

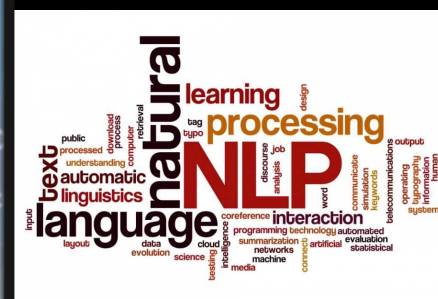
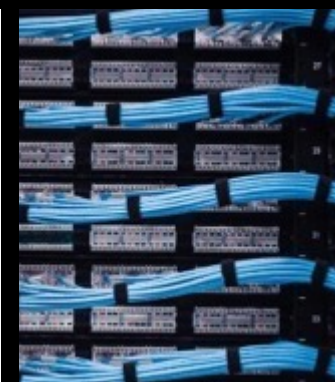
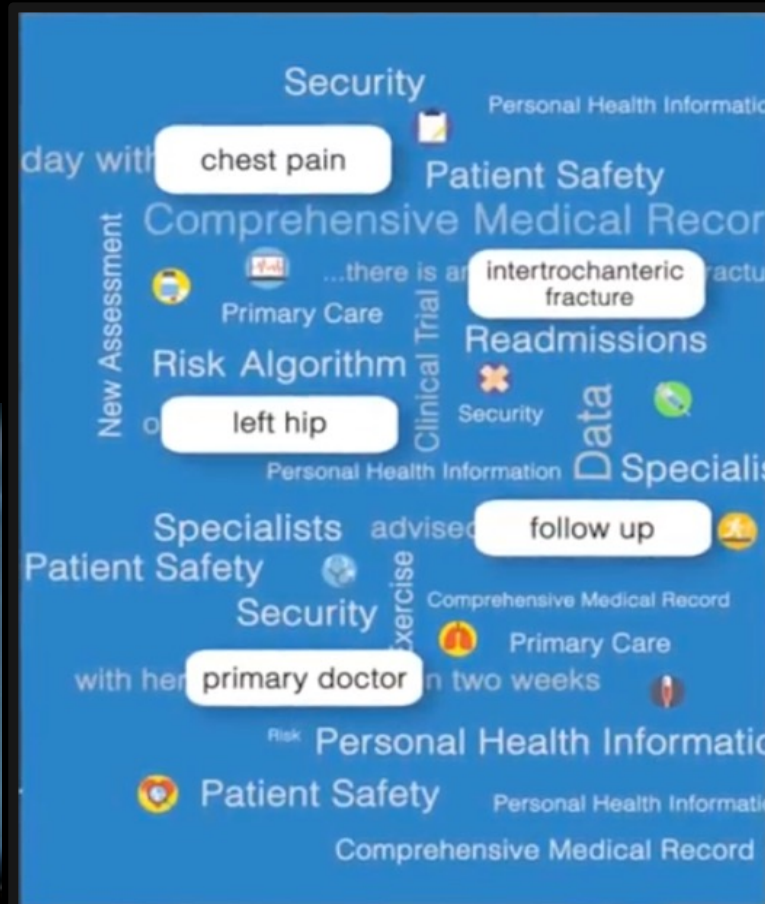
NLP

Natural Language Processing



Sylvain Guilbaud

Sales Engineer, InterSystems France



iKnow sur Github <https://github.com/intersystems/iknow>

Installation du module iknowpy

- <https://pypi.org/project/iknowpy/>

```
$ pip install iknowpy
Collecting iknowpy
  Downloading iknowpy-1.1.0-cp36.cp37.cp38.cp39-cp36m.cp37m.cp38.cp39-macosx_10_9_x86_64.whl (28.3 MB)
    |████████████████████| 28.3 MB 9.1 MB/s
Installing collected packages: iknowpy
Successfully installed iknowpy-1.1.0
```

- <https://community.intersystems.com/post/iknow-our-open-source-python-library-hits-v10>
- https://www.youtube.com/watch?v=VkJ6rcv4_aA

Using the iKnow Engine in Python



demo notebook (unsaved changes)

