



HAL
open science

L'Écosystème R-IDEA 1.0: Un environnement de calcul automatisé pour la mise en œuvre de la méthode IDEA V4

David Carayon

► **To cite this version:**

David Carayon. L'Écosystème R-IDEA 1.0: Un environnement de calcul automatisé pour la mise en œuvre de la méthode IDEA V4. [Rapport Technique] INRAE. 2020, pp.24. hal-02983957v1

HAL Id: hal-02983957

<https://hal.inrae.fr/hal-02983957v1>

Submitted on 30 Oct 2020 (v1), last revised 2 Feb 2021 (v2)

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.



Distributed under a Creative Commons Attribution 4.0 International License

L'Écosystème R-IDEA

Un environnement de calcul automatisé pour la mise en œuvre de la méthode IDEA V4

-

Version 1.2 du 03 septembre 2020

INRAE

Institut national de recherche pour
l'agriculture, l'alimentation et l'environnement



David Carayon

Document distribué sous licence CC-BY



<https://creativecommons.org/licenses/by/4.0/fr/legalcode>

1 - Introduction	1
1.1 - Une nouvelle méthode...	1
1.2 - ... Qui demande de nouveaux outils	1
2 - Développement des outils : Le package R IDEATools	3
2.1 - Description des modules	4
2.1.1 - ImportIDEA ()	4
2.1.1.1 - Aspects techniques	4
2.1.1.2 - Pré-calcul à l'importation	5
2.1.2 - MakeTrees ()	7
2.1.3 - dimensionsPlots ()	8
2.1.4 - radarPlots ()	9
2.1.5 - metaIDEA ()	10
2.1.6 - exportIDEA ()	11
2.1.7 - Programmation complète d'un diagnostic	12
3 - Mise à disposition des outils : ShinyIDEA	13
3.1 - Développement de l'interface web	13
3.1.1 - L'analyse individuelle	15
3.1.2 - L'analyse de groupe	18
3.2 - Hébergement serveur	19
3.3 - Apport à la plateforme WEB-IDEA (BSA)	20
4 - Stratégie de captation des données	21
4.1 - Objectifs	21
4.2 - La base de données PostgreSQL « IDEA_V4 »	21
4.2.1 - La table "evenement"	22
4.2.2 - La table "exploitation"	22
4.2.3 - La table "label_metadonnees"	22
4.2.4 - La table "notes_idea"	22
4.2.5 - La table "label_indicateurs"	23
5 - Références	24

1 - Introduction

1.1 - Une nouvelle méthode...

La nouvelle version de la méthode IDEA (**I**ndicateurs de **D**urabilité des **E**xploitations **A**gricoles) introduit un tout nouveau cadre conceptuel qui s'ancre dans le champ de la durabilité forte, de la multifonctionnalité et prend en compte les enjeux globaux d'une agriculture durable. Ce cadre conceptuel a permis de construire 53 indicateurs permettant d'analyser la durabilité de l'exploitation agricole selon deux approches complémentaires (Zahm et al., 2020).

La première évalue la durabilité en organisant ces 53 indicateurs selon les 3 dimensions normatives du développement durable (agroécologique, socio-territoriale, économique), structurées en 13 composantes ; l'évaluation repose sur un système de notation basé sur 100 unités de durabilité pour chacune des 3 dimensions qui ne se compensent pas entre elles. La seconde évalue la durabilité en organisant les 53 indicateurs selon les 5 propriétés des systèmes agricoles durables (autonomie, robustesse, capacité productive et reproductive de biens et services, ancrage territorial et responsabilité globale) qui sont structurées de manière arborescente en 15 branches ; l'agrégation des indicateurs y suit une démarche qualitative et hiérarchique mobilisant l'outil DEXi.

Le potentiel pédagogique du concept de propriétés des systèmes favorise une approche transdisciplinaire de l'exploitation agricole. À la suite de ses trois précédentes versions, IDEA 4 renouvelle son potentiel d'usage pour accompagner la transition agroécologique (Zahm et al., 2019).

1.2 - ... Qui demande de nouveaux outils

Là où les précédentes versions de la méthode IDEA pouvaient être mises en œuvre "à la main", ce qui présentait un certain intérêt pédagogique, cette nouvelle version se veut plus complète et, par conséquent, bien plus difficile à mettre en œuvre sans des outils automatisés proposés par les concepteurs de la méthode.

L'approche par les dimensions étant la plus proche des anciennes version d'IDEA, basée sur un système de scoring et d'agrégations plafonnées, la technologie adoptée est restée la même : Un calculateur développé sous Microsoft Excel permettant la saisie de données, le calcul automatique des notes IDEA de l'exploitation et l'édition des résultats sous la forme de graphiques simples.

L'approche par les propriétés est quant à elle une agrégation qualitative multi-critères suivant une carte heuristique présentant 46 nœuds et dont les règles de décision ont été définies nœud par nœud à l'aide du logiciel DEXi. Implémenter cette méthodologie complexe dans un environnement logiciel Excel opérationnel et accessible à l'ensemble des publics concernés par la méthode IDEA4 s'est rapidement révélée très délicat pour son usage par le plus grand nombre

(enseignants, élèves, conseillers agricoles). Elle demanderait, pour l'analyse de plusieurs diagnostics, plusieurs heures d'actions répétitives et d'allers et retours entre deux logiciels (avec à chaque fois un risque d'erreur humaine/technologique) pour le diagnostic d'une seule exploitation.

Outre ces besoins techniques, la méthode IDEA V4 accusait un besoin d'automatisation, de fiabilité et se devait de proposer des restitutions graphiques modernes. Ainsi, les travaux de la fin d'année 2019 et 2020 se sont consacrés au développement d'outils informatiques permettant de :

- Mettre en œuvre concrètement les nouvelles approches proposées par ce nouveau cadre conceptuel
- Rendre l'ensemble de la procédure fiable et automatisable
- Proposer de nouveaux modes de représentation des résultats à la fois précis, synthétiques et pédagogiques

Ces développements se sont appuyés en majeure partie sur le gratuiciel "R" et ses nombreuses librairies open-source.

Le présent rapport présente cet "écosystème" informatique automatisé développé dans le cadre de ces travaux.

2 - Développement des outils : Le package R IDEATools

IDEATools est un programme développé sur le logiciel R (sous forme de package) dont l'objectif est de proposer des outils et ressources mobilisables sous R afin de manipuler des données issues d'enquêtes réalisées dans le cadre de la méthode IDEA4 (Figure 1). Cette application a d'abord été développée afin de répondre aux « besoins techniques » soulevés par la mise en place de l'évaluation par les propriétés de la durabilité.

Afin de répondre à ces besoins, l'ensemble des règles de catégorisation et des 46 tables de décisions basées sur DEXi ont été implémentée sous forme d'algorithmes en langage R dans IDEATools. Cet algorithme propose l'évaluation des 5 propriétés d'une exploitation agricoles en seulement quelques secondes et nécessite seulement de disposer du calculateur "approche par les dimensions" correctement rempli pour une exploitation agricole donnée. En plus de cette évaluation, l'algorithme propose également d'éditer les arbres éclairés (issus de la carte heuristique) en quelques secondes.

Au delà de cette approche par les propriétés, des modules ont ensuite été développés pour éditer des sorties graphiques similaires à celles proposées par le calculateur (excel), mais améliorées grâce aux bibliothèques graphiques disponibles sous R telles que {ggplot2}.

Il est donc possible, grâce à cette première version d'IDEATools, de produire des graphiques relatifs à l'approche par les dimensions (score des dimensions, des composantes, des indicateurs) ainsi qu'à l'approche par les propriétés (arbres éclairés, diagrammes radars) pour 40 exploitations (40 calculateurs) en moins d'une minute.

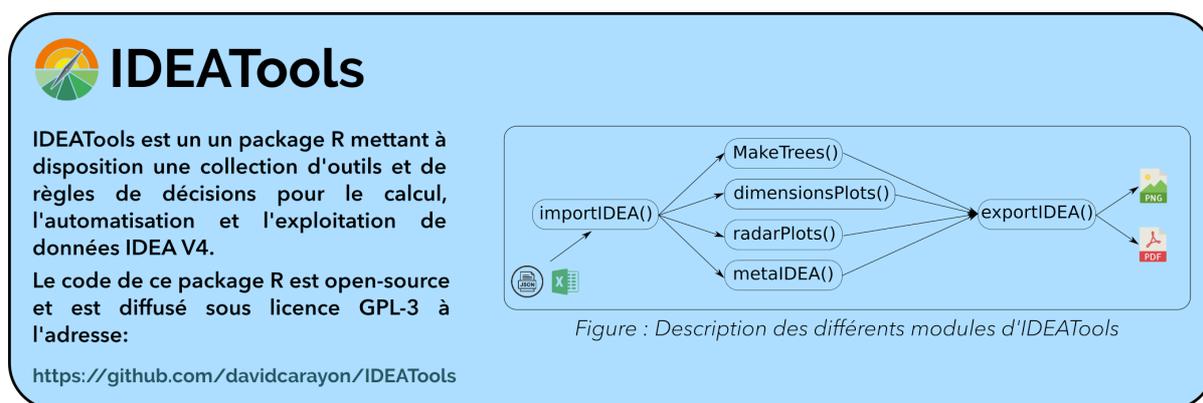


Figure 1 : Présentation d'IDEATools

IDEATools est un programme libre pouvant être modifié ou redistribué selon les termes de la GNU General Public License (GPL) telle que publiée par la Free Software Foundation; soit la version 3 de la licence. Ce programme est diffusé via la forge logicielle GitHub (<https://github.com/davidcarayon/IDEATools>) dans l'espoir qu'il sera utile, mais sans aucune garantie de qualité marchande ou d'adéquation à un but particulier.

