

## Aplicação da mineração de processos e computação cognitiva na indústria: um estudo comentado

Vinicius Moura Pertel (Pontifícia Universidade Católica do Paraná - PPGEPS) [vinicius@dominus-eng.com](mailto:vinicius@dominus-eng.com)  
Fernando Deschamps (Pontifícia Universidade Católica do Paraná - PPGEPS) [fernando.deschamps@pucpr.br](mailto:fernando.deschamps@pucpr.br)  
Eduardo de Freitas Rocha Loures (Pontifícia Universidade Católica do Paraná - PPGEPS)  
[eduardo.loures@pucpr.br](mailto:eduardo.loures@pucpr.br)  
Eduardo Alves Portela Santos (Pontifícia Universidade Católica do Paraná – PPGEPS)  
[eduardo.portela@pucpr.br](mailto:eduardo.portela@pucpr.br)

### Resumo:

Com o mundo da indústria passando por uma nova revolução, chamada indústria 4.0, o uso de tecnologias e ferramentas mais avançadas e sofisticadas tem sido olhado de forma cada vez mais cautelosa, pois estas tecnologias serão os pontos-chaves para o sucesso desta nova etapa. O artigo aqui apresentado faz uma análise de um processo de produção industrial através do uso da ferramenta chamada *Process Mining*. Esta ferramenta permite extrair informações-chaves e auxiliar a fazer inferências sobre o processo atual analisado. Junto a esta análise, também está abordado o tema da utilização de softwares capazes de fazer o uso da computação cognitiva para avaliar os resultados dos chamados “log” utilizados na mineração de processos. O uso de softwares detentores de recursos avançados, como o processamento de forma inteligente, nos permite extrair informações de valores dos dados analisados, como por exemplo algumas sugestões ou até ações automáticas que os equipamentos podem tomar para atingir determinado objetivo. Estes softwares inteligentes junto a outras ferramentas de suporte, como o *Process Mining*, serão de imensa importância para a concretização da manufatura avançada.

**Palavras-chave:** Mineração de Processos, Computação cognitiva, Indústria 4.0, Manufatura Avançada.

## Application of process mining and cognitive computing in industry: a commented study

### Abstract

With the industry's world facing a new revolution, called industry 4.0, the use of tools and technologies more advanced and sophisticated has been looked more cautiously, because these artifices will be the key for the success of this new step. The article presented here makes an analysis of an industrial production process through the use of the tool called Process Mining. This tool allows you to extract key information and assist in making inferences about the current process being analyzed. In addition to this analysis, we also discuss the use of software capable of using cognitive computing to evaluate the results of the so-called logs used in process mining. The use of software with advanced features, such as intelligent processing, allows us to extract information from values of analyzed data, such as some suggestions or even automatic actions that the equipment can take to achieve a certain objective. These intelligent software along with other support tools, such as Process Mining, will be of immense importance for the achievement of advanced manufacturing.

**Key-words:** Process Mining, Cognitive Computation, Industry 4.0, Advanced Manufacturing.

### 1. Introdução

Ao longo dos anos, o uso de tecnologias nos negócios em geral tem sido intensificado e não somente em setores high-tech, mas nos mais variados nichos existentes. A principal motivação

que estimula os usuários a estarem sempre atrás de novas tecnologias, é o fato delas serem facilitadores de atividades de rotina que geralmente toma tempo e está sujeita a erros humanos. O mundo industrial está presenciando o acontecimento da quarta revolução, originada na Alemanha e disseminada aos demais países. Segundo a Capgemini, empresa mundial de serviços de consultoria e tecnologias, a chamada indústria 4.0 tem como objetivo trazer a digitalização ao meio industrial, podendo ser apontada como uma revolução digital também. Junto com toda esta digitalização, está incluso o uso das ferramentas e tecnologias inteligentes. Com a i4.0 os negócios atingirão níveis de eficiência, variedade e qualidade jamais vistos.

De acordo com a Confederação Nacional da Indústria (CNI), na publicação Sondagem Especial Indústria 4.0, apesar de toda esta preocupação com a melhora dos negócios através da implantação de novas tecnologias, existe também uma brecha para uma avaliação interna do processo já existente nas empresas. A indústria 4.0 tem um desafio de aprimorar os processos atuais através de parametrizações de seus equipamentos, gestão dos recursos e trocas de informações entre diversos níveis da empresa. Existem diversas maneiras de fazer uma avaliação dos processos atuais de um negócio, sendo que uma delas é utilizar uma ferramenta de mineração de processos. Com este recurso e através de um log de dados da organização, é possível fazer uma engenharia reversa para observar as atividades presentes do processo deste negócio.

Considerando este cenário de indústria 4.0, não é difícil encontrar menções ao uso do data mining por diversas gigantes do mercado para a extração de valores aos seus negócios, porém existe também o *Process Mining* como uma ferramenta extremamente poderosa para permitir se fazer inferências sobre um processo atual. Esta técnica por si só é capaz de retornar diversas informações importantes, mas em conjunto com outras pode se tornar ainda mais eficiente e relevante.

