

# Projet COMECO

*Brun, P.,  
Descombes P.,  
Zurell D.,  
Karger D. N.,  
Wegner J. D.,  
Zimmermann N. E.*

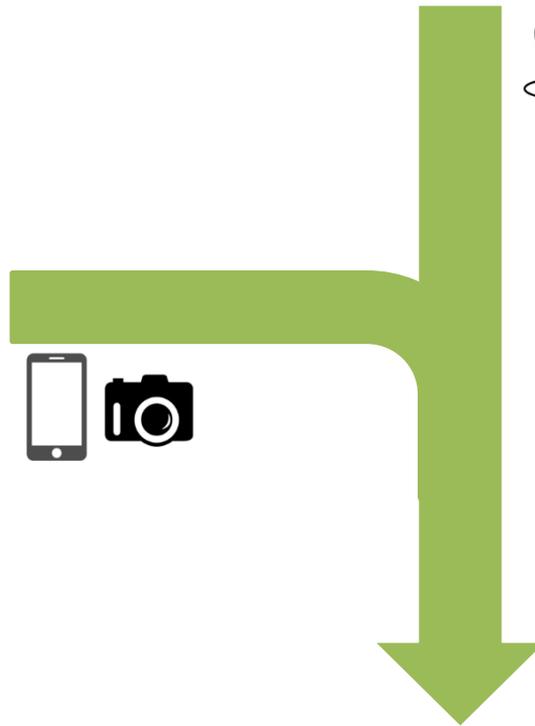
Séminaire validation  
Flora Jurana  
25.11.2020

   
**info flora**

   
**ETH** **EcoVision**

Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich

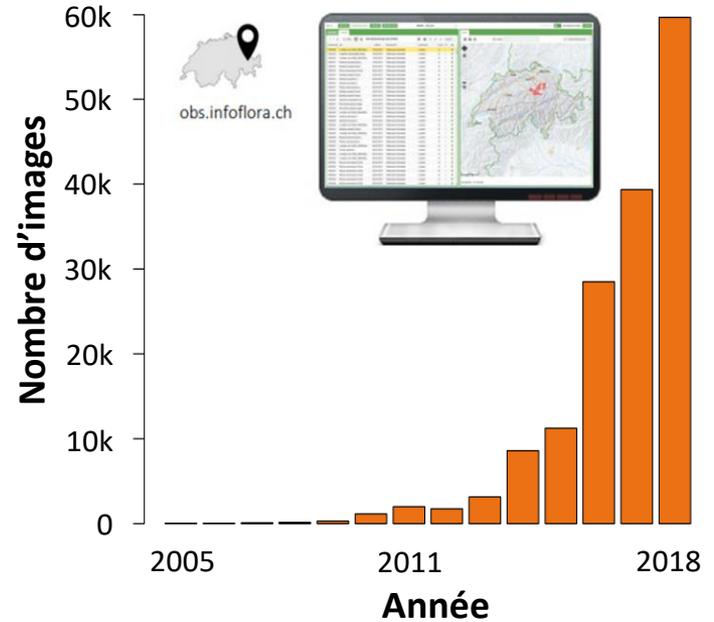
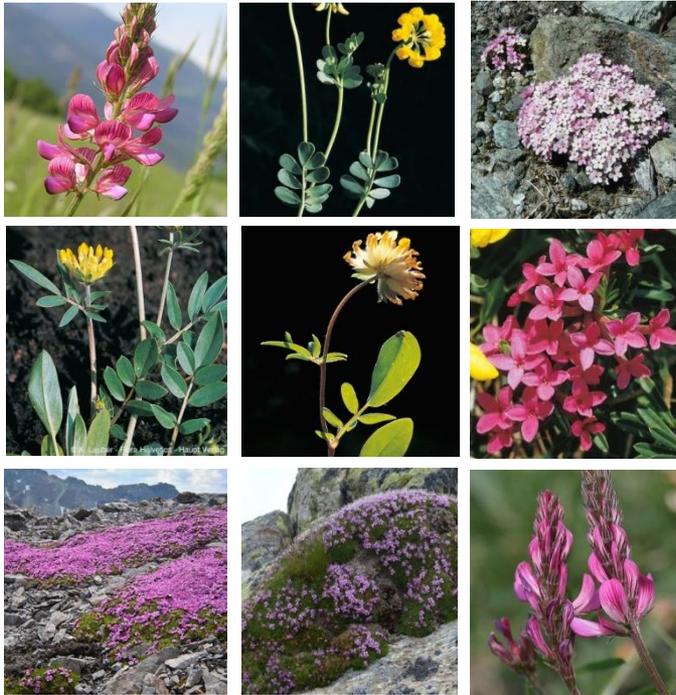
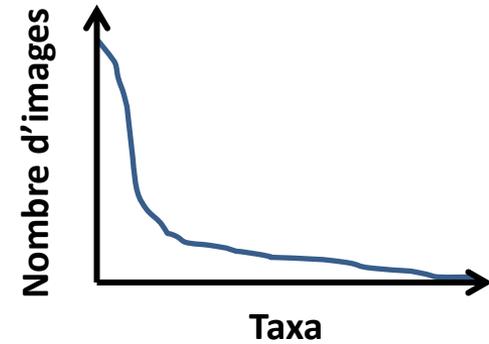
  