View Insert Cell Kernel Widgets Help Trusted

Run Code

```
3
4 #from colorama import init
5 #init() # init colorama - only when running outside notebook
6
7 iknow.index(text, language="en", iknow=iknowpy.iKnowEngine()):
8
9 for s in iknow.m_index['sentences']:
10
11     # first figure out where negation spans are and tag
12     # those entities
13     for a in s['path_attributes']:
14
15         # path attributes are expressed as positions within s
16         # which in turn keys into the s['entities'] array
17         for ent in range(s['path'][a['pos']],
18                         s['path'][a['pos']+a['span']-1]+1):
19             if a['type']=="Negation":
20                 s['entities'][ent]['colour'] = Fore.RED
21             if a['type']=="Certainty":
22                 s['entities'][ent]['colour'] = Fore.GREEN
```

InterSystems IRIS Text Analytics

Détection de Concepts Multi-Langages

Contrairement à la plupart des autres technologies sémantiques, IRIS Text Analytics indique automatiquement les **concepts** les plus **significatifs**, avec **gestion de la négation**.

Suite à ce **bilan**, une **prise en charge** en **urgence** par l'**équipe de chirurgie cardio-vasculaire** est organisée, qui procède à une **embolectomie** du **membre inférieur gauche** avec un **bon résultat**.

Pas d'anomalie de la **dénomination** **ni** de la **répétition**.

Nerfs crâniens : **Pas** de **paralysie faciale**.

Acuité visuelle très diminuée mais **pas** de **trouble évident** de l'**oculomotricité** **ni** du **champ visuel**.

Voies longues : **Pas** de **déficit sensitif** **ni** **moteur**.

ROT discrets mais **symétriques**.

Pas de **Babinski**.

Pas de **clonus de cheville**.

Pas d'**ataxie**.



InterSystems IRIS Text Analytics

Natural Language Processing

IRIS Text Analytics permet aux développeurs de construire des applications à partir des données non-structurées contenues dans les documents de textes internes ou externes à l'entreprise (rapports, contrats, emails, notes, blogs, réseaux sociaux, presse en ligne, dépêches, etc.)

Il s'agit d'un ensemble d'API pour:

- la collecte de documents depuis une variété de sources
- le traitement sémantique des textes
- l'enrichissement
- la catégorisation de texte

A graphic with a dark blue background. The text "EMBEDDING" is in yellow, "TEXT ANALYSIS" is in white, "INTO YOUR" is in white, and "SOLUTIONS" is in yellow. The text is arranged in four lines, centered horizontally.

EMBEDDING
TEXT ANALYSIS
INTO YOUR
SOLUTIONS

Analyses par croisements de concepts

TableConceptsSets

Sets

diabète type 2	5	9	...		
hématocrite	6	16	...		
hématome	3		...		
hypertension	54		...		
insuffisance cardiaque	14	18	...		
Metformine	3	4	...		
pression atriale	1		...		
pression partielle	6	10	...		
pulmonaire	8	34	...		
ventricule gauche	9	17	...		
VOUX	4		...		

Overlapping sources for sets 'diabète type 2' and 'hypertension'

467:
... Mode de sortie : sortie / Destination : Clinique de Joli-Mont ... • **Hypertension artérielle**. ... • **Hypertension artérielle**. ... Il s'agit du patient de 74 ans connu pour le **diabète de type 2 non-daté**. ... L'examen artériel des membres inférieurs de contrôle a montré des valeurs de pression post-opératoires ...

