

Extracting Info from CAO Documents

BeCode Use Case - December 2022

Agenda

The Stakeholders	5'
CAO Deep-dive	20'
The Challenge	20'
Q&A	10'

"The Consultant"



Sarah Michels

Manager



Jonas Vanden Branden

Senior Advisor



Kristof Bosmans

Advisor



Sara Silvente

Advisor



Yasser Barona

Advisor

KPMG Belgium - Lighthouse Capabilities

Data Strategy & Management



Creating a vision and strategy around data, supporting the digital transformation of our clients, drive data monetization and help building trust in data.

- Data Strategy & Monetization
- Data-Driven Transformation
- Governance & Organisation
- (Master) Data Management



BI & Analytics



Enabling businesses and people by delivering actionable insights to help driving business decision making and defining the right KPIs

- KPI / Performance Management
- Data Platform
- Dashboarding & visualisation
- ERP Analytics
- Process Mining



Advanced Analytics & Artificial Intelligence



Leveraging advanced analytics by **understanding the strengths and limitations** of such techniques, to **maximize return** from analytics and AI initiatives.

- Use Case Exploration
- Predictive & Prescriptive Modelling
- Machine Learning
- Model Review
- MLOps



Intelligent Automation and NewTech



Create the **strategy** and **operating model** for **sustainable** human/robot co-existence through radical innovation and leveraging new technologies.

- Robotics Process Automation (RPA)
- Intelligent Document Processing
- BPMN Automation
- Cognitive Automation and
- Digital Assistants

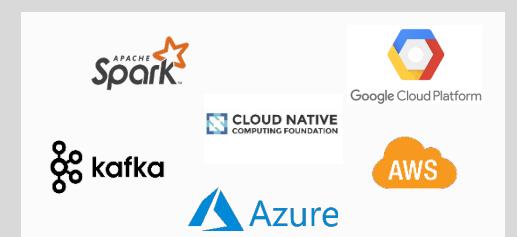


Digital Solution Architecture

Digital Solution Architecture is a modern business and technology architecture practice that is driven by customer experience and recent technologies. Its main objective is to provide the technical support of the 4 pillars here above and to drive the Next Generation Architectures.



- (Big) Data Architecture
- Solution Architecture
- DataOps & DevOps and Agile Software Development
- API Management



"The Client"





CLA Deep-dive

By KPMG Law

Agenda

Sources of Belgian employment law

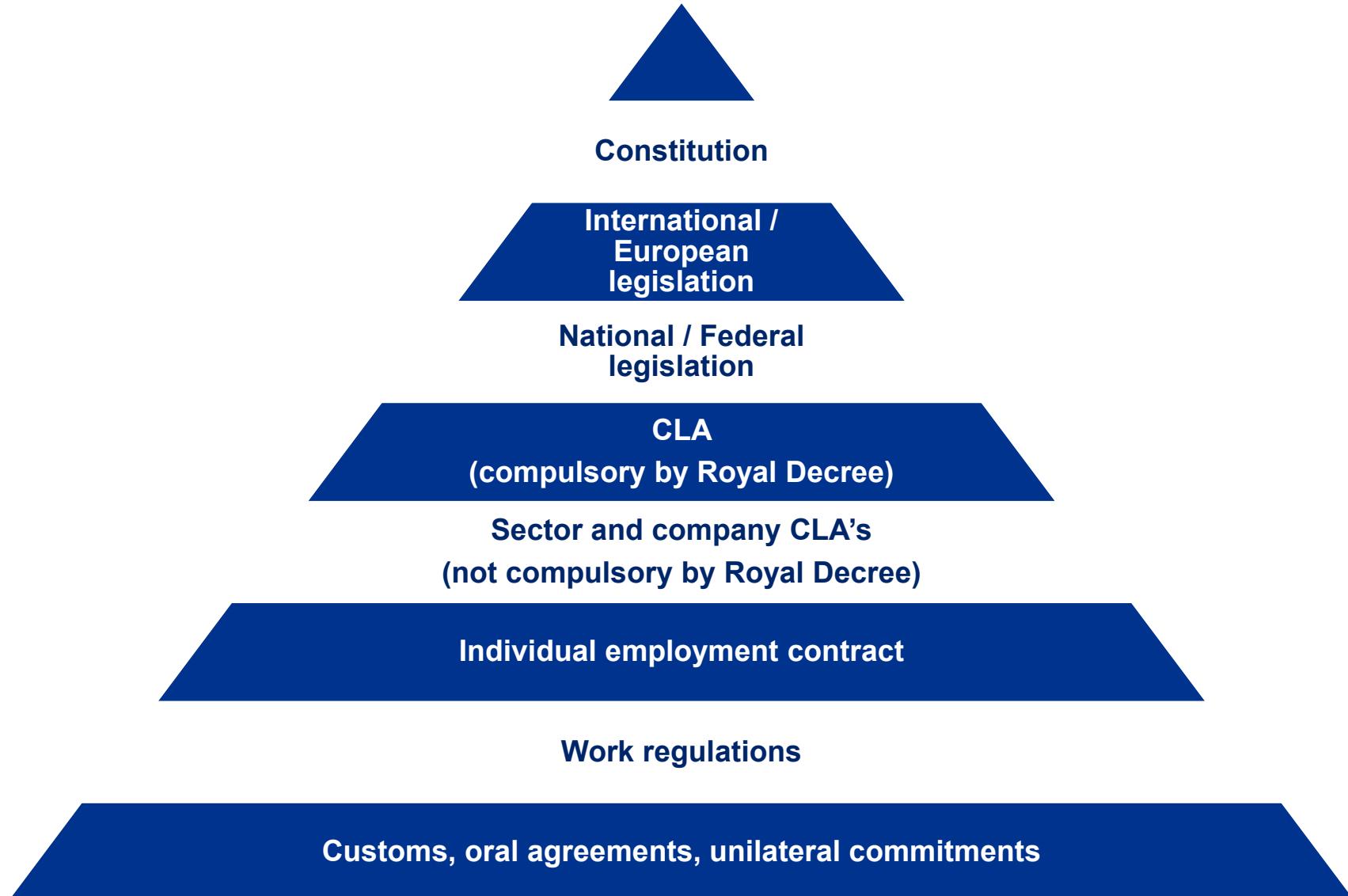
Collective Labour Agreements (CLA)





Sources of Belgian employment law

Sources of Belgian employment law





Collective labour agreements

Collective labour agreement: notion

Written agreement

between one (or more) employer
(organization(s)) and one (or
more) employee representatives

COLLECTIVE BARGAINING

Salaries Employment Health
Grievance Protection Law
Employees Representative
Issues
Healthcare Cost
Process Negotiation
Agreement Productive
Working Hours Terms Rules
Employer Union Wages Group
Overtime Safety Establish Cost

Employment conditions such as
salary, working time...

