différentes formes. Tout d'abord les fichiers OWL et les fichiers qui ont permis sa création sont disponibles dans un dépôt git <https://gitlab.irstea.fr/copain/phenologicalstages>. L'ontologie est également chargée dans un triple store et peut être interrogé sur : <http://ontology.inrae.fr/ppdo/sparql/>. L'ontologie est disponible sur AgroPortal [13] à l'adresse <http://agroportal.lirmm.fr/ontologies/PPDO>.¹² Le suivi des problèmes fourni par Gitlab est l'outil utilisé pour recevoir des commentaires et contrôler les modifications. Dans

8. <http://labs.sparna.fr/skos-play>

9. <http://librdf.org/raptor/>

10. <https://jena.apache.org/documentation/fuseki2/>

11. Produire une seule base de connaissances avec les instances de toutes les échelles de toutes les cultures n'aurait pas de sens pratique, les usages se faisant culture par culture.

12. A terme, AgroPortal stockera également les bases de connaissances par variété sous forme de vues ; mais à ce jour le portail ne permet pas de visualiser les instances ni de les regrouper dans différents `skos:ConceptScheme`.

un proche avenir, nous prévoyons d'ajouter de nouvelles bases de connaissances dédiées à de nouvelles cultures dans le SPARQL endpoint et le dépôt git.

5 Discussion et Conclusion

Cette transformation (RDFisation) des échelles phénologiques de la vigne est un travail collaboratif entre des experts de l'IFV et des ontologues. Le résultat améliorera l'usage et l'interopérabilité des échelles phénologiques actuellement existantes. Par exemple, les labels des échelles de l'IFV ont été améliorés car les labels français et les définitions ont été traduits en anglais grâce à la monographie BBCH. De plus, certaines incohérences et erreurs ont été détectées. Par exemple, l'échelle IFV-Labels mentionnait une référence au stade BBCH 88 de vigne qui n'existe pas dans la monographie BBCH. Au cours de la phase de spécification des alignements les experts de IFV ont décidé d'enrichir l'échelle IFV-Epicure en ajoutant de nouveaux stades. Les alignements entre les stades des différentes échelles ont été clarifiés par l'emploi des propriétés SKOS. Les 6 échelles ont été alignées entre elles alors qu'il n'existait jusqu'à présent que des alignements partiels entre les échelles prises deux à deux.

Nous avons peuplé l'ontologie PPDO avec toutes les échelles de la vigne préconisées par l'IFV. À l'avenir, nous prévoyons de publier des échelles relatives au blé (par exemple, Zadoks) en utilisant le même cadre méthodologique. Nous allons aussi publier les échelles individuelles BBCH de plusieurs légumes.

Le mélange de SKOS et d'OWL permet de regrouper les instances similaires dans une classe pour faciliter leur recherche ultérieure. L'ensemble des stades d'une échelle phénologique donnée est représenté dans un modèle SKOS. En effet, la notion de Scheme au sens de SKOS permet d'identifier une échelle phénologique et de lui associer des métadonnées spécifiques (auteurs, dates, documents sources, etc.). Le mélange des hiérarchies `owl:subClassOf` et `skos:broader/skos:narrower` permet de représenter respectivement des groupes de stades principaux et secondaires et des liens hiérarchiques entre les stades d'une même échelle.

La base de connaissance produite pour la vigne est actuellement disponible via un SPARQL endpoint à l'INRAE. À la fin du projet D2KAB, nous prévoyons de publier sur le Web de données liées notre archive BSV et les annotations associées. Ensuite, afin d'améliorer la découverte et l'interopérabilité de l'ontologie PPDO, nous alignerons formellement ses classes avec des concepts du thésaurus Agrovoc représentant ces stades (pour le moment nous avons déclaré simplement des liens `rdf:seeAlso`). Les alignements pourraient être complétés en alignant PPDO avec PPO. Le type d'alignement devra être étudié avec soin car les points de vue de ces deux ontologies divergent : PPO et PO représentent des processus de développement précis de chaque organe végétal. En revanche, les stades des échelles phénologiques ne sont qu'un instantané de certaines phases typiques de développement des plantes. Ces stades sont utiles pour planifier les pratiques agricoles.