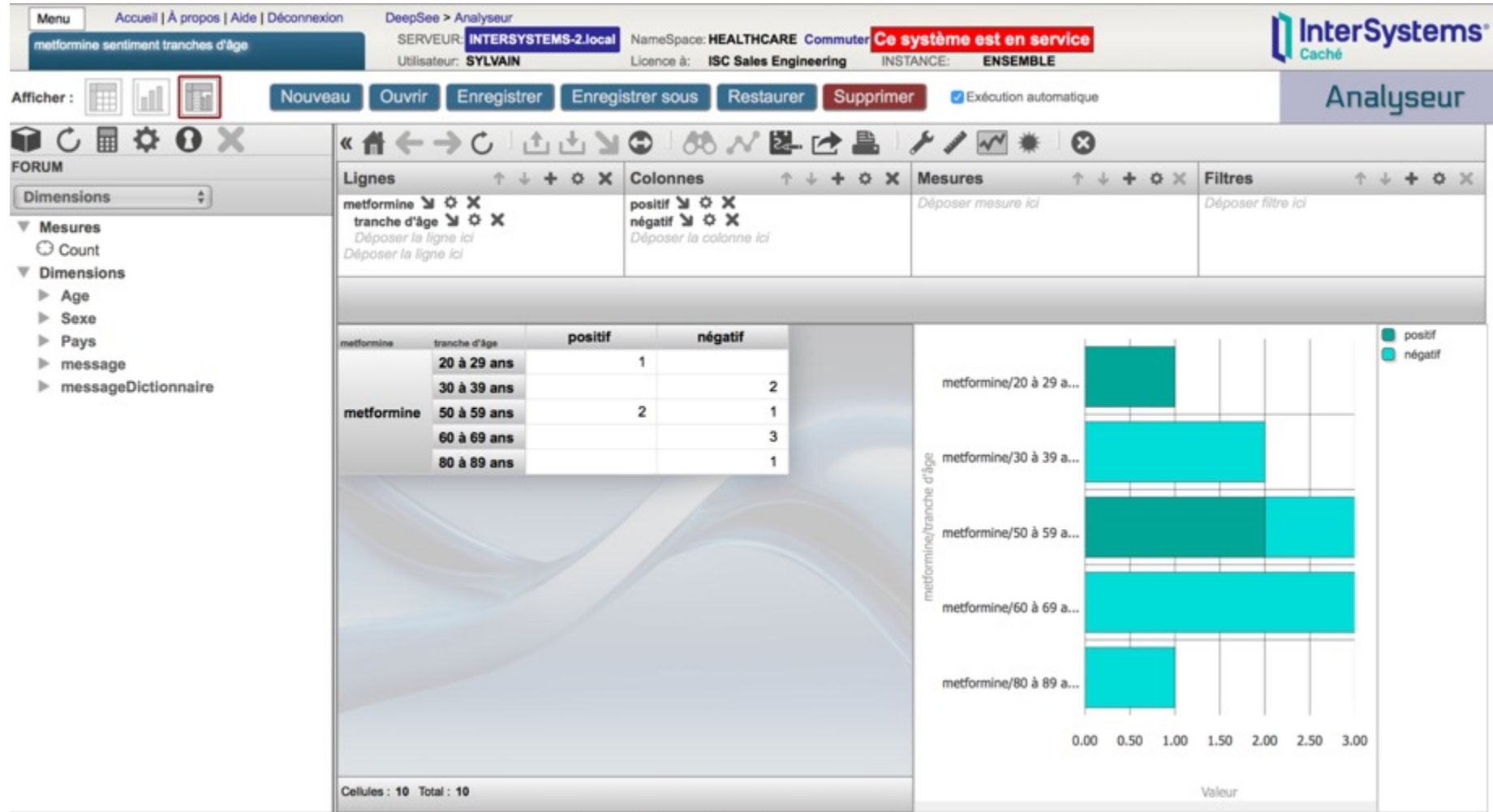
364:
... **Hypertension artérielle**. ... - **Diabète sucré type 2 non insulino-requérant** (avec rétinopathie, IRC et PNP). ... Pas de signe de mise en jeu sous réserve d'une artère cérébrale antérieure droite non perçue. ... éventuellement d'une rééducation à l'Hôpital de Beau-Séjour de préférence, pour des raisons de ... dose de benzodiazépines le 31.01.2012, alors qu'il est fébrile et en cours de traitement pour une ...

360:
... Mode de sortie : sortie / Destination : Clinique de Joli-Mont ... • **Hypertension artérielle**. ... • **Hypertension artérielle**. ... Il s'agit du patient de 74 ans connu pour le **diabète de type 2 non-daté**. ... L'examen artériel des membres inférieurs de contrôle a montré des valeurs de pression post-opératoires ...

164:
... <FORMIDOC_7 LIBELLE="motif d'hospitalisation">Motif d'hospitalisation : Cure d'événement par laparotomie élective, prévue initialement le 25.11.2011, reportée au 4.01.2012. **Hypertension artérielle traitée**. Etat dépressif. Obésité



Analyses croisées: structurées / non-structurées



Analyses avancées - analyse multi sources

The screenshot displays the InterSystems iKnow 'Analyseur' (Analyzer) interface. The top navigation bar includes a menu, user information (SYLVAIN), and system status ('Ce système est en service'). The main workspace is divided into several panes:

- Dimensions:** A tree view on the left showing hierarchical dimensions like 'Age', 'Sexe', 'Pays', 'message', and 'messageDictionnaire'. The 'Age' dimension is expanded, showing age ranges from '20 à 29 ans' to '80 à 89 ans'.
- Lignes:** A table view showing data rows. The first row has ID 2760 and the second row has ID 2764. The table columns are 'ID', 'message', and 'message'.
- Colonnes:** A pane for defining columns, currently showing 'sentiment'.
- Mesures:** A pane for defining measures, currently showing 'Déposer mesure ici'.
- Filtres:** A pane for defining filters, currently showing 'Déposer filtre ici'.