Pour citer ce programme : Carayon, D., Girard, S., Zahm, F. (2020). IDEATools: Un applicatif pour le calcul, l'automatisation et l'exploitation de données IDEA4. R package version 1.0.

2.1 - Description des modules

2.1.1 - ImportIDEA()

Ce premier module sert à l'import des données issues d'un calculateur IDEA4.

2.1.1.1 - Aspects techniques

Les formats d'entrée acceptés sont le format natif du calculateur (.xls), mais également le format Excel plus récent (.xlsx) ainsi qu'un format de fichier standardisé directement exportable depuis les dernières versions du calculateur > 4.2.1 (.json).

La logique appliquée par l'algorithme pour récupérer les différentes données est la suivante :

- S'il s'agit d'un fichier .json, alors `jsonlite::fromJSON()` permet d'importer le fichier sans autres besoins de transformations.
- S'il s'agit d'un fichier Excel (.xls ou .xlsx), utilisation de `readxl::read_excel()`:
 - Si le numéro de version indiqué dans l'onglet 'Notice' est supérieur à la version 4.2.0, alors l'algorithme ira chercher les données dans l'onglet 'Renvoi BDD'. La procédure est dans ce cas identique à l'import JSON.
 - Si le numéro de version est inférieur à 4.2.0 ou inexistant, alors un algorithme alternatif ira chercher les métadonnées dans l'entête de l'onglet "Saisie et calculateur", puis les scores déplafonnés dans les bilan de chaque dimension et les résultats plafonnés dans l'onglet du bilan global de la durabilité. Selon l'ancienneté dans le développement (2017 à 2020) du calculateur inséré, il est donc possible que certaines données soient manquantes (en particulier les métadonnées).

Les imports issus de fichiers Excels ont été standardisés pour être identiques à l'import JSON. Ainsi, quel que soit le format d'entrée, le fonctionnement des modules suivants reste toujours identique.

Attention : L'import sera interrompu dès que la moindre donnée relative aux indicateurs ou à certaines métadonnées clés (ex : élevage) sera manquante.

En ce qui concerne le nom de l'exploitation (ou l'id si le format est de type json ou calculateur Excel > 4.2.0), une absence d'information sera traduite par la création d'un code aléatoire à 5 lettres. Si l'utilisateur décide de forcer un import anonyme (en utilisant le second argument du module, `anonymous` qui est un booléen), alors ce code aléatoire sera affecté à l'exploitation même si il existe un identifiant renseigné.

Dans le cadre du développement de ce package, 3 fichiers de tests (correspondants aux 3 possibilités d'import mentionnées plus haut) ont été utilisés et sont accessibles dans le répertoire source du package :

```
# Renvoie la localisation du fichier. Ce chemin peut directement être inséré
  dans le module d'import IDEA
chemin <- file.path("example_json.json", package = "IDEATools")
IDEAdata <- importIDEA(chemin)
```

En plus d'un import "simple", ce module permet également d'importer un répertoire de calculateurs entier, mêlant les 3 types de fichiers de données (xls, xlsx et json).

Dans ce module, comme pour tous les autres, le traitement de fichiers multiples a été résolu par le développement d'un sous-module de traitement simple (un fichier à la fois) qui est appliqué de manière itérative sur chaque fichier grâce aux fonctionnalités du package `{purrr}`.

Les autres packages R mobilisés dans ce module sont :

- `{readxl}` pour la lecture de fichiers excels et `{jsonlite}` pour les fichiers .json
- `{dplyr}`, `{tidyr}`, `{stringr}` et `{janitor}` pour la manipulation de données
- `{purrr}` pour les traitement itératifs dans le cas de traitements multiples

2.1.1.2 - Pré-calcul à l'importation

Les calculateurs Excel sont uniquement développés pour traiter l'approche par les dimensions. De ce fait, à l'import de ces fichiers, aucun résultat n'est disponible pour l'approche par les propriétés. Plutôt que de développer un nouveau module qui procéderait à cette évaluation, il semblait plus judicieux d'implémenter cette évaluation dès l'importation. De ce fait, en fin d'import, toutes les données liées à l'évaluation sont disponibles, les modules suivants ne se concentrent que sur les restitutions graphiques.

Cette évaluation par les propriétés repose sur deux types de tables qui ont été construites par le CS IDEA et qui sont stockées en tant que données internes du package. Ces règles de décisions sont donc tout à fait consultables (et prochainement modifiables) par un utilisateur du package.

- Une table fixant les seuils permettant d'attribuer une des quatre modalités "Favorable", "Intermédiaire", "Défavorable" ou "Très défavorable" à chacun des 53 indicateurs :

indicateur	nom_indicateur	TD	D	I	F
A1	Diversité des espèces cultivées	NA	0	3	4
A2	Diversité génétique	NA	0	2	4
A3	Diversité temporelle des cultures	NA	0	3	5
A4	Qualité de l'organisation spatiale	NA	0	3	5
A5	Gestion des insectes pollinisateurs et des auxiliaires des cultures	NA	0	2	4
A6	Autonomie en énergie, matériaux, matériels, semences et plants	NA	0	3	5

- L'ensemble des 46 tableaux de contingences élaborés par le biais du logiciel DEXi, puis également stockés sous forme de liste sous R :

```
data("decision_rules_total")
names(decision_rules_total)
#> [1] "node_1" "node_2" "node_3" "node_4" "node_5" "node_6" "node_7"
#> [8] "node_8" "node_9" "node_10" "node_11" "node_12" "node_13" "node_14"
#> [15] "node_15" "node_16" "node_17" "node_18" "node_19" "node_20" "node_21"
#> [22] "node_22" "node_23" "node_24" "node_25" "node_26" "node_27" "node_28"
#> [29] "node_29" "node_30" "node_31" "node_32" "node_33" "node_34" "node_35"
#> [36] "node_36" "node_37" "node_38" "node_39" "node_40" "node_41" "node_42"
#> [43] "node_43" "node_44" "node_45" "node_46"
```

```
head(decision_rules_total$node_1)
```

A1 - Diversité des cultures	A3 - Diversité temporelle des cultures	A4 - Qualité spatiale du territoire	Diversité de l'organisation spatiale et temporelle
défavorable	défavorable	défavorable	très défavorable
défavorable	défavorable	intermédiaire	très défavorable
défavorable	défavorable	favorable	défavorable
défavorable	intermédiaire	défavorable	très défavorable
défavorable	intermédiaire	intermédiaire	défavorable
défavorable	intermédiaire	favorable	défavorable

La numérotation des noeuds a été réalisée selon la Figure 2 :