A quarta revolução aborda a ideia de tomadas de decisões independentes, e isto é possível através da utilização de softwares com grande inteligência computacional ou de computação cognitiva, o qual será combinado com o *Process Mining*, a fim de torna-los um conjunto de recursos muito importante quando se trata de melhoria em cima dos processos atuais de uma organização. Com o poder de análise das ferramentas de *Process Mining*, se tem o retorno de pontos a serem melhorados, e com as ferramentas inteligentes, estas melhoras podem ser feitas de forma autônoma, como por exemplo a mudança da rotina de produção devido a um gargalo identificado. Um dos desafios propostos é a otimização dos processos para tornar as empresas cada vez mais eficientes, diminuindo gastos desnecessários, desperdício, tempos ociosos, etc. Para isso, este artigo corre atrás de uma aplicação mais prática para entender de que forma é possível superar este desafio proposto, através da mineração de processos e utilização de softwares mais inteligentes. Os dados aqui utilizados foram cedidos pela empresa Tampaflex, uma indústria nacional de médio porte que atua no ramo de injeção plástica e possui uma vasta gama de produtos.

## 2. Método

Como objetivo geral deste artigo se tem a oportunidade de mesclar duas ferramentas para uma maior potencialização dos resultados obtidos. São estas ferramentas o *Process Mining* e o uso de softwares inteligentes de computação cognitiva. O objetivo é apresentar de uma forma mais prática o processo para trabalhar com a mineração de processos, e junto avaliar quais as informações desta ferramenta são interessantes para utilizar como *input* nos softwares inteligentes e quais seriam seus possíveis *outputs*. Para compor a metodologia deste artigo estão previstos os seguintes passos para serem seguidos a fim de atingir esse objetivo.

- Extrair informações do processo a ser analisado em formato Excel.

- Tratar estas informações para deixá-las na correta formatação para que o software de *process mining* consiga ler e interpreta-las, e assim fazer as devidas análises.
- Inserir o log tratado e aplicar uma série de filtros para extrair as informações relevantes sobre o processo.
- Pesquisar na literatura sobre softwares inteligentes que serão capazes de atender as necessidades da indústria 4.0.
- Analisar quais informações do programa de mineração de processos é interessante inserir no software inteligente, e de que forma estes dados poderiam agregar ao processamento para trazer uma saída de maior valor para a empresa.
- Analisar todos os resultados obtidos versus os objetivos propostos.
- Projetar as prospecções futuras para o trabalho.

### 3. Process Mining

#### 3.1 Extração do Log e Tratamento

Os dados cedidos pela Tampaflex possuem a única função de auxiliar na pesquisa e desenvolvimento deste estudo. Para a seleção dos dados foram escolhidas cinco ordens de produção (OP) e retirados os apontamentos feitos em cada. Para a extração dos dados brutos foi utilizado a ferramenta de *Select* do MySQL, com a parametrização para extração de cada informação pertinente a ser analisada, que neste caso foram ordem de produção, produto, descrição do produto, revisão do produto, data e hora do registro, usuário, máquina, descrição da máquina, tipo do apontamento feito e quantidade apontada. Esta extração gerou um arquivo .csv que precisou ser tratado ainda para a inserção no software Disco da Fluxicon, responsável por converter o .csv em um arquivo log de extensão compatível com o ProM, este por sua vez responsável por fazer as análises do processo.

ODF	PRODUTO	DESCRIÇÃO	REVISÃO	DATA HORA	USUÁRIO	MÁQUINA	DESCRIÇÃO	TIPO APONTAMENTO	QUANTIDADE
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 17:47:34.393	DJANYFER	6	Injetora 06 Dongshin 180	Setup Iniciado	
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 17:47:42.463	DJANYFER	6	Injetora 06 Dongshin 180	Setup Finalizado	
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 17:47:59.890	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 18:29:43.707	ALTIERES	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 19:29:06.637	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 20:18:10.060	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 21:07:47.180	DJANYFER	6	Injetora 06 Dongshin 180	Troca de matéria-prima Iniciado	
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 21:34:50.383	ELIFELITES	6	Injetora 06 Dongshin 180	Troca de matéria-prima Finalizado	
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 22:12:21.450	ELIFELITES	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 22:37:38.103	RICARDO	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 23:16:23.707	ELIFELITES	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-04 23:58:46.770	RICARDO	6	Injetora 06 Dongshin 180	Peças Boas	3,303
66730	651R2438000	PE - Tampa 024 mm Retrátíl - Natural (molde 1)	11	2017-09-05 00:38:46.080	RICARDO	6	Injetora 06 Dongshin 180	Peças Boas	3,303

Figura 1 - Dados apontamento OP

A figura acima apresenta um pedaço dos dados brutos retirados já com uma formatação para que as informações que antes eram separadas por vírgula, fossem apresentadas separadas por colunas. Também foi unificado em uma única tabela todas as informações das OPs, sendo feita uma disposição dos eventos de acordo com a data e hora da ocorrência.