A modal window titled 'Valeur de mesure iKnow' is open, displaying the text: 'Record #1' followed by a paragraph of text: 'ma diabète à essayé 2 fois en l'espace de 2 ans de me donner metformine 850 mais cela n'a pas marché, à cause d'effets indésirables intestinaux je suis donc de nouveau au 500 mg, 1 matin et soir en + de l'insuline.' The text is highlighted in blue. The modal window also includes a 'Longueur du récapitulatif' field set to 'all' and buttons for 'Annuler' and 'OK'.

UIMA Standard - Background

InterSystems IRIS: UIMA Integration

Unified Information Management Architecture

Industry standard for developing & deploying components that *process* unstructured data

- Open-source framework implementation available from Apache
- Serves as part of the core infrastructure for IBM's Watson
- Supports any type of unstructured content, typically text

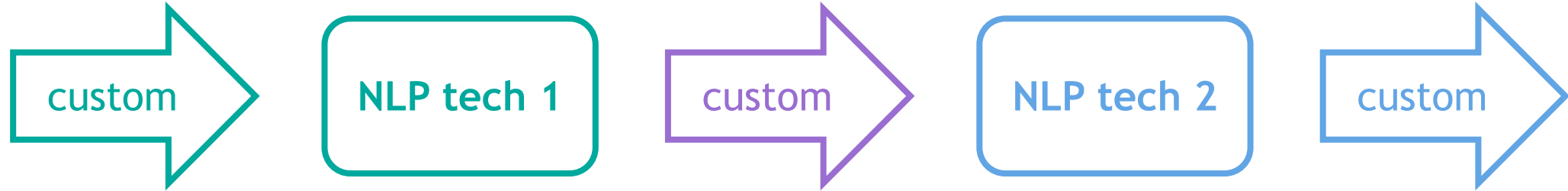
Easily combine different components that each focus on one task

- **Easily**: deployment & scaling is taken care of by the framework
- **Different components**: third-party, custom, ISC, ...
- **One task**: Core NLP, NER, rule-based information extraction, speech-to-text, ...



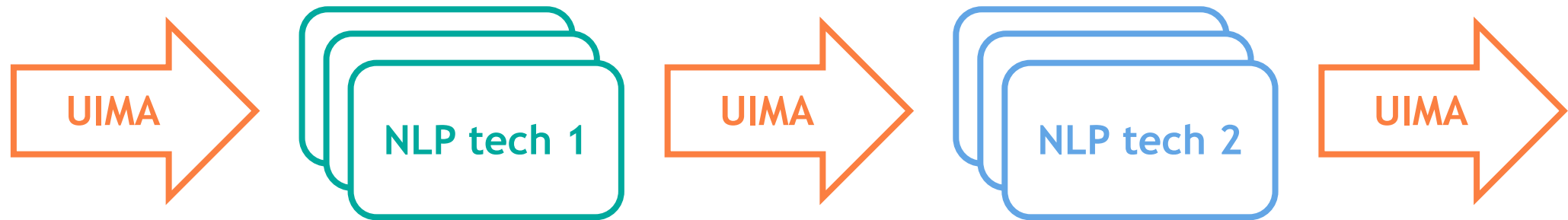
Unstructured Data Processing: Challenges