No interference of government

Collective labour agreements: different types

(1) CLA concluded on national level (*within the National Labour Council*)



National CLA: example

Tekst	Inhoudstafel	Begin
<p>HOOFDSTUK I. - Draagwijde.</p> <p>Artikel 1. § 1. Deze collectieve arbeidsovereenkomst heeft tot doel het grondrecht van de werknemers op de verhödiging van hun persoonlijke levensfeer in het kader van de dienstbetrekking te waarborgen door, rekening houdend met de behoeften voor een goede werking van de onderneming, te bepalen voor welke doeleinden en onder welke proportionaliteits- en transparantievoorzieningen een controle op de elektronische on-linecommunicatiegegevens kan worden geïnstalleerd en volgens welke regels de individualisering van deze gegevens is toegestaan.</p> <p>Zij doet geen afbreuk aan gunstige bepalingen op het niveau van de paritaire comités en de ondernemingen.</p> <p>§ 2. Deze collectieve arbeidsovereenkomst heeft betrekking op de regels voor de toegang tot en/of het gebruik van de elektronische on-linecommunicatiemiddelen van de onderneming, welke regels het prerogatief van de werkgever zijn. Deze overeenkomst laat dus de eventueel in de ondernemingen geldende regels en praktijken op het gebied van voorlichting en zelfs raadpleging onverlet.</p> <p>Zij doet evenmin afbreuk aan de in de ondernemingen bestaande regels en praktijken wat de uitoefening van de vakbondsactiviteiten betreft.</p> <p>Commentaar: Deze collectieve arbeidsovereenkomst ligt in de sfeer van artikel 109ter D van de wet van 21 maart 1991 betreffende de hervorming van sommige economische overheidsbedrijven en de wet van 8 december 1992 tot bescherming van de persoonlijke levensfeer ten opzichte van de verwerking van persoonsgegevens; zij heeft niet betrekking tot de controle op de elektronische on-linecommunicatiegegevens die onder het toepassingsgebied ervan valt, alleen tot doel de toepassing van deze wetten in het kader van de dienstbetrekking te waarborgen en de beginselen ervan te concreteren, inzonderheid de beginselen bepaald in de wet van 8 december 1992, namelijk het finaliteitsbeginsel, het proportionaliteitsbeginsel en het transparantiebeginsel.</p> <p>Zij heeft tot doel een strikt evenwicht te stand brengen tussen deze beginselen van de voornoemde wetten van 21 maart 1991 en 8 december 1992, die als essentiële garanties worden beschouwd voor de bescherming van de persoonlijke levensfeer van de werknemer op de werkplek, en de prerogatieven die het de werkgever mogelijk maken de goede werking van de onderneming te waarborgen.</p> <p>HOOFDSTUK II. - Definitie.</p> <p>Art. 2. Voor de toepassing van deze collectieve arbeidsovereenkomst wordt onder elektronische on-linecommunicatiegegevens verstaan, de elektronische on-linecommunicatiegegevens s.l. ongeacht de drager via welke een en ander door een werknemer wordt overgebracht of ontvangen in het kader van de dienstbetrekking.</p> <p>Commentaar: Deze collectieve arbeidsovereenkomst wil een kader vaststellen dat ruim genoeg is om alle on-linetechnologieën te omvatten, rekening houdend met de toenemende verwevenheid en de snelle ontwikkeling van deze technologieën en de drager die wordt gebruikt. Zij is dan ook van toepassing ongeacht deze drager.</p> <p>Zij viseert bovenindien de elektronische on-linecommunicatie, zowel intern als extern.</p> <p>HOOFDSTUK III. - Verbitamen van de partijen.</p> <p>Art. 3. De ondertekende partijen bevestigen de volgende beginselen :</p> <ul style="list-style-type: none"> - de werknemers erkennen het beginsel volgens hetwelk de werkgever het recht heeft controle uit te oefenen op het werkinstrument en op het gebruik dat de werknemer ervan maakt in het kader van de uitvoering van zijn contractuele verplichtingen, ook wanneer dit gebruik binnen de persoonlijke levensfeer valt, rekening houdend met de in deze overeenkomst bepaalde toepassingsregels; - de werkgevers verhouden het recht van de werknemers op bescherming van hun persoonlijke levensfeer in het kader van de dienstbetrekking en de rechten en verplichtingen die er voor iedere partij uit voortvloeien. <p>HOOFDSTUK IV. - Toepassingsregels.</p> <p>Afdeling 1. - Regels voor de controle op de elektronische on-linecommunicatiegegevens.</p> <p>Onderafdeling 1. - Algemene bepalingen.</p> <p>Art. 4. De controle op de elektronische on-linecommunicatiegegevens wordt slechts toegestaan voorzover voldaan wordt aan de finaliteits- en proportionaliteitsbeginselen als bepaald in de punten 1 en 2 van onderafdeling 2 verderop, alsook aan het transparantiebeginsel, zoals gewaarborgd door de procedurevoorwaarden in punt 3 van dezelfde onderafdeling.</p> <p>Commentaar: De werkgever heeft de mogelijkheid controle te verrichten op persoonlijke elektronische on-linecommunicatiegegevens, voor een of meer doeleinden die door deze overeenkomst als gerechtvaardigd worden erkend.</p> <p>Onderafdeling 2. - Beginselen.</p> <p>1. Finaliteitsbeginsel.</p> <p>Art. 5. § 1. De controle op de elektronische on-linecommunicatiegegevens is maar toegestaan mits een of meer van de volgende doeleinden worden nastreefd :</p> <ol style="list-style-type: none"> 1^o het voorkomen van ongoedoorloofd of lasterlijke feiten, feiten die strijdig zijn met de goede zeden of de waardigheid van een andere persoon kunnen schaden; 2^o de bescherming van de handels- en financiële belangen van de onderneming die vertrouwelijk zijn aangehoekt het tegengaan van erme in strijd zijnde praktijken; 3^o de veiligheid en/of de goede technische werking van de IT-netwerksystemen van de onderneming, met inbegrip van de controle op de kosten die ermee gepaard gaan alsook de fysieke bescherming van de installaties van de onderneming; 4^o het te goede trouw naleven van de in de onderneming geldende beginselen en regels voor het gebruik van on-linetechnologieën. <p>§ 2. De werkgever omschrijft duidelijk en expliciet de doelstelling(en) van de controle.</p> <p>Commentaar: a. Ongoorloofde of lasterlijke feiten, feiten die strijdig zijn met de goede zeden of de waardigheid van een andere persoon kunnen schaden, als bedoeld in § 1.^o van dit artikel, kunnen met name bestaan in het kraken van computers, waaronder het op ongoedoorloofde wijze kennis nemen van elektronische on-linecommunicatiegegevens inzake personeelsbeheer of vertrouwelijke medische bestanden of nog het raadplegen van pornografische of pedofiele sites alsook van sites die aanzetten tot discriminatie, rassenscheiding, haat of geweld jegens een groep, een gemeenschap of de ledenvan een ras, huiskleur, afkomst, religie of nationale of etnische afstamming van deze ledenvan een sommigen onder hen.</p> <p>b. De praktijken die in strijd zijn met de financiële, economische en handelsbelangen van de onderneming, als bedoeld in § 1.^o van dit artikel, kunnen met name de vorm aannemen van afbrekende reclame als bepaald in artikel 23, 6^o van de wet van 14 juli 1991 betreffende de handelspraktijken en de voorlichting en bescherming van de consument, verspreiding van bestanden en schending van zakengemeinen, met inbegrip van onderzoek en ontwikkeling, productieprocessen en alle mogelijke vertrouwelijke gegevens.</p> <p>b. Paragraaf 2 van dit artikel laat de mogelijkheid onverlet om controles op elektronische on-linecommunicatiegegevens uit te voeren voor opleidingsdoeleinden, vermits het hier niet om toezicht gaat.</p> <p>In geval van geheim toezicht op de elektronische on-linecommunicatiegegevens zijn de bepalingen van het Strafwetboek van toepassing en deze vorm van toezicht kan alleen worden ingevoerd in overeenstemming met de voorschriften van het Wetboek van Strafvordering.