Este arquivo foi inserido no Disco, onde foi necessário acertar alguns parâmetros básicos para converter os dados em um fluxo de processos. Os parâmetros básicos são case ID, atividade, *timestamp* e recurso, mas caso queira levar mais alguma informação para fazer análises é possível indicando estas informações como “outros”. Estes parâmetros são divididos pelas colunas dos seus dados, onde cada coluna deve se encaixar em algum deles. Se o arquivo de dados possuir colunas que não são interessantes então deve-se remover estas colunas no próprio software da Fluxicon.

Neste caso, as colunas selecionadas para entrar na avaliação do software foram ODF, Data Hora, Usuário, Tipo Apontamento e Quantidade distribuindo os parâmetros case ID, *timestamp*, recurso, atividade e “outros”, respectivamente, como mostrado na figura abaixo. As demais colunas contendo alguma informação foram desconsideradas, preliminarmente, para esta análise de processo.

Com os parâmetros selecionados, o software está apto a fazer a importação dos dados. O Disco oferece alguns filtros para *TimeFrame*, *Performance*, *Endpoints*, *Attribute* e *Follower*, além de dados estatísticos sobre o processo. A figura 3 é a página gerada após a importação dos dados, já apresentando uma representação de fluxograma do processo a ser analisado.

ODF	PRODUTO	DESCRIÇÃO	REVISÃO	DATA HORA	USUÁRIO	MÁQUINA	DESCRIÇÃO	TIPO APONTAMENTO	QUANTIDADE
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 17:47:34.393	DJANYFER	6	Injetora 06 Dongshin 180	Setup Iniciado	QUANTIDADE
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 17:47:42.463	DJANYFER	6	Injetora 06 Dongshin 180	Setup Finalizado	QUANTIDADE
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 17:47:59.890	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 18:29:43.707	ALTIERES	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 19:29:06.637	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 20:18:10.060	DJANYFER	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 21:07:47.180	DJANYFER	6	Injetora 06 Dongshin 180	Troca de matéria-prima Inicial	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 21:34:50.383	ELIFELITES	6	Injetora 06 Dongshin 180	Troca de matéria-prima Final	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 22:12:21.450	ELIFELITES	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 22:37:38.103	RICARDO	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 23:16:23.797	ELIFELITES	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-04 23:58:46.770	RICARDO	6	Injetora 06 Dongshin 180	Peças Boas	3.303
66730	651R2438000	PE - Tampa 024 mm Retrátil - Natural (molde 1)	11	2017-09-05 00:38:46.080	RICARDO	6	Injetora 06 Dongshin 180	Peças Segregadas	3.023

Figura 2 - Seleção dos parâmetros básicos

O processo analisado consiste em uma sequência completa de atividades que ocorreram para produzir determinados produtos plásticos, através do processo de injeção em moldes com um operador disposto, no mínimo, para cada injetora. A representação do processo gerado consiste em setup inicial e final de máquina, para prepara-la para produção. Apontamento de peças boas, onde cada apontamento representa o registro da produção de uma caixa de peças boas e a quantidade destas peças por caixa varia de acordo com o produto a ser produzido. Entre o início da OP e sua finalização existem diversas ocorrências, onde algumas já são esperadas devido a pré programação. As ocorrências esperadas neste log analisado seriam a parada para que o operador faça um lanche, manutenção preventiva nas injetoras, *Try out* da engenharia e análise de laboratório. As demais ocorrências são, em sua maioria, prejudiciais ao desempenho do processo e serão alvos de análises para possíveis melhorias.