Dans le cadre de l'évolution vers l'agriculture numérique, une autre perspective pour ce travail serait de rendre plus interopérables les données produites, par exemple en viticulture, par l'arrivée prochaine sur le marché de piquets connectés capables d'observations fines et répétées transmises en temps réel. Des prises de vue quotidiennes permettront assez prochainement de caractériser très finement la croissance et le développement des plantes voisines de ces postes d'observations par des algorithmes de reconnaissances de forme à partir d'images appropriées.

Remerciements

Ce travail a été réalisé avec le soutien du projet "Des Données aux Connaissances en Agronomie et Biodiversité (D2KAB—www.d2kab.org) financé par l'Agence Nationale de la Recherche (ANR-18-CE23-0017); ainsi qu'avec l'aide de l'Institut Français de la Vigne et du vin www.vignevin.com. Nous remercions également Thibaut Scholasch (Fruition Sciences) pour son aide à la traduction des labels.

Références

- [1] Miles Alistair and Bechhofer Sean. SKOS Simple Knowledge Organization System. W3C Recommendation, W3C, August 2009.
- [2] Robert Arp, Barry Smith, and Andrew D Spear. *Building ontologies with basic formal ontology*. MIT Press, 2015.
- [3] Sophie Aubin, Pierre Bisquert, Patrice Buche, Juliette Dibie, Liliana Ibanescu, Clement Jonquet, and Catherine Roussey. Recent progresses in data and knowledge integration for decision support in agri-food chains. In *30èmes Journées Francophones d'Ingénierie des Connaissances, IC'19*, pages 43–59, Toulouse, France, July 2019.
- [4] M Baggiolini. Les stades repères dans le développement annuel de la vigne et leur utilisation pratique. *Revue Romande d'Agriculture, de Viticulture et d'Arboriculture*, 1 :4–6, 1952.
- [5] M Baillod and M Baggiolini. Les stades repères de la vigne. *Rev. Suisse Vitic. Arboric. Hortic.*, 25(1) :7–9, 1993.
- [6] Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked Data - The Story So Far. *Semantic Web and Information Systems*, 5(3) :1–22, 2009.
- [7] Bryan G Coombe. Growth stages of the grapevine : adoption of a system for identifying grapevine growth stages. *Australian journal of grape and wine research*, 1(2) :104–110, 1995.
- [8] Laurel Cooper, Ramona L. Walls, Justin Elser, Maria A. Gandolfo, Dennis W. Stevenson, Barry Smith, Justin Preece, Balaji Athreya, Christopher J. Mungall, Stefan Rensing, Manuel Hiss, Daniel Lang, Ralf Reski, Tanya Z. Berardini, Donghui Li, Eva Huala, Mary Schaeffer, Naama Menda, Elizabeth Arnaud, Rosemary Shrestha, Yukiko Yamazaki, and Pankaj Jaiswal. The Plant Ontology as a Tool for Comparative Plant Anatomy and Genomic Analyses. *Plant and Cell Physiology*, 54(2) :e1, Dec. 2012.
- [9] Biswanath Dutta, Anne Toulet, Vincent Emonet, and Clement Jonquet. New Generation Metadata vocabulary for Ontology Description and Publication. In *11th Metadata and Semantics Research Conference, MTSR'17*, volume 755 of *CCIS*, pages 173–185, Tallinn, Estonia, Nov. 2017. Springer.
- [10] Daniel Garijo and M. Poveda Villalon. A checklist for complete vocabulary metadata. Technical report, WIDOCO, April 2017.
- [11] IFV. Les stades phénologique de la vigne, 2017.
- [12] P. Jaiswal, L. Cooper, J. L. Elser, A. Meier, M-A. Laporte, C. Mungall, B. Smith, E. KS. Johnson, M. Seymour, J. Preece, X. Xu, R. S. Kitchen, B. Qu, E. Zhang, E. Arnaud, S. Carbon, S. Todorovic, and D. Wm. Stevenson. Planteome : A resource for Common Reference Ontologies and Applications for Plant Biology. In *24th Plant and Animal Genome Conference, PAG'16*, San Diego, USA, January 2016.
- [13] Clement Jonquet, Anne Toulet, Elizabeth Arnaud, Sophie Aubin, Esther Dzalé Yeumo, Vincent Emonet, John Graybeal, Marie-Angélique Laporte, Mark A. Musen, Valeria Pesce, and Pierre Larmande. AgroPortal : a vocabulary and ontology repository for agronomy. *Computers and Electronics in Agriculture*, 144 :126–143, January 2018.
- [14] D.H. Lorenz, K.W. Eichhorn, H. Bleiholder, R. Klose, U. Meier, and E. Weber. Growth Stages of the Grapevine : Phenological growth stages of the grapevine (*Vitis vinifera* L. ssp. *vinifera*)—Codes and descriptions according to the extended BBCH scale. *Australian Journal of Grape and Wine Research*, 1(2) :100–103, July 1995.
- [15] Uwe Meier. *Entwicklungsstadien mono- und dikotyler Pflanzen : BBCH Monografie*. Open Agrar Repository, 2018.