</p> <p>Ongeoorloofde of lasterlijke feiten, feiten die strijdig zijn met de goede zeden of de waardigheid van een andere persoon kunnen schaden, als bedoeld in § 1.^o van dit artikel, kunnen met name bestaan in het kraken van computers, waaronder het op ongoedoorloofde wijze kennis nemen van elektronische on-linecommunicatiegegevens inzake personeelsbeheer of vertrouwelijke medische bestanden of nog het raadplegen van pornografische of pedofiele sites alsook van sites die aanzetten tot discriminatie, rassenscheiding, haat of geweld jegens een groep, een gemeenschap of de ledenvan een ras, huiskleur, afkomst, religie of nationale of etnische afstamming van deze ledenvan een sommigen onder hen.</p> <p>De praktijken die in strijd zijn met de financiële, economische en handelsbelangen van de onderneming, als bedoeld in § 1.^o van dit artikel, kunnen met name de vorm aannemen van afbrekende reclame als bepaald in artikel 23, 6^o van de wet van 14 juli 1991 betreffende de handelspraktijken en de voorlichting en bescherming van de consument, verspreiding van bestanden en schending van zakengemeinen, met inbegrip van onderzoek en ontwikkeling, productieprocessen en alle mogelijke vertrouwelijke gegevens.</p> <p>b. Paragraaf 2 van dit artikel laat de mogelijkheid onverlet om controles op elektronische on-linecommunicatiegegevens uit te voeren voor opleidingsdoeleinden, vermits het hier niet om toezicht gaat.</p> <p>Commentaar: La possibilité est reconnue à l'employeur d'effectuer un contrôle, sur les données de communication électroniques en réseau à caractère personnel pour une ou plusieurs finalités reconnues comme légitimes aux termes de la présente convention.</p> <p>Section 1^o. — Modalités de contrôle des données de communication électroniques</p> <p>Sous-section 1^o. — Dispositions générales</p> <p>Art. 4. Le contrôle des données de communication électroniques en réseau n'est autorisé que pour autant qu'il soit satisfait aux principes de finalité et de proportionnalité énoncés aux points 1 et 2 de la sous-section 2 ci-après ainsi qu'au principe de transparence que les conditions de procédures définies dans le point 3 de la même sous-section ci-après, visent à garantir.</p> <p>Commentaire: De werkgever heeft de mogelijkheid controle te verrichten op persoonlijke elektronische on-linecommunicatiegegevens, voor een of meer doeleinden die door deze overeenkomst als gerechtvaardigd worden erkend.</p>	<p>§ 2. Deze collectieve arbeidsovereenkomst heeft geen betrekking op de regels voor de toegang tot en/of het gebruik van de elektronische on-linecommunicatiemiddelen van de onderneming, welke regels het prerogatief van de werkgever zijn. Deze overeenkomst laat dus de eventueel in de ondernemingen geldende regels en praktijken op het gebied van voorlichting en zelfs raadpleging onverlet.</p> <p>Zij doet evenmin afbreuk aan de in de ondernemingen bestaande regels en praktijken wat de uitoefening van de vakbondsactiviteiten betreft.</p> <p>Commentaire: Elle ne porte pas non plus préjudice aux règles et pratiques existant dans les entreprises en ce qui concerne l'exercice des activités syndicales.</p> <p>Section 2. — Application des règles et pratiques existantes dans les entreprises</p> <p>§ 2. La présente convention collective de travail ne vise pas les modalités d'accès et/ou l'utilisation des moyens de communication électronique en réseau de l'entreprise au sens de la définition de l'employeur. Cette convention laisse donc en l'état les règles et pratiques d'information voire de consultation éventuellement en vigueur dans les entreprises.</p> <p>Elle ne porte pas non plus préjudice aux règles et pratiques existant dans les entreprises en ce qui concerne l'exercice des activités syndicales.</p> <p>Commentaire: La présente convention collective de travail s'inscrit dans le cadre de l'article 109ter D de la loi du 21 mars 1991 portant réforme de certaines entreprises publiques économiques et de la loi du 8 décembre 1992 relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel; elle a pour seul objet, en ce qui concerne le contrôle des données de communication électroniques en réseau qui relèvent des changements d'application, de garantir, dans la relation de travail, l'application de ces lois, d'en concrétiser les principes et notamment par la loi du 8 décembre 1992, le principe de finalité, le principe de proportionnalité et le principe de transparence.</p> <p>Elle s'attache à établir un strict équilibre entre d'une part, ces principes contenus dans les lois précitées du 21 mars 1991 et du 8 décembre 1992, qui sont des garanties jugées nécessaires pour assurer la protection de la vie privée du travailleur sur le lieu de travail, et d'autre part, les prérogatives de l'employeur lui permettant d'assurer le bon fonctionnement de l'entreprise.</p> <p>CHAPITRE II. — Définition</p> <p>Art. 2. Pour l'application de la présente convention collective de travail, sont entendues les données de communication électroniques en réseau, dont toutes les données de communication électroniques transitant par rapport, entendues au sens large et indépendamment du support par lequel elles sont transmises ou reçues par un travailleur dans le cadre de la relation de travail.</p> <p>CHAPITRE III. — Définition</p> <p>Art. 3. Les organisations signataires affirment les principes suivants :</p> <ul style="list-style-type: none"> - les travailleurs reconnaissent le principe selon lequel l'employeur dispose d'un droit de contrôle sur l'outil de travail et sur l'utilisation de cet outil par le travailleur dans le cadre de l'exécution de ses obligations contractuelles, y compris lorsque cette utilisation relève de la sphère privée, compte tenu des modalités d'application prévues par la présente convention; - les employeurs respectent le droit des travailleurs à la protection de leur vie privée dans le cadre de la relation de travail et des droits et obligations qui collaborent à assurer la sécurité sociale. <p>CHAPITRE III. — Engagement des parties</p> <p>Art. 3. Les organisations signataires affirment les principes suivants :</p> <ul style="list-style-type: none"> - les travailleurs reconnaissent le principe selon lequel l'employeur dispose d'un droit de contrôle sur l'outil de travail et sur l'utilisation de cet outil par le travailleur dans le cadre de l'exécution de ses obligations contractuelles, y compris lorsque cette utilisation relève de la sphère privée, compte tenu des modalités d'application prévues par la présente convention; - les employeurs respectent le droit des travailleurs à la protection de leur vie privée dans le cadre de la relation de travail et des droits et obligations qui collaborent à assurer la sécurité sociale. <p>CHAPITRE IV. — Modalités d'application</p> <p>Section 1^o. — Modalités de contrôle des données de communication électroniques</p> <p>Sous-section 1^o. — Dispositions générales</p> <p>Art. 4. Le contrôle des données de communication électroniques en réseau n'est autorisé que pour autant qu'il soit satisfait aux principes de finalité et de proportionnalité énoncés aux points 1 et 2 de la sous-section 2 ci-après ainsi qu'au principe de transparence que les conditions de procédures définies dans le point 3 de la même sous-section ci-après, visent à garantir.</p> <p>Commentaire: La possibilité est reconnue à l'employeur d'effectuer un contrôle, sur les données de communication électroniques en réseau à caractère personnel pour une ou plusieurs finalités reconnues comme légitimes aux termes de la présente convention.</p>	<p>Example (2002)</p>