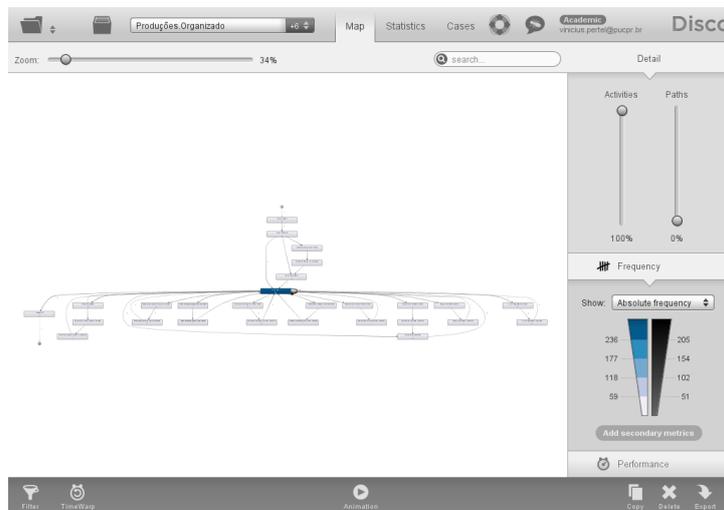


Figura 3 - Tela após importação dos dados

Após esta importação é possível exportar o arquivo gerado para o formato .MXML, com o qual se carrega no ProM para análises mais profundas. Carregando este arquivo no ProM tem-se algumas informações iniciais sobre este log que são geradas automaticamente, figura 4, em quatro possíveis abas a esquerda da janela. Estas informações são número de processos, número de instâncias de processo e quais são, quantidade de entidades, quantidade de atividades e quais são elas, apontando as atividades de início e fim, e por último quem são os recursos utilizados neste processo. Esta janela serve também como uma ferramenta de auxílio para aplicação de filtros no log inserido antes que sejam feitas as análises. Com isto é possível retirar amostras específicas para melhorar a avaliação como um todo. Uma maior exploração do log será apresentada no capítulo 3.2.

### 3.2 Avaliação Log

Aplicando o plugin chamado Fuzzy Miner neste log, figura 5, podemos observar como o processo de apontamento das OPs tem se comportado. O processo está com clusters para ajudar na representação, devido à grande quantidade de atividades e complexidade para representar todas as sequencias ocorridas neste log.



Figura 4 - Tela Básica de Informações

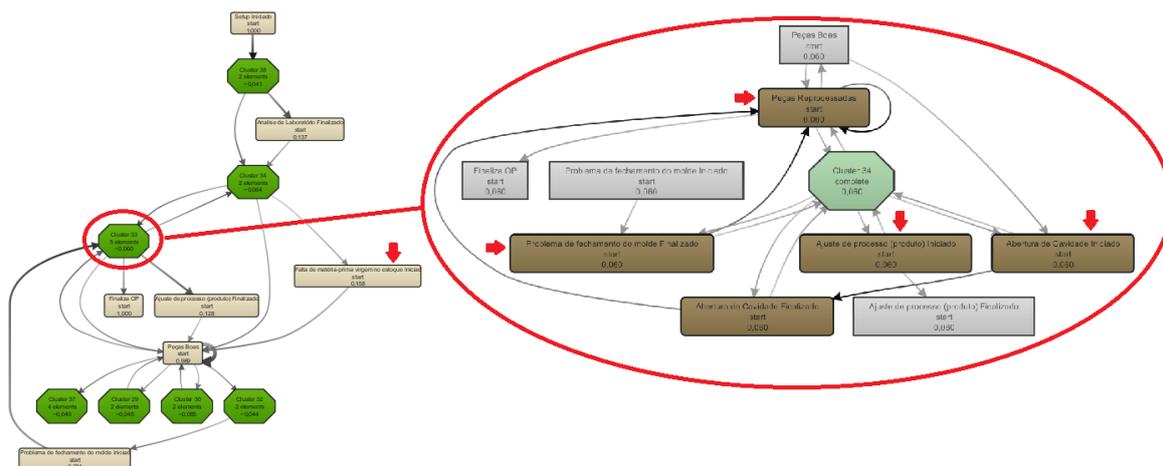


Figura 5 - Plugin Fuzzy Miner com Clusters

Olhando o processo gerado pelo plugin Fuzzy Miner, com ressaltado sobre o cluster 33 expandido ao lado, é possível notar a presença de algumas das diversas atividades, consideradas indesejáveis e indicadas pelas setas vermelhas, que ocorreram durante a produção destas OPs. Uma atividade é considerada prejudicial se ela ocorre de maneira não prevista e prejudica o OEE da produção. Estas atividades prejudiciais ao desempenho são mostradas na Tabela 1.

As atividades diferenciadas apenas por “iniciado” e “finalizado” foram unificadas em uma única linha tendo seus dados de tempo retirados através da diferença entre os dois tempos, e sua representatividade somada. De um total de 330 horas trabalhadas, 43 horas foram destinadas as atividades consideradas como prejudiciais ao processo, isto representa em torno de 13% de “perda de produtividade”. Porém, nem todas estas ocorrências representam 100% de perdas, como por exemplo as peças segregadas, peças reprocessadas e peças refugadas.

As peças que foram segregadas são levadas a chamada área amarela para uma vistoria mais afundo a fim de encontrar o defeito ou motivo pelo qual foram segregadas. Caso seja encontrado

o defeito que estava como hipótese, a produção destas caixas se tornam peças para serem encaminhadas ao reprocessamento. As peças que foram refugadas não passam pela área amarela e são encaminhadas diretamente para serem reprocessadas.