**SDSC**



# Base de données d'images InfoFlora

- 350'000 images géoréférencées (obs ID + coord XY)
- 4157 taxa (1-5295 images/taxa)

*Etat au 09.07.2020*



## (1) Recherche et développement (R&D)

- Développer de nouvelles approches afin d'améliorer la reconnaissance d'espèce sur image



- Comprendre les sources de confusions (quelles espèces sont confondues et pourquoi)
  - Morphologie / phylogénie / écologie / taxonomie

## (2) Application pratique (FloreApp InfoFlora)

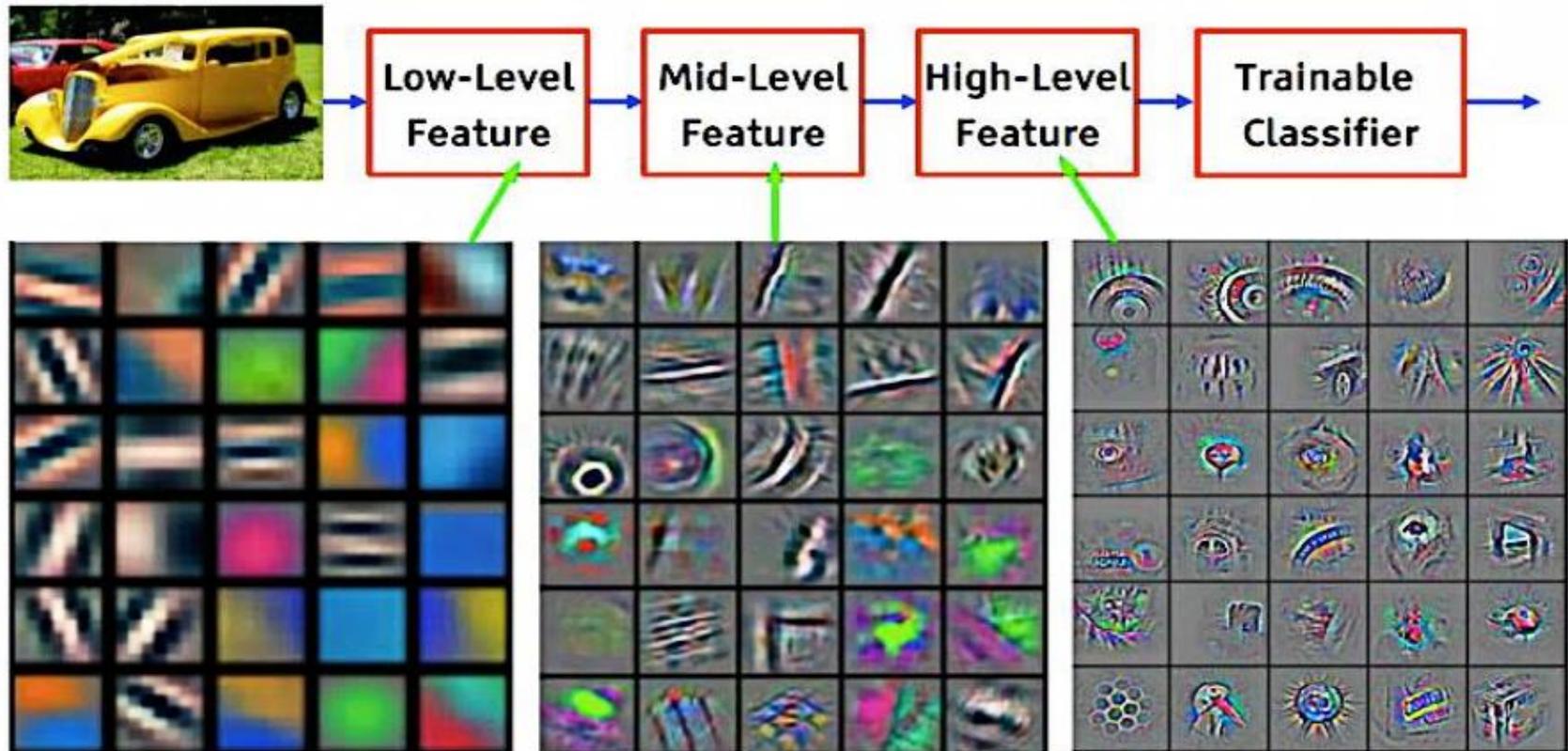


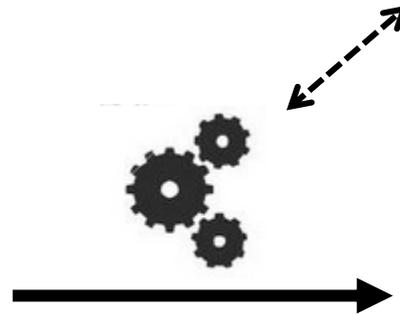
- Validation des données
- Aide aux botanistes sur le terrain (détermination, feedback)