Collective labour agreements: different types

(2) CLA concluded on industry level (*between the trade unions and employer organizations*)



BBTK
ABVV



CGSLB
SYNDICAT LIBERAL

ACLVB
LIBERALE VAKBOND



unizo
VERSTERKT ONDERNEMERS

VOKA
Vlaams netwerk van ondernemingen



uwe
union wallonne des entreprises

Beci
BELGAESSE ENTREPRISES COMMERCE ET INDUSTRIE

unisoc

Boerenbond FWA
AGROFront

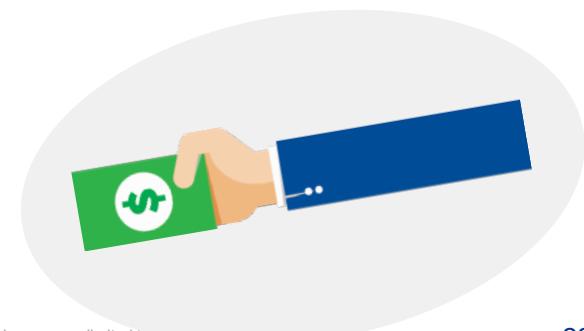
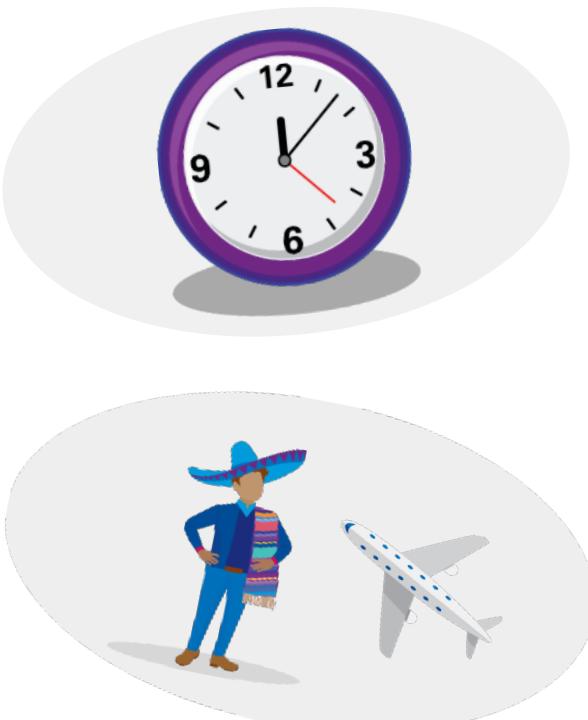
Collective labour agreements: different types

(3) CLA concluded on a company level (*between the trade unions and an individual employer*)



Collective labour agreement: which topics?

What kind of topics are regulated by CLA's?