Atividade	Frequência	Tempo (Hrs)	Representatividade
Peças Reprocessadas	17	7,7	4,6%
Peças Refugadas	12	8,9	3,2%
Peças Segregadas	5	3,8	1,4%
Vazamento de água no Molde	2	7,4	1,1%
Falta de matéria-prima virgem no Estoque	1	12,5	0,54%
Ajuste de processo (Produto)	1	0,8	0,54%
Abertura de Cavidade	1	0,5	0,54%
Problema de fechamento do Molde	1	0,9	0,54%
Manutenção Corretiva Molde	1	0,2	0,54%
Limpeza de Molde	1	0,3	0,54%
TOTAL	42	43	13,54%

Fonte: Análises a partir do ProM

Tabela 1 - Atividades Indesejadas

As peças selecionadas para serem reprocessadas são moídas e embaladas para serem usadas como matéria prima para a produção de novas peças. Desta forma a perda que ocorre nestes processos são de tempo, custos com produção, custos para reprocessar o material e uso de colaboradores que poderiam estar dedicando o tempo com produção efetiva de produtos. Porém, neste processo o material utilizado para a injeção (material virgem) é reaproveitado para ser injetado novamente, mas sem os defeitos que ocorreram preliminarmente.

Outro ponto interessante a ser notado é referente a relação entre o tempo que cada atividade consumiu do processo versus sua representatividade em relação ao número de ocorrências. Esta observação demonstra que nem todas as atividades com maior representação são as que mais tomaram tempo do processo, indicando assim aquelas que podem ser consideradas como as mais críticas e que merecem uma atenção maior.

Por meio do plugin de análise chamado *Performance Sequence Diagram* é possível extrair diversas informações sobre a performance dos colaboradores, equipamentos, atividades, etc. Para esta avaliação é muito interessante utilizar este plugin para acompanhar, por exemplo, as horas de trabalho de cada injetora, figura 6, podendo esta informação auxiliar na conciliação entre produção e manutenção preventiva. Desta forma o Planejamento e Controle de Produção (PCP) possui uma visão mais clara sobre quando terão de abrir mão daquela injetora para que os responsáveis da manutenção executem seus serviços, e com isso evitar paradas forçadas no meio de uma produção ou ainda evitar que as máquinas operem em situações não ideais, podendo acarretar em quebras e danos maiores.

Para uma análise em que se deseja avaliar quais foram as OPs que mais se distanciaram do processo “padrão” de produção, pode-se utilizar o plugin *Trace Clustering*. Este plugin possui a função de clusterizar os processos, reconhecendo aqueles que possuem maior semelhança e separando-os em clusters. No caso deste log, o processo de clusterização retornou dois clusters, onde um possui a representação de quatro OPs e o outro apenas de uma. Desta forma é possível retirar do cluster com maior número de ordens de produção um processo muito mais limpo e que representa a maioria das atividades avaliadas como indica a figura 7.

Estes são apenas alguns dos possíveis plugins a serem utilizados para extrair informações e fazer análises sobre um determinado log, variando de caso a caso dependendo da informação

que se necessita analisar. No capítulo 4 será abordado o tema de softwares inteligentes e como funcionam.