**Améliorer/favoriser la qualité et l'acquisition de données**

# Classification des images par machine learning

- Algorithmes développés dans les années 1980
- Puissance informatique + Big data ont favorisé la technique
- Images décomposées et formes «comparées» à une BD référence
- BD référence : > 100 images /objets = bon





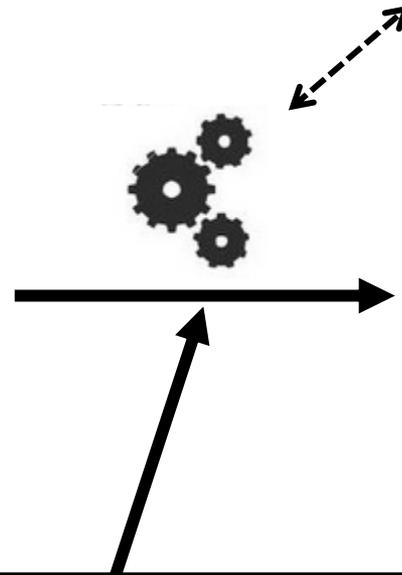
*BD réf.*



*Androsace vandellii (Turra) Chiov.*  
*Androsace helvetica (L.) All.*  
*Androsace alpina (L.) Lam.*

Approche simple

BD réf.

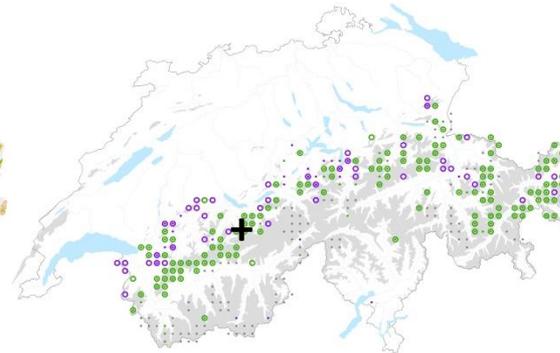


*Androsace vandellii* (Turra) Chiov.  
*Androsace helvetica* (L.) All.  
*Androsace alpina* (L.) Lam.