Combining different NLP tools is challenging at best



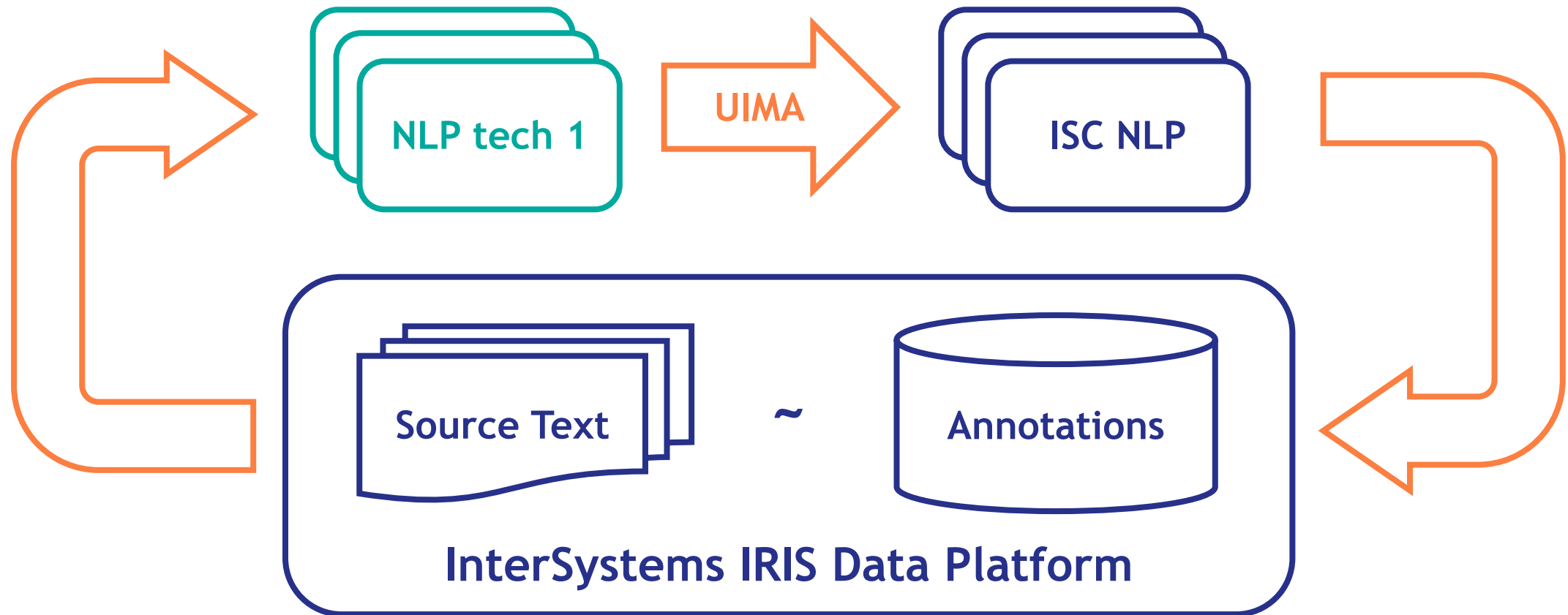
Unstructured Data Processing: UIMA

UIMA standard offers interoperability & scalability for NLP application developers



What's new in InterSystems IRIS?

InterSystems IRIS UIMA Integration marries data platform and UIMA data processing infrastructure for increased developer productivity and a quicker path to analytics