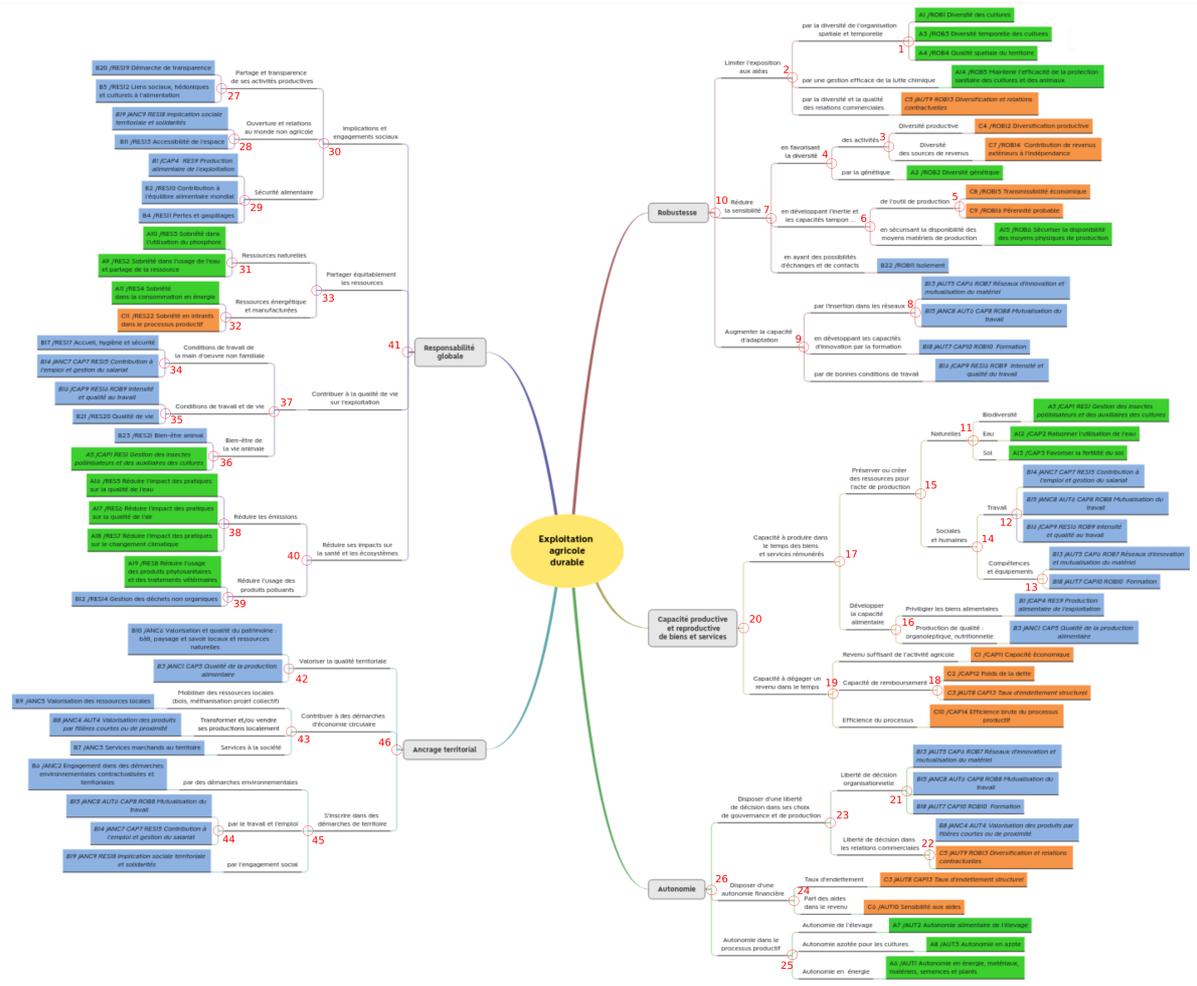


Figure 2 : Carte heuristique complète avec la numérotation des noeuds employée dans le code source d'IDEATools

2.1.2 - MakeTrees ()

Ce module permet la construction de cartes heuristiques colorées à partir de données importées par le module précédent. Différentes approches faisant appel à la librairie {ggplot2} ont d'abord été explorées pour cette fonctionnalité, avant d'être remplacées par une procédure plus complexe s'appuyant sur du dessin vectoriel .svg.

Des modèles (ou canvas) non colorés ont d'abord été dessinés sous le logiciel Inkscape (Figure 3). Un identifiant unique a été assigné à chaque rectangle afin de pouvoir lier chaque rectangle à l'indicateur qu'il représente. Le code source svg (type XML) de chacun de ces canvas a été sauvegardé sous forme de liste comme ressource interne de ce package.

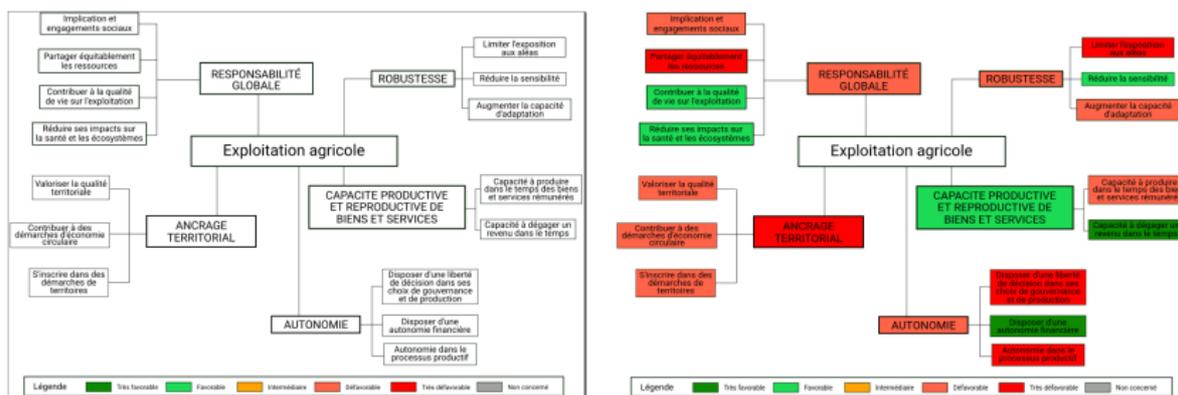


Figure 3 : Modèle vide (gauche) puis coloré (droite) de l'arbre éclairé synthétique d'une exploitation

L'algorithme va ensuite détecter dans le code source de chaque canvas les balises relatives aux rectangles (<rect[...]/>), y chercher l'identifiant de l'objet et puis la ligne relative au remplissage de l'objet (style="fill:#ffffff") avant de remplacer ce code hexadécimal par la couleur correspondant à l'évaluation proposée par IDEA de l'indicateur concerné.

Pour information, une balise de rectangle au format svg s'écrit comme suit :

```
<rect
  style="fill:#ffffff;fill-opacity:1;stroke:#001800;stroke-width:0.14778921"
  id="rect2-16"
  width="17.833706"
  height="4.9643545"
  x="370.30774"
  y="143.9761"
  inkscape:label="#rect2-16" />
```

Ce code source "modifié" est stocké dans la liste renvoyée par ce module. Celui-ci sera ensuite transformé en image réelle (png/pdf) par le dernier module d'export.

Les packages ici utilisés sont principalement {stringr} pour la manipulation des chaînes de caractères puis {dplyr} pour la manipulation des tableaux de manière générale. Enfin, {purrr} est ici aussi utilisé dans le cas du traitement simultané de plusieurs fichiers de données.

2.1.3 - dimensionsPlots ()

De la même manière que le module de production des cartes heuristiques, ce module s'appuie sur les données importées par le premier module. En revanche, celui-ci est très majoritairement basé sur la librairie {ggplot2} pour la production de 5 graphiques relatifs à l'approche par les dimensions (Figure 4):

- Vision globale des 3 dimensions
- Vision des 13 composantes
- 3 graphiques de vision des indicateurs de chaque dimension, regroupés par composante.

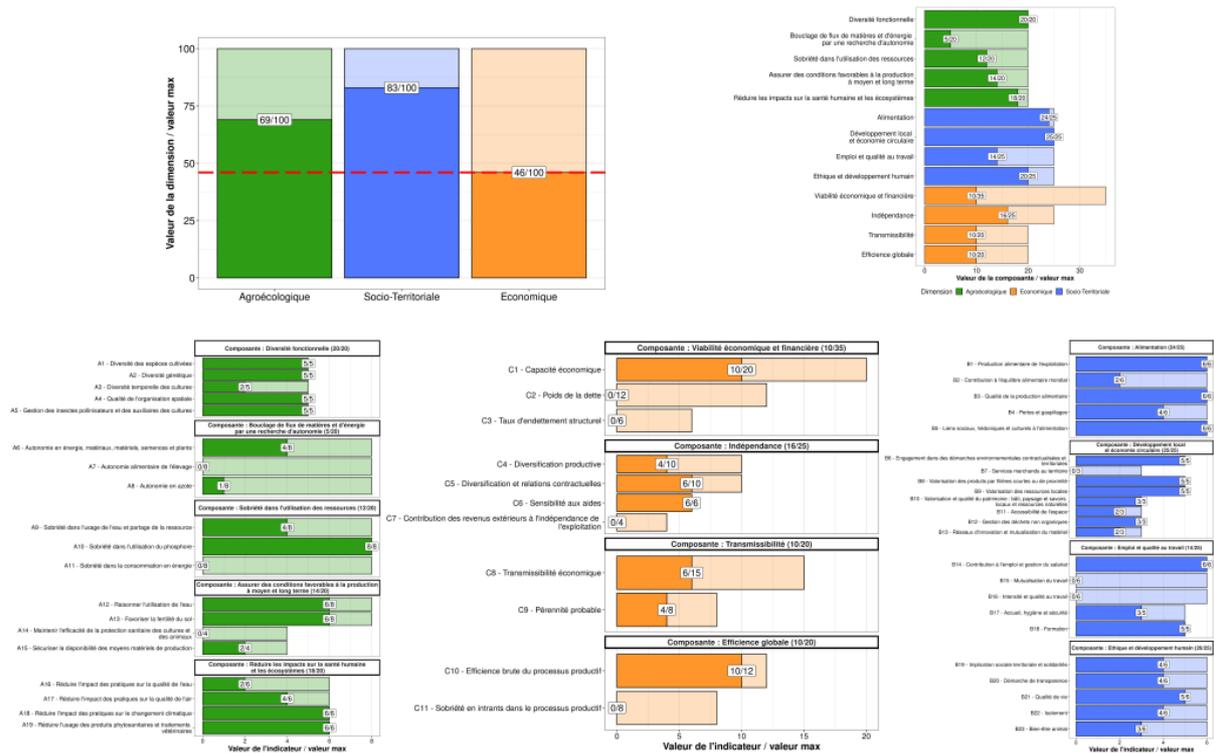


Figure 4 : Exemple de graphiques produits par ce module

Ce module renvoie une liste de ggplots, qui pourront être exportés en fichier concret (png/pdf) par le module d'export.

Ici encore, `{purrr}` est utilisé dans le cas du traitement simultané de plusieurs fichiers de données. `{dplyr}` et `{tidyr}` pour le remaniement des données.

2.1.4 - radarPlots ()

Derrière l'intitulé de "radar", ce ne sont pas des diagrammes radar au sens strict qui sont réalisés, mais plutôt des diagrammes en barres ayant subi une rotation polaire. En effet, les radars sont de moins en moins utilisés car il est admis que leur interprétation (notamment en termes d'aire) peut totalement varier selon le positionnement des variables et créer des différences d'interprétation trompeuses selon la façon dont chaque utilisateur construit son radar.