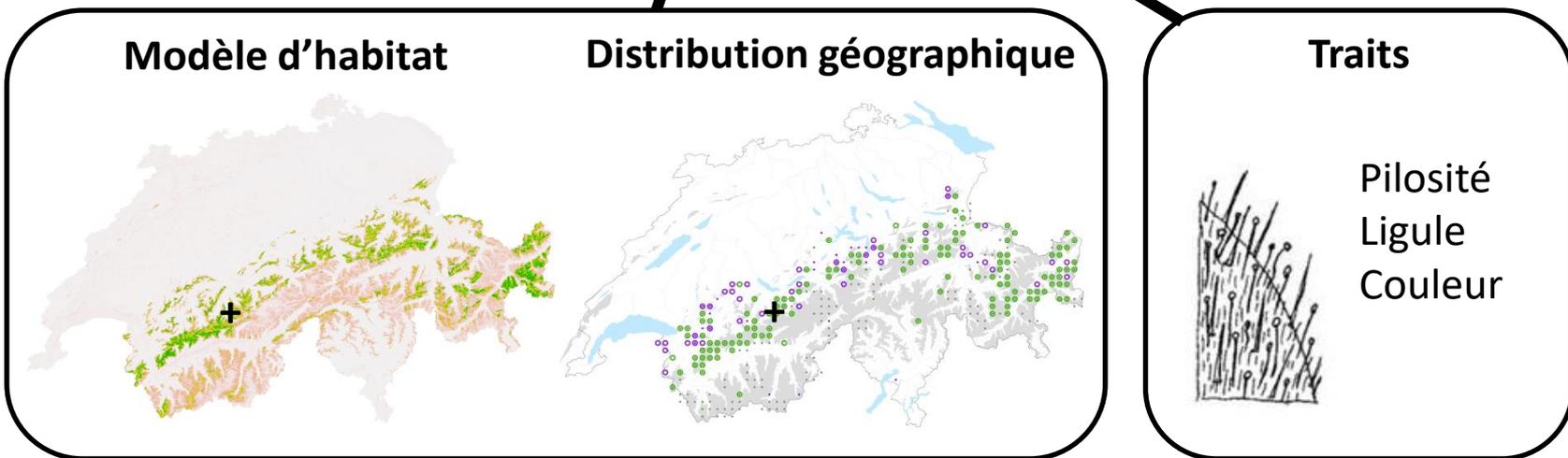
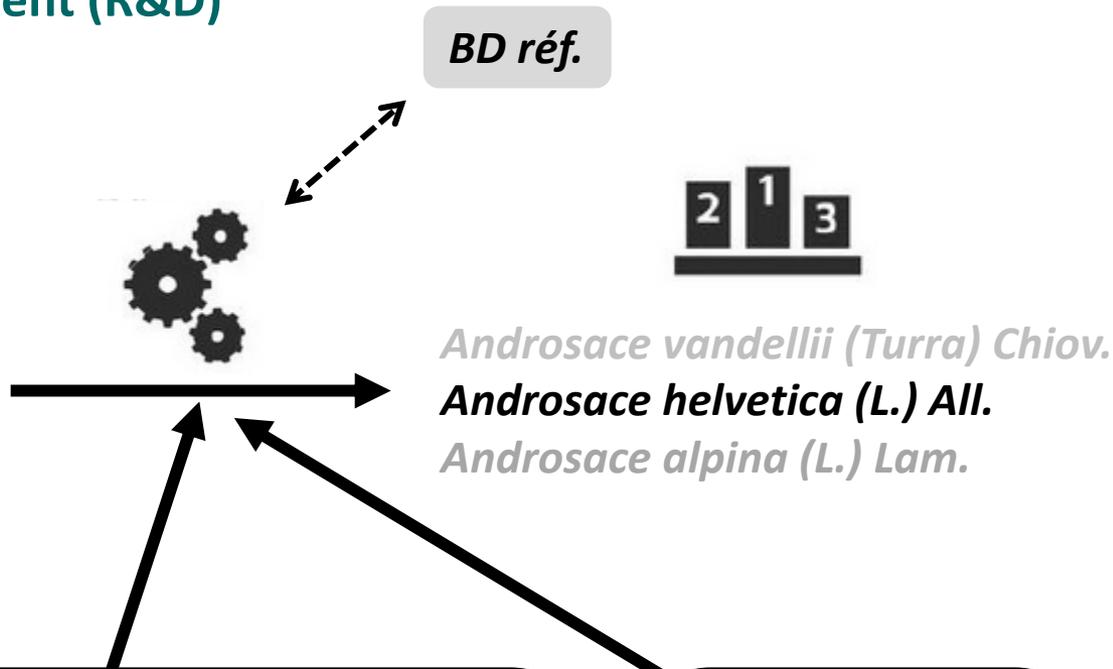
Modèle d'habitat