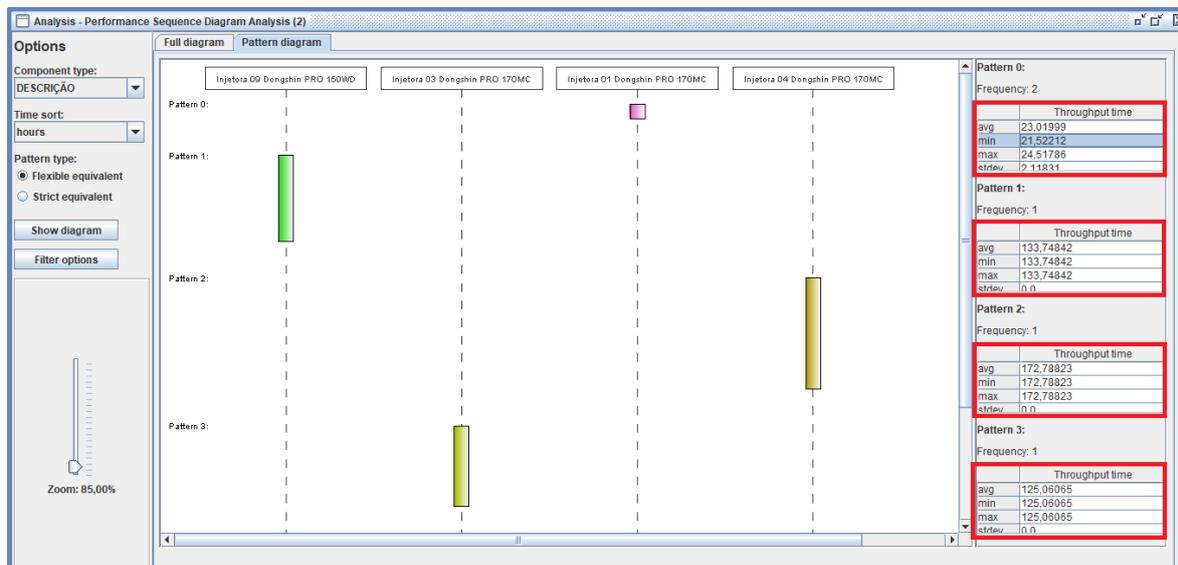


Figura 6 - Plugin Performance Sequence Diagram

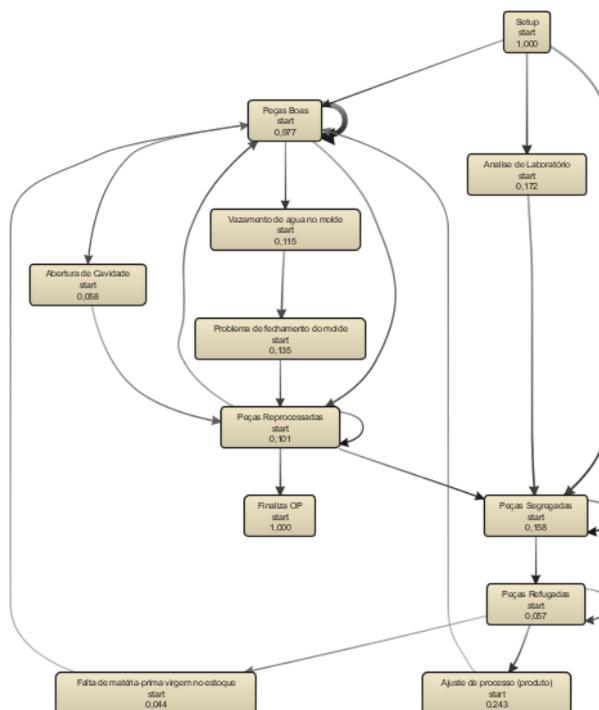


Figura 7 - Fuzzy Miner após Clusterização

#### 4. Softwares Inteligentes

O uso de softwares inteligentes não surgiu recentemente. Segundo pesquisa realizada nas bases de dados da Universidade de Oxford, Alan M. Turing publicou um artigo sobre a inteligência da computação em 1956 em (TURING, A. M., 1956). Desde então o tema tem sido alvo de discussões e objeto de estudo.

Entre as diversas companhias do setor, uma grande referência hoje na área da computação cognitiva é a empresa *International Business Machines*, também conhecida como IBM. Um

marco para a sua divulgação ao público foi o uso do seu software, chamado Watson, para competir em um programa de perguntas, conhecido como Jeopardy, em 2011 contra dois competidores conhecidos que já haviam ganho o desafio do programa em outras oportunidades.

Outra grande empresa que se destaca tanto no desenvolvimento de novas tecnologias, quanto no investimento em pesquisas é o Google. Uma parte dos seus desenvolvimentos de inteligência artificial foram focadas nos processamentos de vídeos, traduções e ferramentas de pesquisa. Segundo o Google, este mecanismo de pesquisa, denominado *RankBrain*, tem sido um forte aliado na otimização das buscas feitas em seu site, se tornando o terceiro sinal mais importante para conformar o resultado das buscas.

Através de pesquisas sobre estas gigantes e suas ferramentas disponíveis ou ainda em processo de desenvolvimento, apurou-se um recurso desenvolvido pela IBM chamado de *Business Process Insight* (BPI), como um protótipo. O BPI é um software que contém uma ferramenta de inteligência sobre processos e possui diversas técnicas oriundas das disciplinas de negócios, monitoramento das atividades do negócio, processamento de eventos complexos, BPM e mineração de processos.

Os pesquisadores Rozsnyai S., Lakshmanan G., Muthusamy V. no artigo publicado na Conferência global SRII em 2012, chamado de *Business Process Insight: An Approach and Platform for the Discovery and Analysis of End-to-End Business Processes*, falam mais sobre o BPI. Segundo os pesquisadores do BPI, o gerenciamento dos processos de negócios atuais possui estruturas extremamente rígidas e pensadas para atender 100% de um negócio específico, se tornando inútil para outros ramos de negócios, isso sem considerar as diversas empresas que operam sem nenhum controle dos seus processos. Desta forma eles propõem a utilização do BPI para gerar insights sobre processos semiestruturados. O objetivo é permitir o gerenciamento de processos semiestruturados com uma melhor visibilidade do seu comportamento, desta forma irá melhorar o gerenciamento e terá como resultado o aumento da inteligência do processo com análises estatísticas.

Uma representação de todo o funcionamento do BPI está apresentada na figura 8, que foi extraída diretamente do material apresentado pela IBM e está melhor explorado em seguida.

Para entender o funcionamento do BPI, os pesquisadores da IBM afirmam que primeiro a ferramenta descobre o comportamento do processo e subsequentemente fornece conhecimentos sobre o que foi executado nesta sintetização de processo, do começo ao fim, através de um histórico e uma *stream* de eventos observados. Cada instância do processo está relacionada com um subconjunto de eventos que formam um caminho executado do começo ao fim para aquela instância. O gerenciamento do tempo de execução melhora através de uma análise sobre a causa raiz devido a exploração colaborativa do comportamento do processo que foi descoberto, além de um monitoramento contínuo e alertas inseridos no processo.