Ce module est très similaire au module précédent car il est lui aussi basé sur `{ggplot2}`.

Cette représentation graphique circulaire n'étant pas compatible avec un affichage du libellé complet des indicateurs, une table a été adjointe à chaque graphique afin de traduire chaque code indicateur en libellé complet. Les couleurs des lignes de ces tables ont été programmiquement colorées pour correspondre à la dimension d'appartenance de l'indicateur (Figure 5).

Pour ces graphiques, les scores de chaque indicateur sont standardisés en % de la note maximale afin d'être représentés sur une échelle commune.

Ce module renvoie une liste de ggplots, qui pourront être exportés en fichier concret (png/pdf) par le module d'export.

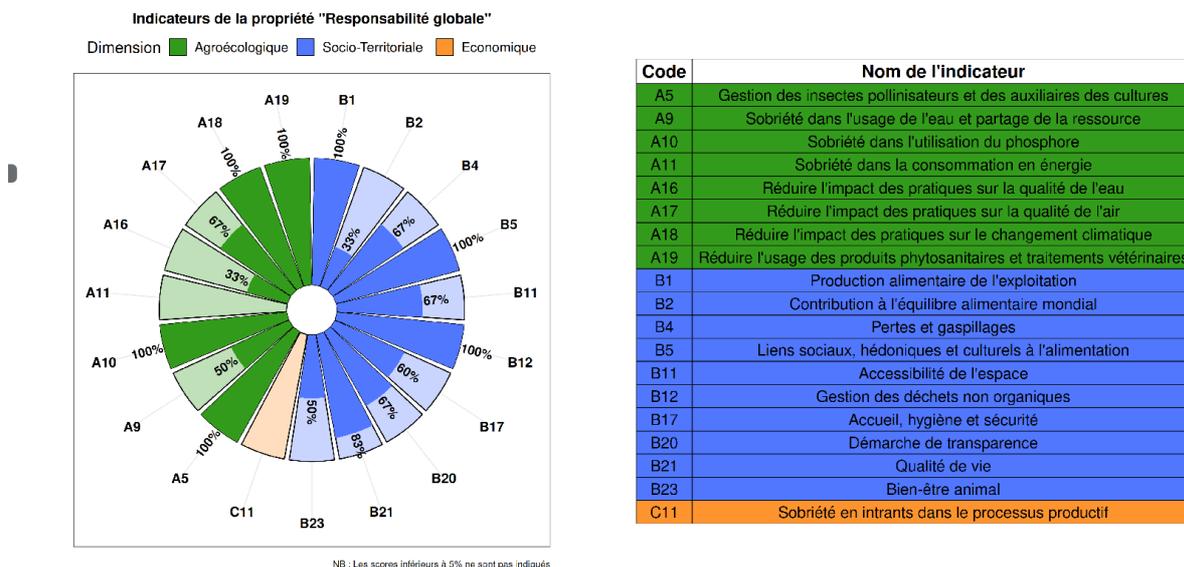


Figure 5 : Exemple de radar produit pour la propriété 'Responsabilité globale'

Le package {purrr} a de nouveau été utilisé dans le cas du traitement simultané de plusieurs fichiers de données. {dplyr} et {tidyr} pour le remaniement des données.

2.1.5 - metaIDEA()

Ce module est le plus récemment développé et est également basé sur {ggplot2}. Il est amené à fortement évoluer.

Ce module est dédié à l'analyse de collectifs et n'est donc utilisable que dans une situation d'import de plusieurs fichiers en simultané. Par soucis de cohérence, le nombre minimal de fichiers a été fixé à 3.

Une matrice colorée (ou heatmap) est proposée pour l'approche par les propriétés, ainsi qu'un diagramme en barres empilées pour l'approche par les dimensions (Figure 6).

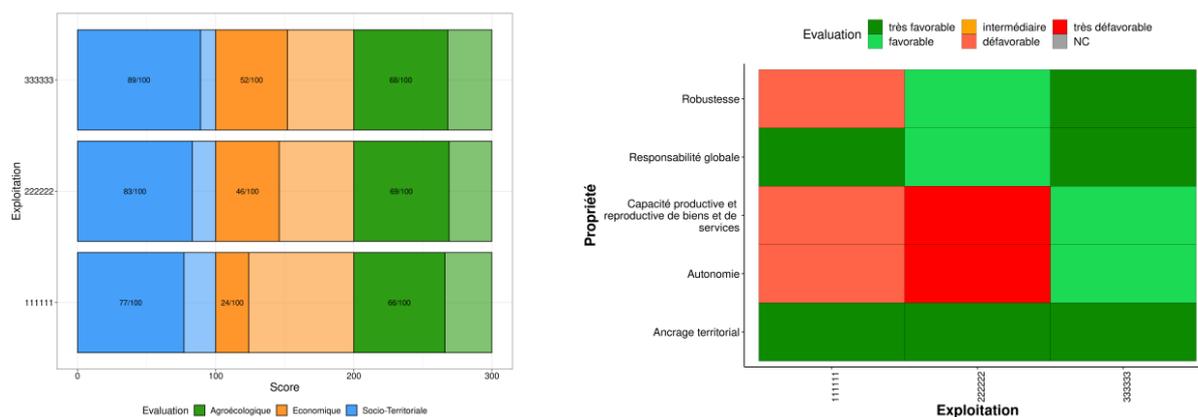


Figure 6 : Exemples d'un diagramme en barres empilées et d'une heatmap pour un groupe de 3 exploitations

Ce module est encore en version provisoire, et sera développé plus en profondeur au fil des réflexions des concepteurs de la méthode sur les analyses de groupe à proposer.

Ce module renvoie une liste de ggplots, qui pourront être exportés en fichier concret (png/pdf) par le module d'export.

Une fois encore, `{tidyr}` et `{dplyr}` ont été utilisés pour la manipulation des données.

2.1.6 - exportIDEA()

Ce dernier module permet d'exporter les résultats issus des 4 précédents modules sous forme de fichiers graphiques (svg, png ou pdf).

Le module d'export va lancer différents algorithmes en fonction du type de résultats qui lui sont injectés :

- S'il s'agit d'une liste de ggplots (produits par `dimensionsPlots()`, `radarPlots()` ou `metaIDEA()`), alors les différents graphiques vont être exportés dans le dossier défini par l'utilisateur dans l'argument `outdir`. Par défaut, si aucune indication n'est donnée par l'utilisateur, ce dossier de sortie prendra la forme "EXPORT_{date}". A noter que les résultats seront ensuite exportés dans un sous-dossier correspondant à l'identifiant de l'exploitation (ce qui permet de séparer les résultats dans le cas d'un traitement multiple).
- S'il s'agit de code source svg produit par `MakeTrees()`, alors les graphiques sont exportés par des fonctions issues du package `{rsvg}`, à savoir `rsvg::rsvg_pdf` et `rsvg::rsvg_png`.

Les dimensions d'export de chaque graphique ont été choisies et implémentées dans une liste interne au package. L'export est ensuite réalisé par l'application des fonctions d'export (`ggplot2::ggsave()` ou `rsvg`) combinées à l'approche itérative de `{purrr}` pour les traitements simultanés.

2.1.7 - Programmation complète d'un diagnostic

Voici un exemple concret de diagnostic complet à partir d'un fichier de données JSON.

```
# Installation
devtools::install_github("davidcarayon/IDEATools")

# Chargement
library(IDEATools)

# Chemin du fichier d'entrée
chemin <- file.path("example_json.json", package = "IDEATools")

# Diagnostic complet
IDEAdata <- importIDEA(chemin)
dimensionsPlots(IDEAdata) %>% exportIDEA()
MakeTrees(IDEAdata) %>% exportIDEA()
radarPlots(IDEAdata) %>% exportIDEA()
```

3 - Mise à disposition des outils : ShinyIDEA

Le package R IDEATools est aujourd’hui opérationnel et diffusé en ligne. Un utilisateur novice de R peut donc aujourd’hui facilement ré-utiliser ces outils depuis son propre ordinateur et réaliser un diagnostic complet IDEA (dimensions, propriétés, arbres éclairés, etc.) avec seulement 7 lignes de codes (cf. section 2.1.7). Ce format de diffusion (package R) lève en partie les inconvénients techniques introduits par la 4^e version de la méthode IDEA, mais seulement à condition de manipuler le logiciel R, ce qui ne contribue pas encore à rendre la méthode opérationnelle en dehors d’un cadre universitaire.

De nouveaux développements ont donc été entrepris pour mettre les outils IDEATools à disposition de tous les utilisateurs, quelque soit leur niveau de maîtrise des outils informatiques. La technologie privilégiée est encore R et le développement s’appuie cette fois sur un package open-source particulier : {shiny}.

Shiny est un cadre conceptuel de développement (ou framework) permettant le développement d’applications interactives pour le web tout en profitant de la puissance calculatoire et des nombreuses bibliothèques graphiques de R. En résumé, L’application va être codée en langage R, via une syntaxe particulière, puis shiny va convertir ce code en un mélange d’HTML (langage de balisage) et de JavaScript (langage de l’interactivité web). Il est également possible d’insérer directement des balises HTML ou des scripts JS dans le code source pour les plus expérimentés, ce qui rend l’application totalement personnalisable et sans limite de développement.