Distribution géographique



Améliorer la reconnaissance des espèces avec des informations complémentaires



Améliorer la reconnaissance des espèces avec des informations complémentaires

## Comparaison entre InfoFlora et iNaturalist

Dataset	InfoFlora	iNaturalist
Number of images	358,952	58,956
Number of images with valid labels	19,837	58,956
Number of valid classes	2,116	2,347
Number of selected classes (num of img $\geq$ 10)	498	959
Number of selected images for training	11,355	50,002
Number of selected images for validation	2,490	4,795

@ Yihang She, Riccardo De Luttio et Wegner D.

## Comparaison entre InfoFlora et iNaturalist

Dataset	InfoFlora	iNaturalist
Number of images	358,952	58,956
Number of images with valid labels	19,837	58,956
Number of valid classes	2,116	2,347
Number of selected classes (num of img >= 10)	498	959
Number of selected images for training	11,355	50,002
Number of selected images for validation	2,490	4,795

@ Yihang She, Riccardo De Luttio et Wegner D.

## Comparaison simple (avec ResNet34) sans distribution écologique/géographique

BD références

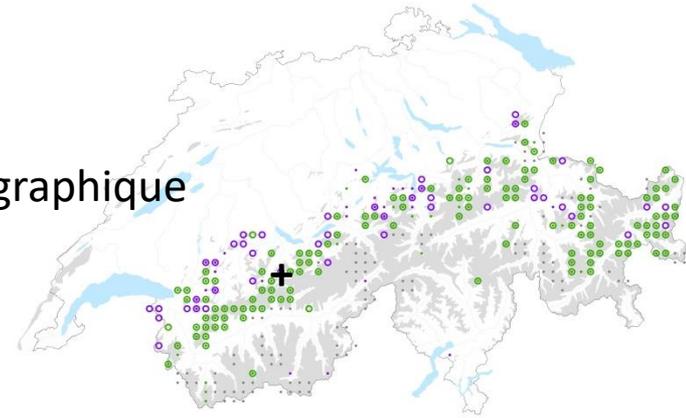
	InfoFlora	iNaturalist
Training accuracy	76.45 %	97.91 %
Validation accuracy	<b>32.37 %</b>	<b>57.62 %</b>

**La qualité de la reconnaissance dépend de la qualité et la quantité d'images validées dans la BD de référence**

Validation par expert ou automatisée (filtrage)

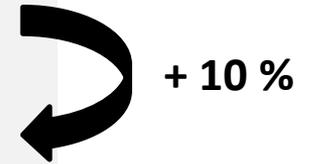
# Résultats préliminaires

Modèle simple (avec ResNet50) + distribution géographique



BD références

	InfoFlora	iNaturalist
Training accuracy	76.45 %	97.91 %
Validation accuracy	<b>32.37 %</b>	<b>57.62 %</b>
Validation accuracy (avec distribution géographique)	-	<b>68.09 %</b>