Collective labour agreement: important topics

Subjects can be found on the website of the [FOD WASO / SPF ETCS](#)

Cao zoeken

Inhoud formulier verwijderen

Paritair comité

Vul het nummer van het paritair comité of de benaming in

Selecteer in de volledige lijst

Thema

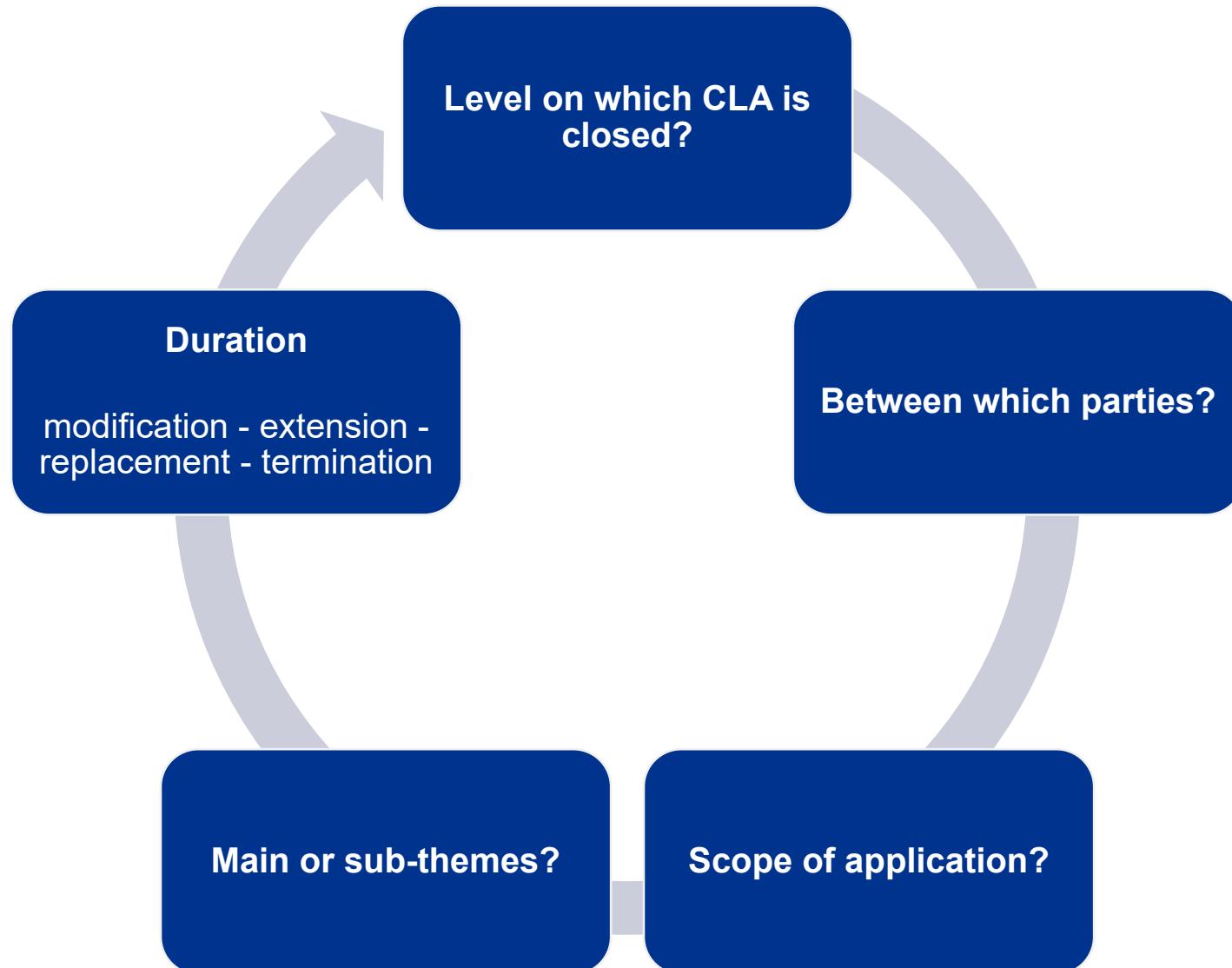
- Verloning
- Functieclassificatie
- Arbeidsduur
- Aanwerving en opleiding
- Werkgelegenheidsmaatregelen en doelgroepen
- Gezondheid en welzijn

Selecteer een thema in de kolom(men)

Woord(en) in de tekst

Zet een spatie tussen de woorden. Gebruik enkel cijfers en letters.

Collective labour agreement: important content



Example

Joint committee

Subject

Date

Aanvullend Paritair Comité voor de Bedienden	Commission paritaire auxiliaire pour employés
<i>Collectieve arbeidsovereenkomst van 9 juni 2016</i>	<i>Convention collective de travail du 9 juin 2016</i>
Arbeidsduur	Durée du travail
Preamble	Préambule
<p>De cao's gesloten op het niveau van het PC 218, zoals toepasselijk en in werking aan de vooravond van de inwerkingtreding van het besluit houdende de benoeming van de leden van het PC 200, werden bij cao van 1.04.2015 (registratienr. 126638) overgedragen naar het PC 200.</p> <p>Deze cao's werden bijgevolg integraal van toepassing op het niveau van het PC 200 en binden de werkgevers en werknemers die ressorteren onder het PC 200.</p> <p>De sluiting van onderhavige cao heeft tot doel om in het kader van de continuïteit de</p>	<p>Les CCT conclues au niveau de la CP 218, telles qu'applicables et en vigueur à la veille de l'entrée en vigueur de l'arrêté portant nomination des membres de la CP 200, ont été transférées à la CP 200 par la CCT du 1.04.2015 (numéro d'enregistrement 126638).</p> <p>Ces CCT sont donc intégralement d'application au niveau de la CP 200 et elles lient les employeurs et les travailleurs qui relèvent de celle-ci.</p> <p>La conclusion de la présente CCT a pour objectif, dans le cadre de la continuité,</p>

Example

Applicable to
which employees
and employers?