Les outils IDEATools ont donc été embarqués dans une application web-interactive portant le nom de “ShinyIDEA” (Figure 7).



ShinyIDEA

ShinyIDEA est une interface web s'appuyant sur la librairie R "shiny" afin de mettre à disposition de tout utilisateur les outils et restitutions graphiques proposées par IDEATools.

L'application propose d'importer un ou plusieurs calculateurs IDEA V4, d'explorer les résultats par le biais de différentes ingographies interactives, puis d'exporter les résultats sous forme d'images ou de rapports automatiques.

https://lien_vers_l_application.fr



The screenshot shows a web interface with a sidebar on the left containing navigation options like 'Dimension ou propriété à afficher', 'Choisir le type de représentation', and 'Afficher les résultats...'. The main area displays several horizontal bar charts with green bars, each representing a different agroecological indicator. The charts are organized into sections with titles like 'Analyse des conditions favorables à la production d'oxygène', 'Analyse des conditions favorables à la production de matière organique', and 'Analyse des conditions favorables à la production de carbone'. Each chart has a y-axis with labels like A101, A102, A103, A104, A105 and an x-axis with values from 0 to 6. A legend at the bottom right explains the color coding: 'Note: couleur indiquant le niveau indicateur (vert ou orange) et la qualité correspondante du graphique'.

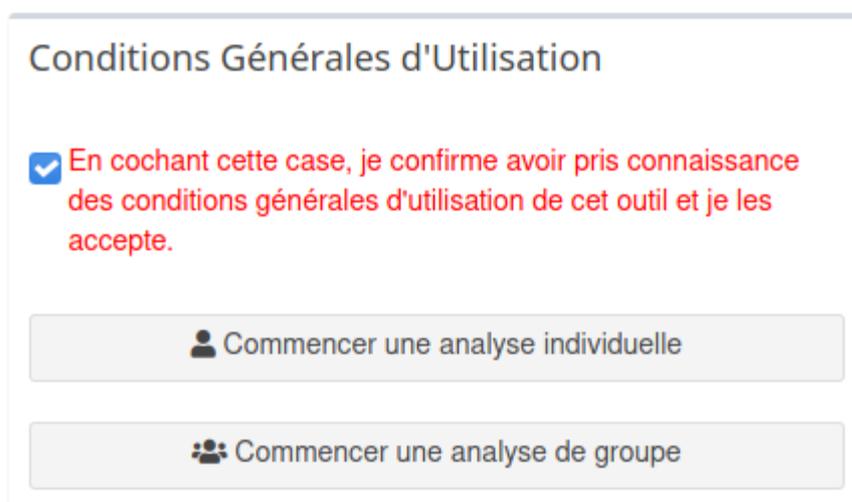
Figure : Exemple de restitution des indicateurs agroécologiques proposées par ShinyIDEA

Figure 7 : Présentation de ShinyIDEA

3.1 - Développement de l'interface web

Différents choix techniques ont été pris autant sur l'aspect "front-end" (ce que l'utilisateur voit) que sur l'aspect "back-end" (le moteur derrière l'application). Le framework [golem](#) a notamment été privilégié pour développer l'application sur le même format qu'un package R.

Dans un premier temps, les fonctionnalités de l'application ne sont débloquées que si l'utilisateur accepte les Conditions Générales d'Utilisation (Figure 8). Ces conditions n'incluent à ce jour aucune contrainte particulière et servent principalement de mode d'emploi de l'application. Il est cependant envisagé que ces conditions évoluent et demandent par exemple à l'utilisateur son autorisation pour capter ses données insérées de manière anonyme.



Conditions Générales d'Utilisation

En cochant cette case, je confirme avoir pris connaissance des conditions générales d'utilisation de cet outil et je les accepte.

 Commencer une analyse individuelle

 Commencer une analyse de groupe

Figure 8 : Interface de validation des CGU

Deux modules sont proposés dans cette application :

- Un module d'analyse individuelle : Demande un calculateur excel en entrée, et permet d'obtenir un certain nombre de résultats/graphiques ainsi qu'un rapport synthétique pour une seule exploitation. Les modules IDEATools mobilisés sont `importIDEA()`, `dimensionsPlots()`, `MakeTrees()`, `RadarPlots()` et `exportIDEA()`.
- Un module d'analyse de groupe : Dans ce module, l'utilisateur peut insérer plusieurs (> 3) calculateurs simultanément afin d'obtenir des résultats de synthèse permettant de comparer ces exploitations entre elles. Différentes sorties graphiques sont ici encore proposées. Les modules IDEATools mobilisés sont `importIDEA()`, `dimensionsPlots()`, `MakeTrees()`, `RadarPlots()`, `metaIDEA()` et `exportIDEA()`.

Ces deux modules sont indépendants et ne communiquent pas entre eux. Ainsi, l'insertion d'un calculateur dans un des modules n'interagit pas avec les calculateurs de l'autre module.

3.1.1 - L'analyse individuelle

Dans ce premier module, l'utilisateur peut insérer un (et un seul) calculateur IDEA correctement rempli. Après un bref temps de chargement, un message va alors apparaître (Figure 9) :

- Si la version du calculateur est supérieure à la 4.2.0, alors toutes les métadonnées doivent être accessibles et le message indiquera "succès"
- Si la version est inférieure à la 4.2.0, alors il est possible que des métadonnées ne soient pas retrouvées par l'algorithme, ce qui sera signalé dans le message par un "Attention !".



Figure 9 : Les deux messages pouvant être renvoyés par l'application, selon l'ancienneté du calculateur importé

Une fois ce message fermé, l'utilisateur pourra alors consulter une rapide synthèse des résultats de son exploitation : Sa note globale IDEA, sa note pour les 3 dimensions et son évaluation globale pour chacune des 5 propriétés (Figure 10).

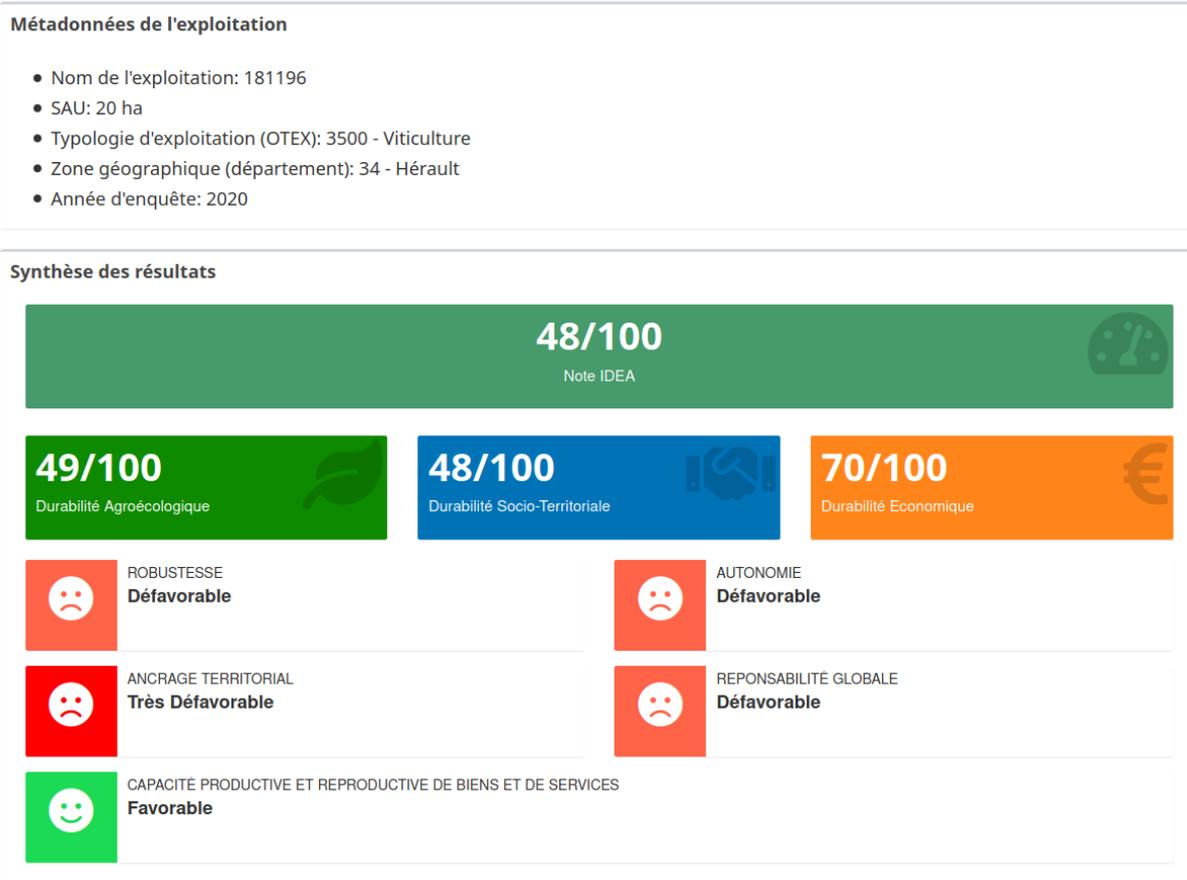


Figure 10 : Exemple d'une synthèse produite par l'application

Pour obtenir plus de détails sur son évaluation, l'utilisateur peut ensuite naviguer vers l'onglet "Détail des indicateurs" pour observer ses notes de composantes et d'indicateurs selon différentes visualisations personnalisables par l'utilisateur (diagrammes radar ou en barres horizontales, résultats exprimés en valeur brute ou en %, indicateurs regroupés par dimensions ou par propriétés, etc.) et vers l'onglet "Détail des arbres éclairés" pour observer le détail de son évaluation agrégative (Figures 11 et 12).