**Fort potentiel des informations supplémentaires pour améliorer la précision de la reconnaissance**

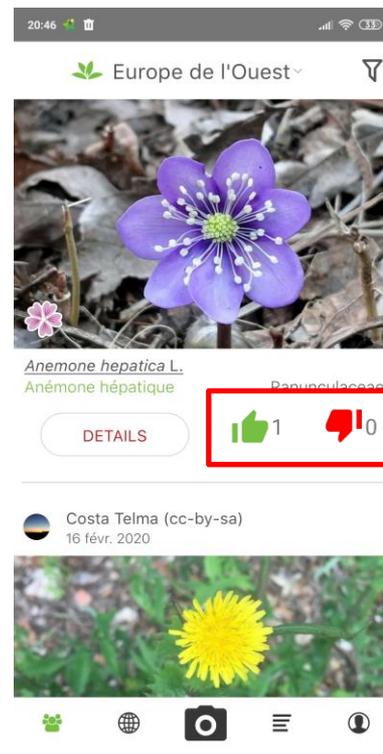
Prochaines étapes:

- Inclure les modèles de niche écologique / traits morphologiques

# Applications pratiques potentielles

Validation des images par les utilisateurs via florapp ou le carnet en ligne

- Identification correct ?
- Qualité et pertinence de l'image ?



**Proposition potentielle permettant d'augmenter la qualité de la base de donnée de référence sur le long terme**

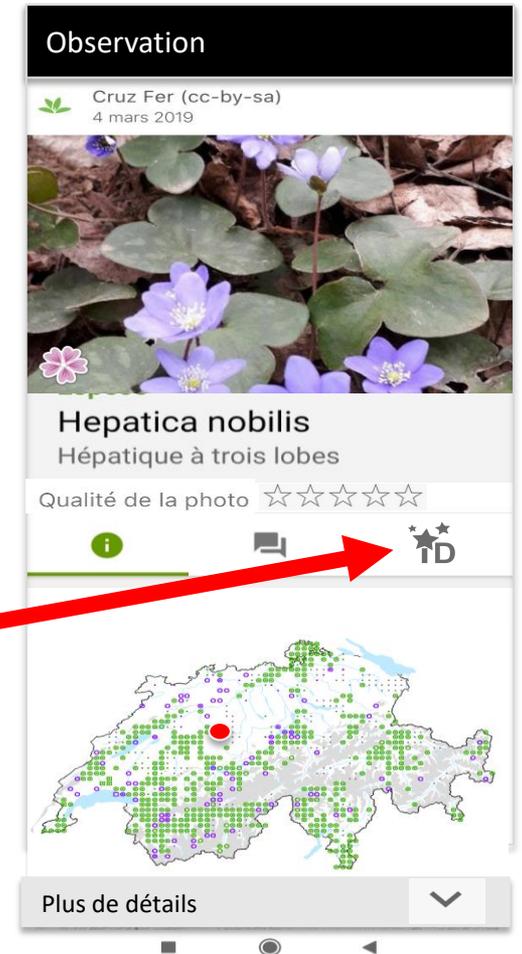
## Applications pratiques potentielles

Validation des observations en analysant les images en interne dans la BD

- Détection d'observations douteuses

Fournir une aide à la détermination sur le terrain

- Famille / Genre / Espèce
- Gain de temps



**Proposition potentielle permettant d'augmenter la qualité des observations transmises à InfoFlora**

# Remerciements

# Projet COMECO

*Brun, P.,  
Descombes P.,  
Zurell D.,  
Karger D. N.,  
Wegner J. D.,  
Zimmermann N. E.*

Séminaire validation  
Flora Jurana  
25.11.2020



Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich



*Descombes Patrice,*  
[patrice.descombes@wsl.ch](mailto:patrice.descombes@wsl.ch)