A origem destes dados vem de diversos formatos e fontes, sendo dados do sistema ERP, e-mails, documentos utilizados na empresa, etc. Para que essas informações cheguem de forma legível ao software, é necessário concretizar e configurar uma integração dos dados em prol de ter uma representação estável que pode ser utilizada em camadas de aplicações mais elevadas.

Com a integração dos dados estruturada, o BPI utiliza-se da descoberta de correlação para detectar as diversas dimensões dos relacionamentos entre os eventos. Como resultado, o software gera regras que expressam como os eventos se relacionam e eles podem ser instâncias isoladas ou dimensões. Através de uma ferramenta de correlação, utilizando as relações descobertas anteriormente, é possível criar grupos relacionados ou gráficos de relacionamentos. Se tratando de processos, as dimensões de relacionamentos representam o comportamento de cada instância do processo através de uma representação gráfica. Com esta representação

gráfica elaborada, o software passa a aprender diferentes modelos através de *process mining*, para explorar o comportamento e treinar para poder fazer previsões no modo de acompanhamento ao vivo.

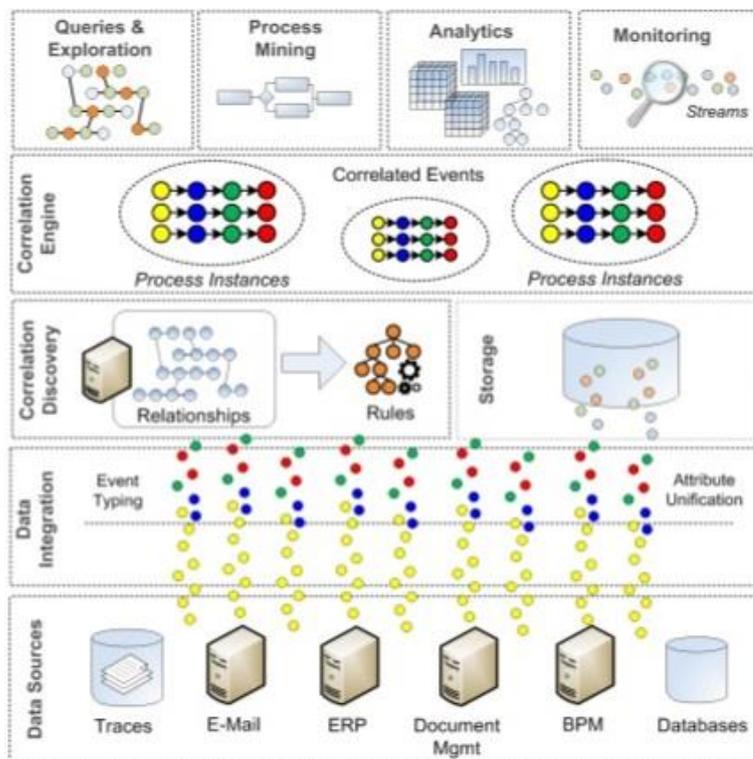


Figura 8 - Funcionamento do BPI

## 5. Softwares Inteligentes x Mineração de Processos

Considerando o cenário atual, existem dois grandes momentos distintos. O primeiro é a área de mineração de processos, que possui uma grande solidez no mercado, mesmo que ainda utilizada apenas por grandes empresas ou em casos específicos. Esta ferramenta já se encontra com uma vasta base de conhecimento que está disseminada. Além disto, existem diversos softwares disponíveis para elaborar a mineração de processos. O segundo cenário é sobre os softwares de computação cognitiva, que ainda se encontram em processo de desenvolvimento e com menor disseminação de conhecimento, ainda mais se for considerar a quantidade de profissionais capacitados nesta área. Desta forma a discussão começa com um desafio de como unir todo o potencial ainda não desenvolvido da computação cognitiva com toda a carga de conhecimento da mineração de processos.

Um ponto inicial foi dado com o protótipo BPI da IBM, mas este software ainda possui diversas barreiras a serem superadas. Por se tratar de protótipo, o acesso ao BPI fica restrito aos desenvolvedores, restando apenas uma compilação sobre os materiais divulgados.

Para entender o que existe e aonde se quer chegar foi elaborado a tabela 2 apresentada logo abaixo. Esta tabela apresenta alguns pontos importantes da mineração de processos e computação cognitiva e posteriormente será tratado como eles, juntos, se tornarem uma ferramenta mais poderosa para análises e gerenciamento de processos.

Os pontos considerados como fracos ou limitadores de cada área foram destacados em vermelho para entender como é possível utilizar uma área para complementar a outra, na questão de análises de processos.