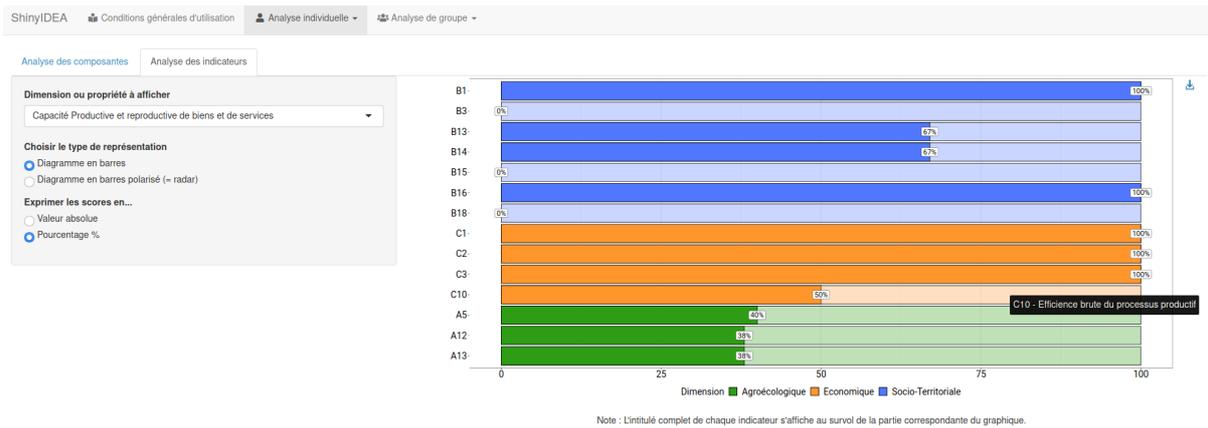


Figure 11 : Exemples de visions détaillées pour les indicateurs appartenant à la propriété "Capacité productive et reproductive de biens et de services"

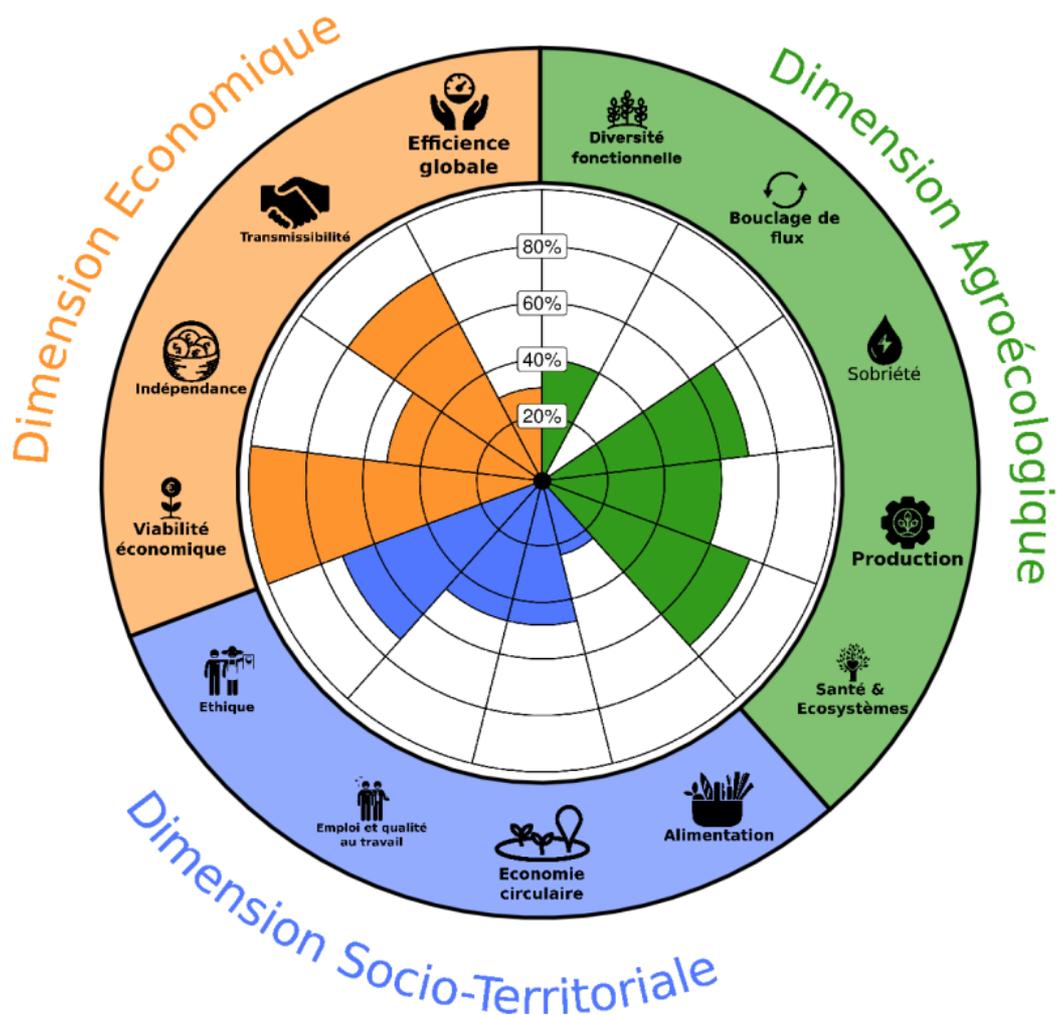
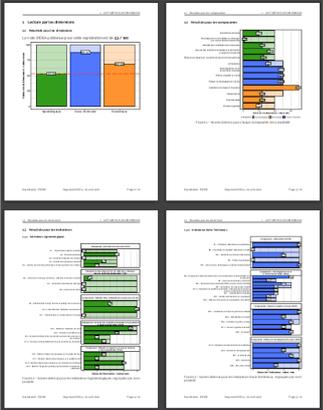


Figure 12 : Exemples de vision polarisée (radar) proposée pour les composantes

Enfin, l'utilisateur peut utiliser deux modules d'exports, lui permettant de récupérer soit un rapport PDF compilant l'ensemble des résultats et figures, soit une archive ZIP contenant l'ensemble des figures (Figure 13)

Rapport automatisé



[Télécharger le rapport](#)

Description : Ce rapport est un document PDF généré automatiquement, comptant 16 pages, compilant l'ensemble des métadonnées extraites dans le calculateur ainsi que l'ensemble des graphiques relatifs à l'approche par les dimensions ainsi que par les propriétés (dont les arbres éclairés). NB : La production du fichier peut nécessiter quelques secondes.

Archive compressée

Dimensions	5 éléments
628934_Composantes.png	436,6 kB
628934_Dimensions.png	105,0 kB
628934_Indicateurs Agroécologiques.png	613,4 kB
628934_Indicateurs Economiques.png	351,6 kB
628934_Indicateurs Socio-Territoriaux.png	601,1 kB
Propriétés	6 éléments
Arbres éclairés	14 éléments
628934_Ancrage Territorial.pdf	9,5 kB
628934_Ancrage Territorial.png	76,5 kB
628934_Autonomie.pdf	8,9 kB
628934_Autonomie.png	75,2 kB
628934_Capacité productive et reproductive de biens et de services.pdf	9,9 kB
628934_Capacité productive et reproductive de biens et de services.png	87,4 kB
628934_Global.pdf	20,7 kB
628934_Global.png	224,1 kB
628934_Global_zoom.pdf	9,1 kB
628934_Global_zoom.png	67,3 kB
628934_Responsabilité globale.pdf	11,2 kB
628934_Responsabilité globale.png	90,8 kB
628934_Robustesse.pdf	10,0 kB
628934_Robustesse.png	103,4 kB
628934_Radar_Ancrage Territorial.png	523,1 kB
628934_Radar_Autonomie.png	438,6 kB
628934_Radar_Capacité productive et reproductive de biens et de services.png	345,4 kB
628934_Radar_Responsabilité globale.png	846,4 kB
628934_Radar_Robustesse.png	594,6 kB

[Télécharger l'archive](#)

Description : Cette archive est un dossier compressé ZIP contenant l'ensemble des graphiques relatifs à l'approche par les dimensions ainsi que par les propriétés (dont les arbres éclairés). NB : La production du fichier peut nécessiter quelques secondes.

Figure 13 : Téléchargements des résultats d'analyse individuelle proposés par l'application

3.1.2 - L'analyse de groupe

Dans ce module, l'utilisateur peut insérer plusieurs (au moins 3) calculateurs dans via le bouton de chargement (Note : Maintenir CTRL enfoncé pour une sélection multiple). Il obtient alors une première synthèse très succincte du groupe inséré : Nombre d'exploitations et scores moyens pour les dimensions (Figure 14).