- [16] Uwe Meier. *Etapas de desarrollo de las plantas monocotiledóneas y dicotiledóneas : BBCH Monografía*. Open Agrar Repositorium, 2018.
- [17] Uwe Meier. *Growth stages of mono- and dicotyledonous plants : BBCH Monograph*. Open Agrar Repositorium, 2018.
- [18] Uwe Meier. *Stades phénologiques des mono-et dicotylédones cultivées : BBCH Monographie*. Open Agrar Repositorium, 2018.
- [19] Mark A. Musen. The protégé project : a look back and a look forward. *AI Matters*, 1(4) :4–12, June 2015.
- [20] María Poveda-Villalón. A reuse-based lightweight method for developing linked data ontologies and vocabularies. In *9th Extended Semantic Web Conference, ESWC'12*, volume 7295 of *LNCS*, pages 833–837. Springer, 2012.
- [21] Catherine Roussey, Stephan Bernard, Francois Pinet, Xavier Reboud, Vincent Cellier, Ivan Sivadon, Danièle Simonneau, and A. Bourigault. A methodology for the publication of agricultural alert bulletins as LOD. *Computers and Electronics in Agriculture*, 142 :632–650, Nov. 2017.
- [22] Catherine Roussey, Xavier Delpuech, Florence Amardeilh, Stephan Bernard, and Clement Jonquet. Semantic Description of Plant Phenological Development Stages, starting with Grapevine. In *14th International Conference on Metadata and Semantics Research (MISR2020)*, Madrid, Spain, November 2020.
- [23] Johannes Keizer Sachit Rajbhandari. The AGROVOC Concept Scheme ; A Walkthrough. *Integrative Agriculture*, 11(5) :694–699, May 2012.
- [24] Rosemary Shrestha, Elizabeth Arnaud, Ramil Mautleon, Martin Senger, Guy F. Davenport, David Hancock, Norman Morrison, Richard Bruskiewich, and Graham McLaren. Multifunctional crop trait ontology for breeders' data : field book, annotation, data discovery and semantic enrichment of the literature. *AoB Plants*, 2010, May 2010.
- [25] Brian J. Stucky, Rob Guralnick, John Deck, Ellen G. Denny, Kjell Bolmgren, and Ramona Walls. The Plant Phenology Ontology : A New Informatics Resource for Large-Scale Integration of Plant Phenology Data. *Frontiers in Plant Science*, page 517, May 2018.
- [26] Ramona L. Walls, Laurel Cooper, Justin Elser, Maria Alejandra Gandolfo, Christopher J. Mungall, Barry Smith, Dennis W. Stevenson, and Pankaj Jaiswal. The Plant Ontology Facilitates Comparisons of Plant Development Stages Across Species. *Frontiers in Plant Science*, 10 :631, June 2019.