HOOFDSTUK I – <i>Toepassingsgebied</i>	CHAPITRE Ier – <i>Champ d'application</i>
Artikel 1. Deze collectieve arbeidsovereenkomst is van toepassing op de werkgevers en bedienden van de ondernemingen die tot de bevoegdheid van het Aanvullend Paritair Comité voor de Bedienden (200) behoren.	Article 1er. La présente convention collective de travail s'applique aux employeurs et employés des entreprises relevant de la compétence de la Commission paritaire auxiliaire pour employés (200).
Onder "bedienden" wordt verstaan, de mannelijke en vrouwelijke bedienden.	On entend par "employés", les employés et les employées.
HOOFDSTUK II – <i>Wekelijkse arbeidsduur op jaarbasis</i>	CHAPITRE II – <i>Durée du travail hebdomadaire sur base annuelle</i>
Art. 2. De gemiddelde arbeidsduur, berekend op jaarbasis, is vastgesteld op 38 uur.	Art. 2. La durée du travail hebdomadaire moyenne, calculée sur base annuelle, est de 38 heures.

Example

Duration

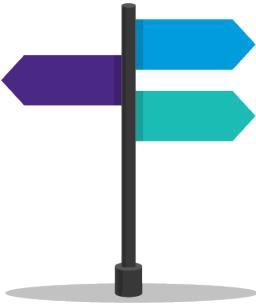
	HOOFDSTUK VI – <i>Duur</i>	CHAPITRE VI – <i>Durée</i>
	Art. 8. Onderhavige collectieve arbeidsovereenkomst treedt in werking op 1 juli 2016 en wordt gesloten voor een onbepaalde duur.	Art. 8. La présente convention collective de travail entre en vigueur le 1 juillet 2016 et est conclue pour une durée indéterminée.
	Ze kan door één der partijen opgezegd worden, mits een opzegging van 3 maanden, gericht bij aangetekende brief aan de Voorzitter van het Aanvullend Paritair Comité voor de Bedienden. <i>en aux deux ouvertekenende organisations,</i>	Elle peut être dénoncée par une des parties, moyennant un préavis de 3 mois, adressé par courrier recommandé au Président de la Commission paritaire auxiliaire pour les employés et aux organisations signataires.

How can the parties terminate the CLA?



The Challenge

Problem Statement



The client analyzes hundreds of **new collective bargaining agreements (CLA)** on an annual basis, in order to interpret practical impact and implement them if necessary.

Due to the **high volume** (>30 / week) of these **extensive documents**, this is a time consuming process. The task of collecting and processing these new collective agreements needs to be **accelerated and automated**, in order to **enable proactive notification**.

Using the available meta-data and Artificial Intelligence (NLP) the **relevant information** will be **extracted** from the documents. The client needs this information in a **structured way**, so **integrations** with existing platforms and/or new **reporting** are possible.

Relevancy of the BeCode case: Business needs



Internal (employment team)

- Be able to be **the first** to inform clients as from the moment a new/modified CLA is published.
- Be able to **proactively** inform clients (not only on the occasion of an *ad hoc* question)
- Receive a short first overview per change (containing: the topic, the joint committee, amounts...) and a short cut to the relevant CLA in order to select in an **efficient** way the relevant changes

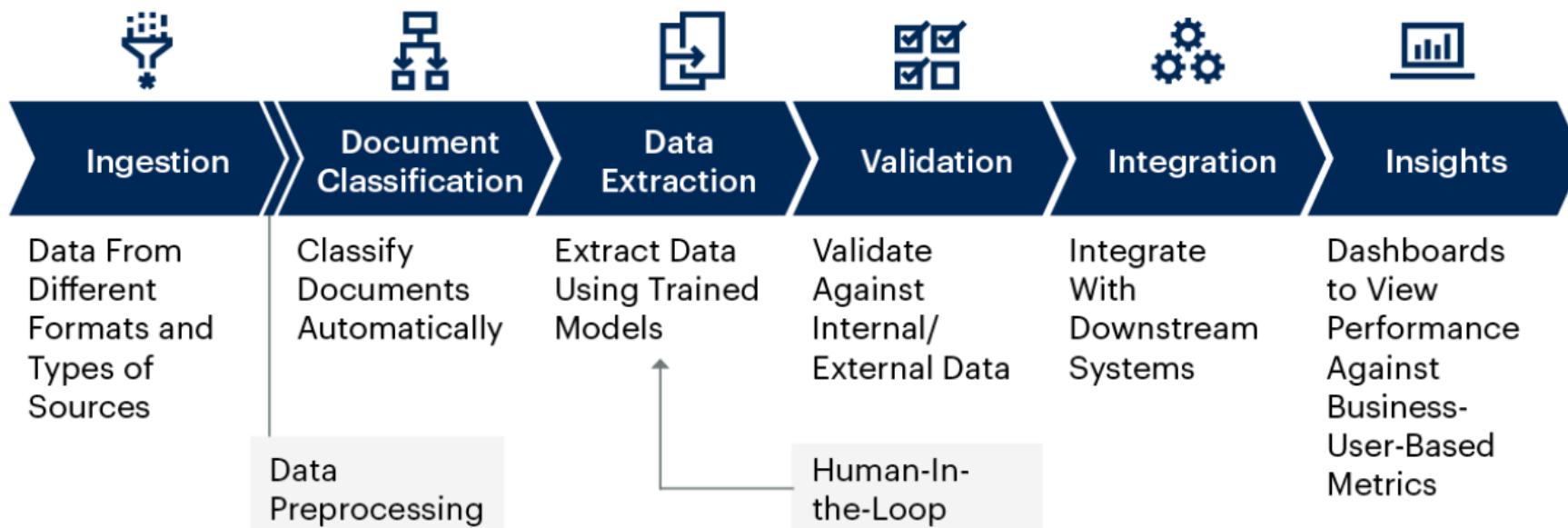
External (clients)

- Be **informed as soon as possible** of latest updates
- Receive updates in **clear language**
- Not to be overloaded by lots of notifications, but only be informed of **relevant** changes