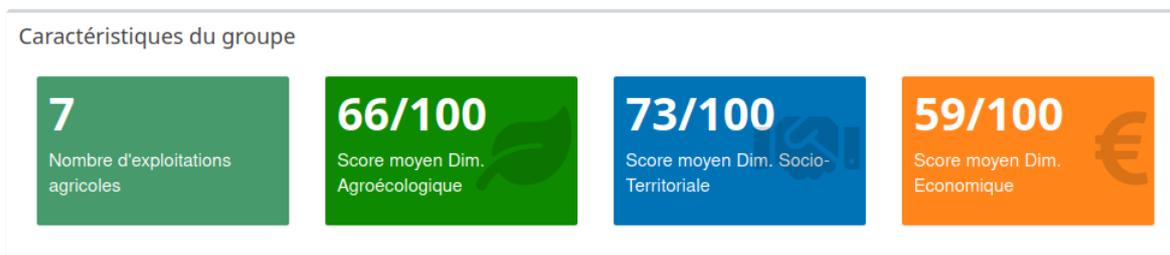
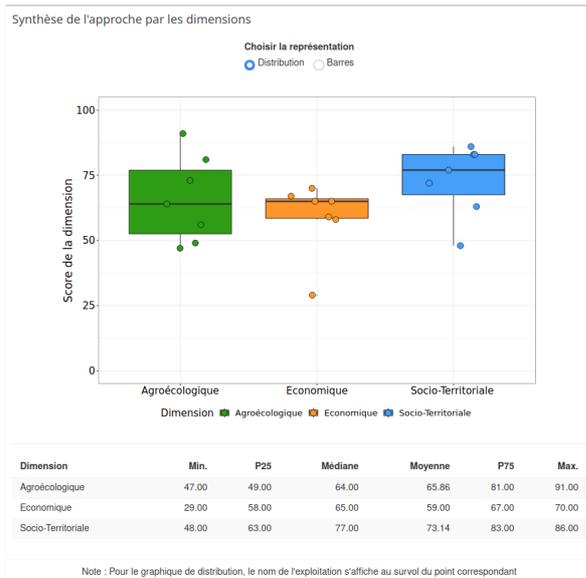


Figure 14 : Exemple d'une synthèse proposée par une analyse de groupe sur 7 exploitations

L'utilisateur peut ensuite naviguer vers un menu proposant le détail des indicateurs de ce groupe. Différentes représentations sont proposées (figures, tableaux) à la fois pour l'approche par les dimensions (à gauche sur la Figure 15) et par les propriétés (à droite sur la Figure 15).



Synthèse de l'approche par les propriétés

Table de synthèse [Version graphique](#)

Show entries Search:

	Ancrage territorial	Autonomie	Capacité productive et reproductive de biens et de services	Responsabilité globale	Robustesse
140196	favorable	favorable	défavorable	favorable	défavorable
181196	très défavorable	défavorable	favorable	défavorable	défavorable
212983	favorable	défavorable	défavorable	favorable	défavorable
333333	favorable	favorable	défavorable	favorable	très défavorable
579272	défavorable	défavorable	défavorable	défavorable	favorable
628934	favorable	défavorable	très défavorable	très défavorable	défavorable
654321	favorable	défavorable	défavorable	très défavorable	défavorable

Showing 1 to 7 of 7 entries Previous Next

Figure 15 : Exemple de résultats détaillés pour une analyse de groupe sur 7 exploitations

Enfin, l'utilisateur peut ici encore utiliser deux modules d'exports (Figure 16), lui permettant de récupérer soit :

- Un fichier excel de synthèse compilant des statistiques descriptives pour les dimensions (quartiles, moyenne, etc.) ainsi que des cases colorées pour l'ensemble des 46 noeuds et 53 indicateurs de l'arbre éclairé des propriétés (avec dénombrement de chaque modalité d'évaluation pour chaque noeud).
- Une archive zip contenant l'ensemble des figures pour chaque exploitation individuelle (un sous-dossier par exploitation) + des figures de synthèses issues du module metaIDEA () .

Fichier excel

	A	B	C	D	E	F
1	Exploitation	Ancrage territorial	Autonomie	Capacité productive et reproductive de biens et de services	Responsabilité globale	Robustesse
2	111111	très favorable	favorable	très favorable	favorable	défavorable
3	123456	défavorable	défavorable	favorable	favorable	favorable
4	140196	très défavorable	favorable	défavorable	très défavorable	défavorable
5	181196	très défavorable	défavorable	favorable	défavorable	défavorable
6	212983	très favorable	favorable	défavorable	favorable	défavorable
7	29 1	défavorable	défavorable	favorable	favorable	favorable
8	29 2	très défavorable	très défavorable	très défavorable	favorable	défavorable
9	29 3	favorable	défavorable	défavorable	défavorable	très favorable
10	29 4	défavorable	favorable	très favorable	très favorable	défavorable
11	56 1	favorable	favorable	favorable	très favorable	défavorable
12	56 2	défavorable	très défavorable	favorable	très favorable	défavorable
13	56 3	très favorable	favorable	très favorable	favorable	favorable
14	579272	défavorable	défavorable	défavorable	très favorable	favorable
15	628934	favorable	défavorable	très favorable	très favorable	défavorable
16	654321	favorable	défavorable	défavorable	très défavorable	défavorable
17	646789	très favorable	favorable	défavorable	défavorable	favorable
18	56 1	favorable	favorable	favorable	très favorable	défavorable
19	56 2	défavorable	très défavorable	favorable	très favorable	défavorable
20	56 3	très favorable	favorable	très favorable	favorable	favorable
21	579272	défavorable	défavorable	défavorable	très favorable	favorable
22	628934	favorable	défavorable	très favorable	très favorable	défavorable
23	654321	favorable	défavorable	défavorable	très défavorable	défavorable
24	646789	très favorable	favorable	défavorable	défavorable	favorable
25	56 1	favorable	favorable	favorable	très favorable	défavorable
26	56 2	défavorable	très défavorable	favorable	très favorable	défavorable
27	56 3	très favorable	favorable	très favorable	favorable	favorable
28	579272	défavorable	défavorable	défavorable	très favorable	favorable
29	628934	favorable	défavorable	très favorable	très favorable	défavorable
30	654321	favorable	défavorable	défavorable	très défavorable	défavorable
31	646789	très favorable	favorable	défavorable	défavorable	favorable
32	654321	défavorable	défavorable	défavorable	défavorable	favorable

Archive compressée

- 732420 2 éléments
 - Dimensions 5 éléments
 - 732420_Composantes.png 440,7 kB
 - 732420_Dimensions.png 102,4 kB
 - 732420_Indicateurs Agroécologiques.png 608,5 kB
 - 732420_Indicateurs Economiques.png 350,4 kB
 - 732420_Indicateurs Socio-Territoriaux.png 606,8 kB
 - Propriétés 1 élément
 - Arbres éclairés 14 éléments
 - 732420_Ancrage Territorial.pdf 9,5 kB
 - 732420_Ancrage Territorial.png 75,0 kB
 - 732420_Autonomie.pdf 8,9 kB
 - 732420_Autonomie.png 71,8 kB
 - 732420_Capacité productive et reproductive de biens et de services.pdf 9,9 kB
 - 732420_Capacité productive et reproductive de biens et de services.png 85,4 kB
 - 732420_Global.pdf 20,7 kB
 - 732420_Global.png 221,6 kB
 - 732420_Global_zoom.pdf 9,1 kB
 - 732420_Global_zoom.png 69,9 kB
 - 732420_Responsabilité globale.pdf 11,2 kB
 - 732420_Responsabilité globale.png 90,1 kB
 - 732420_Robustesse.pdf 10,0 kB
 - 732420_Robustesse.png 102,7 kB
- 881700 2 éléments
 - Dimensions 5 éléments
 - Propriétés 1 élément

Figure 16 : Téléchargements des résultats d'analyse de groupe proposés par l'application

3.2 - Hébergement serveur

La solution sélectionnée pour l'hébergement actuel est un serveur virtuel (Debian 9) hébergé au CINES à Montpellier.

L'URL actuelle de l'application est : <http://10.134.193.81:3838/ShinyIDEA/>

Mais, comme le format de cette adresse ip l'indique (10.xxx.xxx.xx), l'application n'est actuellement accessible que via le réseau interne INRAE (cable réseau ou VPN). Un tunnel sécurisé (https) entre cette adresse privée et une adresse publique est en cours de construction.

3.3 - Apport à la plateforme WEB-IDEA (BSA)

Une partie du code source de ShinyIDEA a été extraite afin de produire un module spécifiquement dédié à la production d'arbres éclairés et qui a été intégré au site développé par Bordeaux Sciences Agro : <http://idea4.agro-bordeaux.fr> En fonction de l'utilisateur connecté, l'appli restreint la sélection de dossiers aux dossiers appartenant à l'organisme de l'utilisateur. Une fois les filtres choisis et la recherche exécutée, une connexion lecture-seule à la base de données derrière le site web est effectuée, les données sont manipulées et transformées pour devenir identiques à un fichier json, puis les modules d'IDEATools permettent de produire les arbres éclairés. Ceux-ci sont présentés directement dans l'interface puis téléchargeables via une archive ZIP (Figure 17).