Integrated Support for UIMA Standard

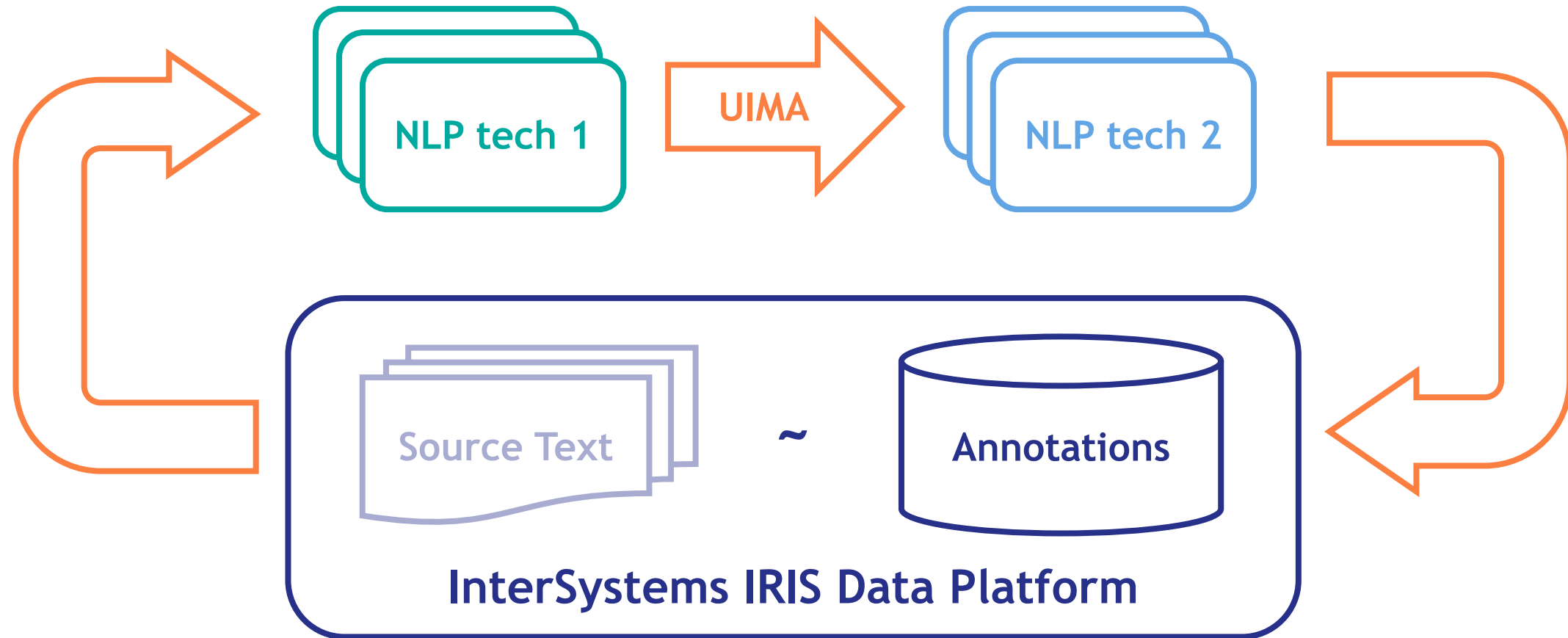
Seamlessly combine diverse annotators

The screenshot displays the UIMA Ruta IDE interface. The main window shows a medical text document titled "i2b2 smoking 2.txt.xmi". The text is annotated with various UIMA annotations, including concepts, negations, and time expressions. Five callout boxes highlight specific features:

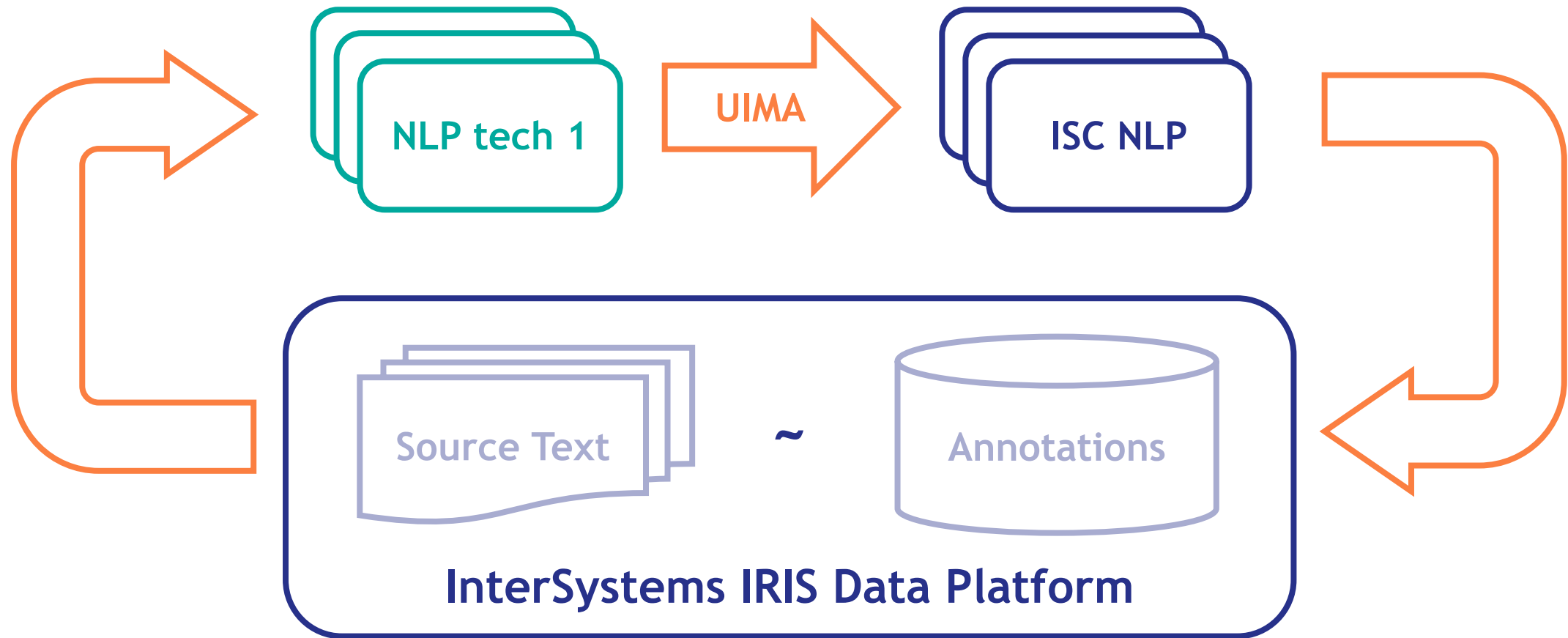
- Custom UIMA Ruta rules**: Points to the "HISTORY OF PRESENT ILLNESS" section header.
- iKnow Concepts**: Points to the "atrial fibrillation" annotation.
- 3rd party Time Annotations (HeidelTime)**: Points to the "two years ago" annotation.
- Simple Dictionary matches**: Points to the "hypertension" annotation.
- iKnow Negation Attributes**: Points to the "no family history of heart disease" annotation.