Solution Design: Intelligent Document Processing

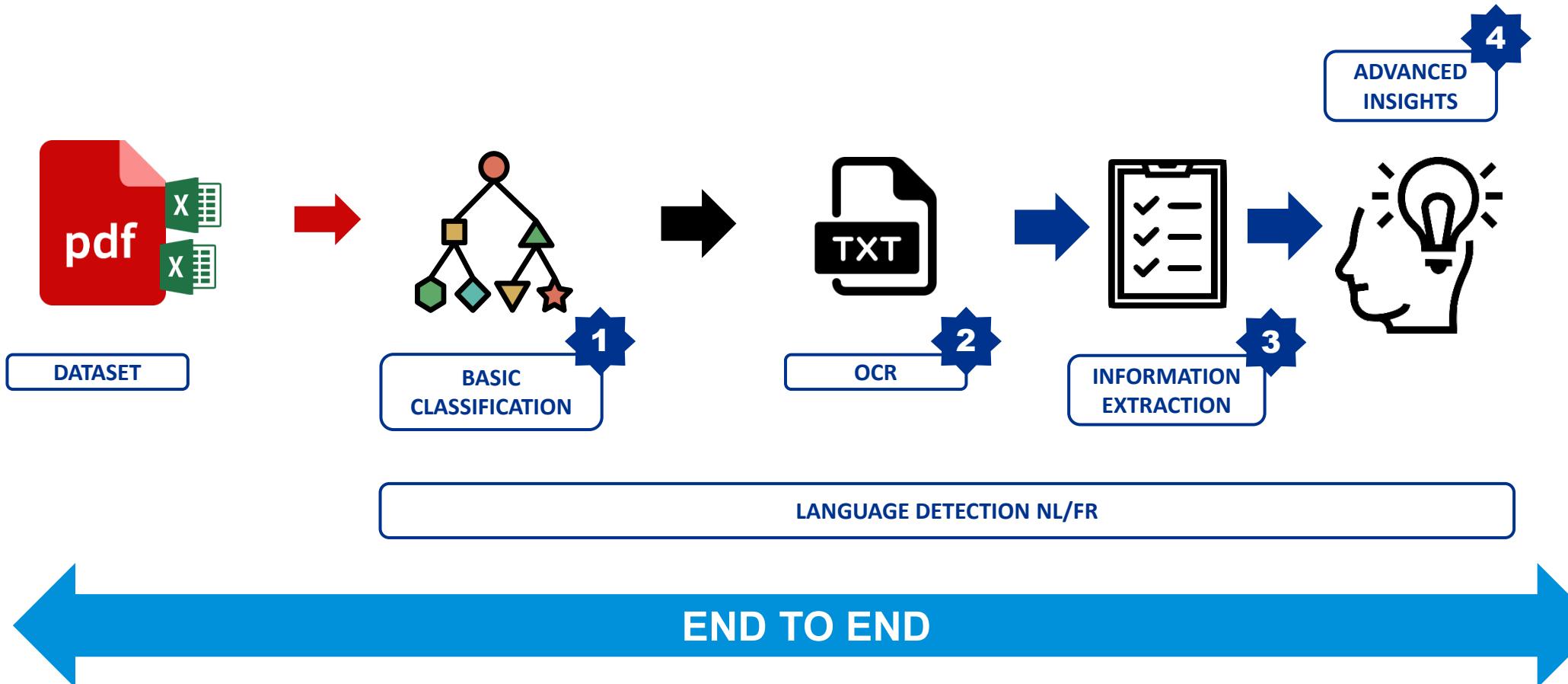
Different Phases in a Baseline IDP Workflow



Source: Gartner

757528_C

Solution Design: Intelligent Document Processing



The Dataset

- ~6000 CLA pdf's (2018- current date)
- Meta Data for each document
- Meta Data for each **joint committee**

Dataset Part A

Dataset Part B

Neerlegging-Dépôt: 21/04/2021
Regist.-Enregistr.: 06/05/2021
N°: 164568/CO/102.02

Paritair Subcomité voor het bedrijf der hardsteengroeven en der groeven van uit te houwen kalksteen in de provincies Luik en Namen

Sous-commission paritaire de l'industrie des carrières de petit granit et de calcaire à tailler des provinces de Liège et de Namur

Collectieve arbeidsovereenkomst van 26 maart 2021

Convention collective de travail du 26 mars 2021

Toekenning van maaltijdcheques

Octroi de titres repas

Rekening houdend met het feit dat er geen restaurant is in *dit* bedrijf werd overeengekomen wat volgt :

HOOFDSTUK I. Toepassingsgebied

CHAPITRE Ier. Champ d'application

Artikel 1. Deze collectieve arbeidsovereenkomst is van toepassing op de werkgevers en de werknemers van de ondernemingen die onder het Paritair Subcomité voor het bedrijf der hardsteengroeven en der groeven van uit te houwen kalksteen in de provincies Luik en Namen ressorteren.

Article 1er. La présente convention collective de travail s'applique aux employeurs et aux travailleurs des entreprises ressortissant à la Sous-commission paritaire de l'industrie des carrières de petit granit et de calcaire à tailler des provinces de Liège et de Namur.

Met "werknemers" worden de werklieden en de werksters bedoeld.

Par "travailleurs" on entend : les ouvriers et les ouvrières.

HOOFDSTUK II. Voorwerp van de overeenkomst

CHAPITRE II. Objet de la convention

Art. 2. Deze overeenkomst wordt gesloten in het raam van het koninklijk besluit tot wijziging van het koninklijk besluit van 28 november 1969 tot uitvoering van de wet van 27 juni 1969 tot herziening van de besluitwet van 28 december 1944 betreffende de maatschappelijke zekerheid der arbeiders. Zij wordt opgesteld met inachtneming van de bepalingen van het koninklijk besluit van 3 februari 1998 betreffende de maaltijdcheques.

Art. 2. La présente convention est conclue dans le cadre de l'arrêté royal modifiant l'arrêté royal du 28 novembre 1969 pris en exécution de la loi du 27 juin 1969 révisant l'arrêté-loi du 28 décembre 1944 concernant la sécurité sociale des travailleurs. Elle est établie en tenant compte des dispositifs de l'arrêté royal du 3 février 1998 concernant les titres repas.