The screenshot shows the 'Production d'arbres éclairés' module in the WEB-IDEA interface. The interface is divided into several sections:

- Navigation:** 'Administration', 'Mon organisme', 'Global', 'Téléchargement', and 'Bordeaux Sciences Agro (4 / 22 dossiers) AD MIN'.
- Search and Filter:** 'VUE PAR LES PROPRIÉTÉS', 'Choisir le numéro de dossier' (111112), 'Choisir l'année' (2019), 'RECHERCHER / ACTUALISER', and 'TÉLÉCHARGER LES ARBRES ÉCLAIRÉS'.
- Diagram:** A central diagram titled 'Exploitation agricole' with sub-diagrams for 'RESPONSABILITÉ GLOBALE', 'ROBUSTESSE', 'CAPACITÉ PRODUCTIVE ET REPRODUCTIVE DE BIENS ET SERVICES', 'ANCRAGE TERRITORIAL', and 'AUTONOMIE'. Each sub-diagram contains specific metrics and their status (e.g., 'Très favorable', 'Favorable', 'Intermédiaire', 'Défavorable', 'Très défavorable', 'Non concerné').
- Legend:** A color-coded legend at the bottom of the diagram.

Figure 17 : Module de production d'arbres éclairés implémenté sur le site WEB-IDEA

4 - Stratégie de captation des données

4.1 - Objectifs

ShinyIDEA est une application web qui met à disposition du grand public les algorithmes automatisés développés dans le package R « IDEATools ».

Lorsqu'un utilisateur insère un calculateur dans cette application, les données sont temporairement stockées sur le serveur le temps des calculs. Ainsi, il serait tout à fait envisageable de capter ces données de manière anonyme afin de construire une large base de donnée pouvant être ré-utilisée dans de futurs travaux de recherche.

Dans cette optique, Une base PostgreSQL dite « opérationnelle », collectant les données relatives aux indicateurs IDEA4, a été créée. A ce jour, l'architecture de cette base n'est structurée que pour le stockage de données relatives à IDEA, sur le modèle de la base WEB-IDEA (107 items et 16 métadonnées).

Cependant, les calculateurs IDEA contiennent un très grand nombre de données brutes qui ne sont pas exploitées, éparpillées dans l'onglet "Saisie et calculateur". Récupérer ces données nécessite encore le développement de nombreux algorithmes complexes tant la disposition des cellules a pu être modifiée au fil des versions du calculateur. Ces travaux sont encore en cours de développement.

4.2 - La base de données PostgreSQL « IDEA_V4 »

La base de données « IDEA_V4 » est actuellement hébergée sur un serveur PostgreSQL d'un datacenter INRAE et a pour objectif de collecter des données relatives aux indicateurs IDEA dans un format structuré et facilement accessible. Celle-ci est actuellement composée de 5 tables (Figure 18).

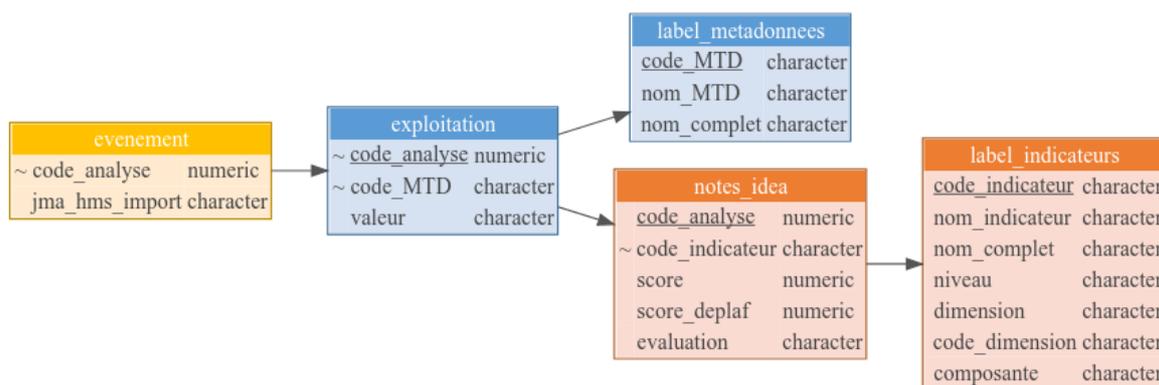


Figure 18 : Schéma des 5 tables de la base de données "IDEA_V4"

4.2.1 - La table "evenement"

Cette table est la première à être alimentée lors d'un import. Elle consigne l'identifiant unique d'import, l'heure précise à laquelle le calculateur a été importé et la version du calculateur. L'identifiant unique d'import est un entier numérique automatiquement incrémenté au fil des imports.

Table 1 : Extrait de la table "evenement"

code_analyse	jma_hms_import	version
10	03-09-2020_09-52-42	4.2.7
11	03-09-2020_09-52-43	4.2.8
12	03-09-2020_09-52-45	4.2.2
42	03-09-2020_09-53-27	4.2.2
81	03-09-2020_09-54-20	4.2.8

4.2.2 - La table "exploitation"

Cette table est automatiquement alimentée au fil des imports de données et répertorie, pour chaque import, les données propres à l'exploitation. L'intitulé correspondant aux codes métadonnées (MTD_**) peut être retrouvé par jointure avec la table "label_metadonnees".

Table 2 : Extrait de la table "exploitation"

code_analyse	code_MTD	valeur
1	MTD_01	100001
1	MTD_02	77
1	MTD_03	2

4.2.3 - La table "label_metadonnees"

Cette table est fixe et n'est pas alimentée au fil des imports. Elle contient des informations relatives aux métadonnées.

Table 3 : Extrait de la table "label_metadonnees"

code_MTD	nom_MTD	nom_complet
MTD_01	Identifiant exploitation	MTD_01 - Identifiant exploitation
MTD_02	SAU	MTD_02 - SAU
MTD_03	UTH	MTD_03 - UTH
MTD_04	UTH F	MTD_04 - UTH F
MTD_05	Tranche d'âge du chef d'exploitation	MTD_05 - Tranche d'âge du chef d'exploitation

4.2.4 - La table "notes_idea"

Cette table est automatiquement alimentée au fil des imports de données et répertorie, pour chaque import, les valeurs attribuées aux indicateurs. L'intitulé

correspondant aux codes indicateurs peut être retrouvé par jointure avec la table "label_indicateurs".

Table 4 : Extrait de la table "notes_idea"

code_analyse	code_indicateur	score	score_deplaf	evaluation
1	A1	4	4	favorable
1	A2	2	2	intermédiaire
1	A3	4	4	intermédiaire
1	A4	3	3	intermédiaire
1	A5	5	5	favorable
1	A6	2	2	défavorable

4.2.5 - La table "label_indicateurs"

Cette table est fixe et n'est pas alimentée au fil des imports : Elle contient des données relatives aux indicateurs et permet de retrouver, à partir des « code_indicateurs », différentes informations (intitulé de l'indicateur, dimension d'appartenance, niveau hiérarchique, etc.) :

Table 5 : Extrait de la table "label_indicateurs"

code_indicateur	nom_indicateur	nom_complet	niveau	dimension	code_dimension	composante
B16	Intensité et qualité au travail	B16 - Intensité et qualité au travail	indicateur	Socio-Territoriale	B	Emploi et qualité au travail
B18	Formation	B18 - Formation	indicateur	Socio-Territoriale	B	Emploi et qualité au travail
B13	Réseaux d'innovation et mutualisation du matériel	B13 - Réseaux d'innovation et mutualisation du matériel	indicateur	Socio-Territoriale	B	Développement local et économie circulaire
B15	Mutualisation du travail	B15 - Mutualisation du travail	indicateur	Socio-Territoriale	B	Emploi et qualité au travail
R8	par l'insertion dans les réseaux	R8 - par l'insertion dans les réseaux	noeud	NA	NA	NA
R9	Augmenter la capacité d'adaptation	R9 - Augmenter la capacité d'adaptation	branche	NA	NA	NA

5 - Références

Carayon, D., Girard, S., Zahm, F. (2020). IDEATools: Un applicatif pour le calcul, l'automatisation et l'exploitation de données IDEA4. R package version 1.0. <https://github.com/davidcarayon/IDEATools>.

Zahm F., Alonso Ugaglia A., Boureau H., Del'homme B., Barbier J.M., Gasselin P., Gafsi M., Girard S., Guichard L., Loyce C., Manneville V., Menet A., Redlingshofer B., 2019, Évaluer la durabilité des exploitations agricoles. La méthode IDEA v4, un cadre conceptuel mobilisant dimensions et propriétés de la durabilité, Cahiers Agricultures, 28, 5, <https://doi.org/10.1051/cagri/2019004>.

Zahm et al. (2019). "Évaluer la durabilité des exploitations agricoles. La méthode IDEA v4, un cadre conceptuel combinant dimensions et propriétés de la durabilité" in Cahiers Agricultures, 8(5):1-10.



Centre Nouvelle Aquitaine - Bordeaux
50 avenue de Verdun, Gazinet
F-33612 CESTAS Cedex

Rejoignez-nous sur :



**Institut national de recherche pour
l'agriculture, l'alimentation et l'environnement**

INRAE
la science pour la vie, l'humain, la terre