The right sidebar shows the "Outline" view, listing various annotations and their counts. The bottom status bar indicates the current document and the Ruta Query.

UIMA Annotation Store

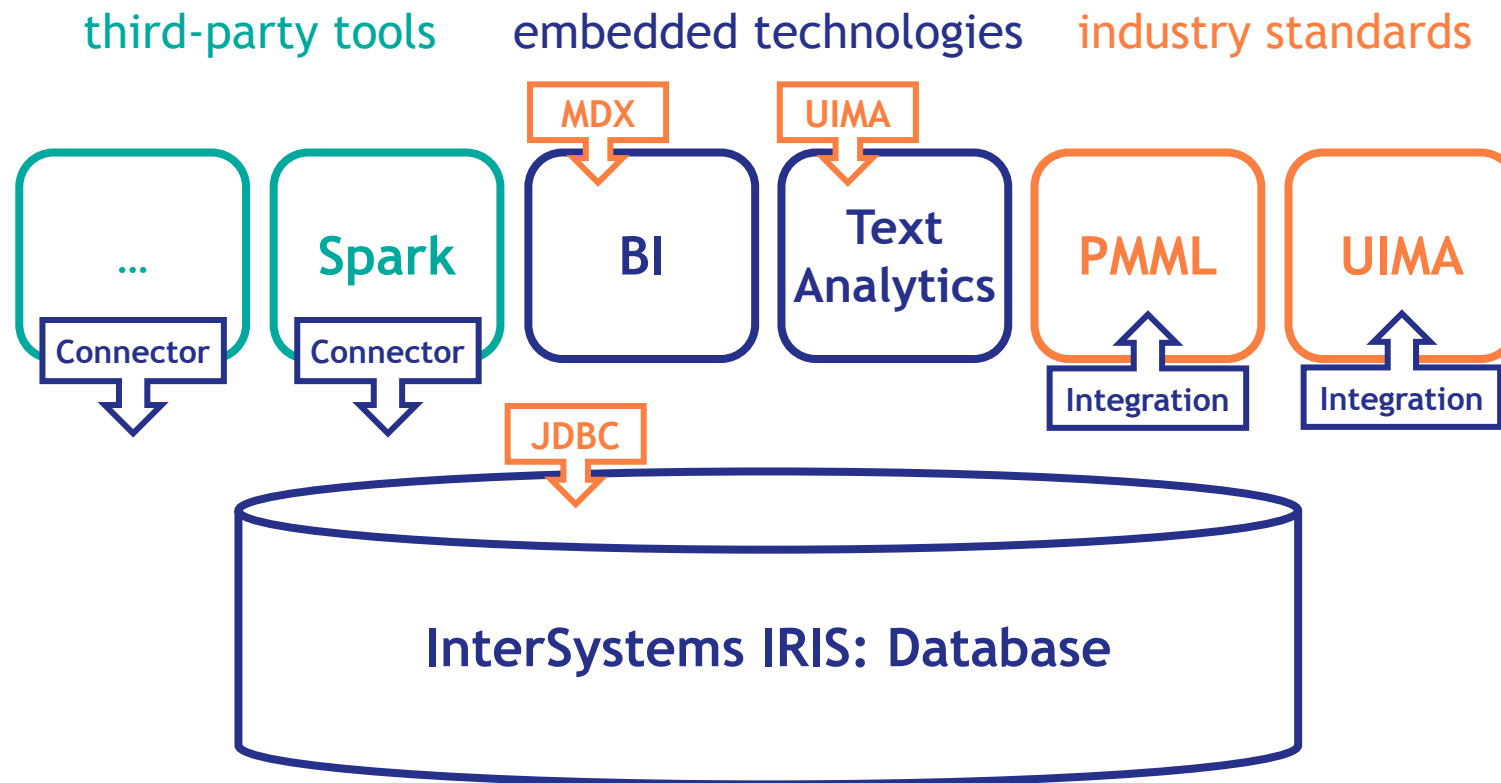


iKnow Analysis Engine



InterSystems IRIS Analytics Roadmap

An **Open Analytics Platform** allows partners and end-users to build solutions and services that leverage best-of-breed analytics technologies, both embedded in the platform and complementary to it.






Developer Community

Ask questions and find answers! Sign up or use your InterSystems Learning / WRC credentials to join the conversation.

<https://community.intersystems.com/>




SOLUTIONS
★ 1



Financial Fraud Prevention with ML and IRIS

Demo of how to apply Machine Learning and Business Rules to


SOLUTIONS
★ 1



Twitter Sentiment Analysis with IRIS

Use IRIS Natural Language Processing and its interoperability


SOLUTIONS



Kano MDM

Kano MDM - is an efficient Master Data Management software product


SOLUTIONS



Background Jobs over ECP

Running a Background Job using JOB command is a well-known feature.


SOLUTIONS



DynamicObject Adapter

An adapter that enables you to "serialize" and "deserialize" class

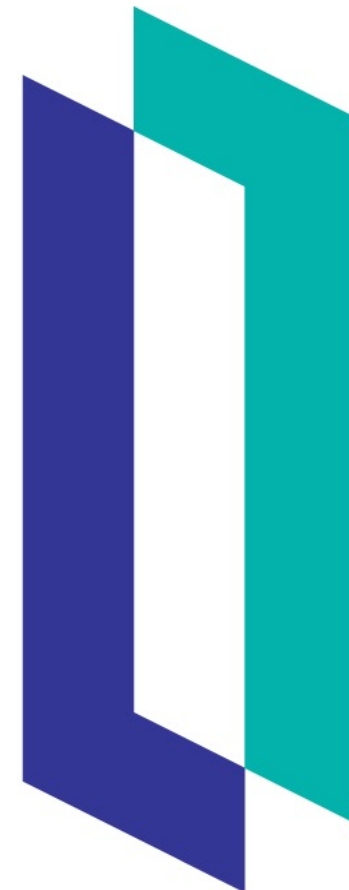
SOLUTIONS



String Datatype - Regular Expression

A string datatype class that implements regular expression

<https://openexchange.intersystems.com/>





Getting Started

Get introduced to freedom of choice with InterSystems IRIS® data platform.



InterSystems IRIS Release Notes

Review new features in this release.



Architecture

Create a resilient architecture that meets your business needs.



Deployment and Installation

Enjoy a choice of deployment options for your solutions.



Application Development

Create innovative applications with InterSystems IRIS, with your choice of languages and data access modes.



Analytics

Combine your favorite tools and technologies for data exploration, business intelligence, and prediction.



System Interoperability

Connect systems with our native integration and interoperability features.



Management and Monitoring

Manage the database, security, backups, and configuration settings.

<https://docs.intersystems.com/>

Documentation





Welcome to the InterSystems Learning System

See at a glance how the InterSystems Learning System can help you learn anytime, anywhere.



Introduction to InterSystems Products and Technologies

New to InterSystems? Start here for a quick look at the different products InterSystems offers.



Featured Overviews

Get a high-level introduction to a particular product or technology with these short, overview materials.



InterSystems IRIS QuickStarts

Explore the benefits of InterSystems IRIS in a five-minute interactive QuickStart.

<https://learning.intersystems.com>





Merci