1. Basic Classification

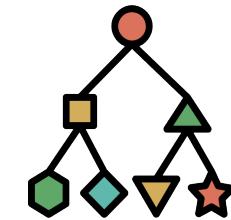
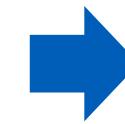
Classify documents into different types or categories

TIPS

- Take advantage of the metadata.
- Apply the basic data analysis stages:
 - Understand
 - Collect
 - Clean
 - Analyze
 - Interpret
- Keep it simple



123456.pdf



Document type

GOAL

Identify document types

2. Text Extraction

Extract the text from the pdf files

TIPS

- Keep ‘newline’ information; the layout matters. (paragraphs, ...)
- Useful python package: **pdfminer.six**
- Useful OCR library: **Tesseract (pytesseract)**
- Cloud solutions



123456.pdf



123456.txt

GOAL

Transform pdf to plain text

3. Unstructured Information Extraction



Mandatory:

- Detect Articles (title + content)
- Find mentions of specific keywords;
 - Prime fin d'année / eindejaarspremie
 - Durée de validité / geldigheidsduur

TIPS

- Focus on **one language (NL/FR)**
- **Start with the article detection**, then work context specific for the others
- **Leverage Snorkel.ai for labelling / entity recognition**
- **Combine both basic tools (like regex) and more advanced (statistical model)**

GOAL

Find structure in the documents

4. Advanced Insights & Visualization

Insights examples:

- Filter articles based on title & contents
- Search for references to (specific) laws or CLA's
- Filter documents based on specific article contents
- Notification system

TIPS

- Focus on one language (NL/FR)
- Start with the article detection, then work context specific for the others
- Combine both basic tools (like regex) and more advanced (statistical model)
- Create manual article types or detect by clustering using keywords

User friendliness is valued

+ More can be defined during 'CLA workshop' today and 10/11

GOAL

Generate interesting insights in the dataset

Language Detection

For each document, detect the languages in the layout and extract all contents.

NL	
Paritair Comité voor het huiden- en lederbedrijf en vervangingsproducten	Commission paritaire de l'industrie des cuirs et peaux et des produits de remplacement
Collectieve arbeidsovereenkomst 6 januari 2021 betreffende de tijdelijke werkloosheid omwille van overmacht	Convention collective de travail du 6 janvier 2021 relative au chômage temporaire pour force majeure.
HOOFDSTUK I. Toepassingsgebied	CHAPITRE I. Champ d'application
Artikel 1. Deze collectieve arbeidsovereenkomst is van toepassing op de werkgevers en op de arbeiders van de ondernemingen die onder het Paritair Comité voor het huiden- en lederbedrijf en vervangingsproducten (PC 128) ressorteren.	Article 1 ^e . La présente convention collective de travail s'applique aux employeurs et aux travailleurs des entreprises ressortissant à la commission paritaire de l'industrie des cuirs et peaux et des produits de remplacement (CP 128).
Onder arbeiders worden zowel mannelijke als vrouwelijke arbeiders verstaan.	Par « ouvriers » on entend les ouvriers et les ouvrières.
HOOFDSTUK II. Bestaanszekerheid	CHAPITRE II. Sécurité d'existence
Artikel 2. Omdat op de uitbreiding van het toepassingsgebied van de tijdelijke werkloosheid wegens overmacht door de COVID-19, komen de sociale partners eveneens in het toepassingsgebied van de bepalingen inzake de bestaanszekerheid van de leersector, voor een bepaalde tijd uit te breiden naar de tijdelijke werkloosheid wegens overmacht, namelijk van 1 januari 2021 tot 31 maart 2021.	Article 2 ^e . En tant qu'il résulte de l'élargissement de l'application de l'accord sur la tijdelijke werkloosheid en raison de la lutte contre le COVID-19, les partenaires sociaux consentent d'élargir l'application des dispositions relatives à la sécurité d'existence prévues dans le secteur du cuir, à la situation de force majeure pour une période de durée déterminée courant du 1 ^e janvier 2021 au 31 mars 2021.
Artikel 3. De algemene beginselen, zoals bepaald in hoofdstuk I.I.1 (uitgezonderd artikel 3 §2 en artikel 5 § 4 tweede lid), hoofdstuk III en hoofdstuk IV van de collectieve arbeidsovereenkomst van 1 oktober 2019 aangaande de tewerkstelling en de tijdelijke werkloosheid (nr 154532/CO/128), voor de toepassing van deze aanvullende bestaanszekerheidsvoegingen blijven ongewijzigd.	Article 3. Les principes généraux, comme prévu en chapitre I.I.1 (sauf article 3 §2 et article 5 § 4 alinéa 2), chapitre III et chapitre IV de la convention collective de travail du 1 ^e octobre 2019 relative à l'emploi et au chômage temporaire (n° 154532/CO/128), pour l'application de ces dispositions relatives à la sécurité d'existence restent inchangés.
Artikel 4. De werkgever zal het totaal bedrag van de dagvergoeding betalen.	Article 4. L'employeur prendra en charge le montant total de l'indemnité journalière.
Het Fonds voor Bestaanszekerheid voor het huiden- en lederbedrijf en vervangingsproducten zal, in functie van de beschikbare financiële middelen, het volledig bedrag besteed aan de arbeiders terugbetaald aan de werkgever. Dit evenwel enkel indien de werkgever aan het Fonds kan aantonen dat zijn onderneming een erkenning tot uitzonderlijk hard getroffen onderneming in het kader van de COVID-19 crisis heeft.	Le Fonds de Sécurité d'Existence de l'industrie des cuirs et peaux et des produits de remplacement remboursera, en fonction des moyens financiers disponibles, la totalité du montant dépensé aux ouvriers à l'employeur. Toutefois, afin de bénéficier de ce remboursement, l'employeur doit prouver auprès du Fonds que son entreprise a été reconnue comme particulièrement touchée dans le cadre de

TIPS

- **Don't reinvent the wheel:** Make use of webservice API's Azure, GCP, AWS) or python packages (e.g. Spacy, spacy-langdetect, ...)
- **Use the meta-data** to identify the Joint Committee and **cluster similar templates**
- **Analyze the bi-lingual layout**

GOAL

Extract all contents of a document in 1 language

General



spaCy



TIPS & TRICKS

- Work with easy to manage 'database' systems (e.g. SQLite, Cloud providers) or even plain text to store your findings and meta-data.
- When training models, split test/validation set and measure specific performance metrics, ... (use meta-data as labels where available)
- Google search is your friend, but also don't hesitate to share your questions in the BeCode channel.



Planning



Expectations & Objectives

- Build **creative** solutions, think outside-the-box!
- **Collaborate** across teams where possible
- **Share** experiences and tips
- **Document** your approaches, where and why you failed / succeeded
- **Present** not only what you found, but **how** you did it, and **why**
- **Learn** to work with new tools