Analisando a Tabela 2, permite-se dizer que a mineração de processo acabará sendo uma das ferramentas utilizadas pela computação cognitiva para evoluir ainda mais sua capacidade de análises. Os pontos fracos da mineração de processos como a limpeza manual dos dados brutos extraídos, ou ainda a necessidade de um conhecimento prévio para extrair informações de valores será compensada pela aplicação automática de filtros para retirar os ruídos dos dados obtidos e ainda com cada iteração do processo de carregar os dados e gerar novas análises, o software se tornará mais inteligente para permitir fazer inferências que não sejam triviais a qualquer pessoa. As limitações de uso de dados apenas estruturados e processos totalmente conhecidos será ultrapassada com o poder de tratamento de dados não estruturados e processos semiestruturados da computação cognitiva, sendo que os dados não estruturados serão todos aquelas informações geradas pela empresa, mas que não são registradas de maneira estruturadas, como por exemplo os e-mails. Como as informações de processo não serão mais fixas, o sistema passa a aprender a cada vez que faz uma análise nova e armazena este conhecimento para que sua estrutura aprenda novos recursos. De uma maneira mais clara a Figura 9 indica uma representação do que seriam as entradas feitas nos sistemas de computação cognitiva que utilizam a mineração de processo para retirar informações de valor, e quais seriam as possíveis saídas destas análises.

As entradas são informações estruturadas e não estruturadas para enriquecer o conhecimento do software e a complexidade das análises. Neste ponto é bem interessante notar a existência de uma entrada sendo os fatores externos. Estes fatores em um futuro próximo serão informações como preços de cotações de matéria prima, com referência do dólar atual, dados estatísticos do consumo histórico do produto a ser produzido e outras informações que o software julgar como relevante para extrair valor. Todos esses dados passam por um filtro para ser analisado apenas aquilo que interessa a empresa. Neste caso para a Tampaflex é interessante receber informações externas como as cotações de preços de matéria prima, consumo do mercado total para retirar a porcentagem de representação da empresa no total, empresas que estão consumindo este tipo de produto, feedback de clientes, etc. As entradas internas interessantes para analisar são as informações do sistema ERP (relatórios de desempenho, gastos, etc), BPM (relatório de apontamento, workflows, etc), trocas de e-mails para extrair informações sobre decisões a serem tomadas e o software estar ciente.

Com a análise do process mining existirão indicadores para auxiliar em uma tomada de decisão inteligente. Esta tomada de decisão será tomada automaticamente pelo sistema a partir do momento que ele obtenha inteligência o suficiente para entender corretamente os dados e fazer previsões de formas mais assertivas. Preliminarmente o software deve atuar como uma ferramenta de auxílio as tomadas de decisões, indicando através do process mining quais as ações precisam ter uma atenção maior em sequência de prioridades. Um exemplo de tomada de decisão inteligente e automática seria o software analisar a quantidade de matéria prima que deve ser comprado, considerando o consumo de mercado, preço do dólar, sazonalidade do produto e desta forma prever uma quantidade X de matéria prima para comprar que terá um melhor equilíbrio entre todos os fatores. Esta decisão coloca um pedido para o fornecedor de forma automática e recebe um retorno de estimativa da chegada para considerar nos cálculos da produção.

## 6. Conclusões

A área de mineração de processos possui diversos recursos interessantes para retornar informações de valor ao negócio, mas para chegar até lá é necessário passar por algumas etapas preliminares. O processo começa com a extração de um log real, o qual vem em uma forma crua sendo necessário a aplicação de alguns filtros para eliminar ruídos e tratar os dados para que fiquem em uma forma interessante de se analisar. Este processo serve para eliminar aquelas

linhas com dados faltando, mesclar ou criar novas colunas e linhas para gerar novas informações, deixar no formato que o software Disco consiga entender.

Mineração de Processos	Computação Cognitiva
Grande disseminação do conhecimento	<b>Conhecimento ainda pouco disseminado</b>
Utilizado por empresas grandes ou com sistemas bem completos	Pouco utilizado ainda por empresas, maior utilização em pesquisas
Diversos softwares de baixo ou sem nenhum custo para utilização	<b>Poucos softwares disponíveis a custos altos</b>
<b>Necessidade de grandes bancos de dados estruturados</b>	Utilização de bancos de dados estruturados e não estruturados
<b>Somente para processos de informações estruturadas</b>	Pode ser utilizado em processos com informações semiestruturadas
<b>Dados provenientes do sistema BPM e ERP</b>	Dados provenientes de diversas bases como BPM, ERP, e-mails, formulários, bases de dados globais, etc.
Moderado número de profissionais no mercado que dominam a área	<b>Poucos profissionais no mercado que dominam a área.</b>
<b>Formatação e limpeza dos dados manuais</b>	Aplicação de filtros para retirar os ruídos automaticamente
<b>Análises geradas a partir de conhecimentos prévios</b>	Análise sempre em evolução de acordo com cada caso novo assimilado
Bom espaço para crescimento no mercado	Grande espaço para crescimento no mercado

Tabela 2 - Mineração de Processos x Computação Cognitiva

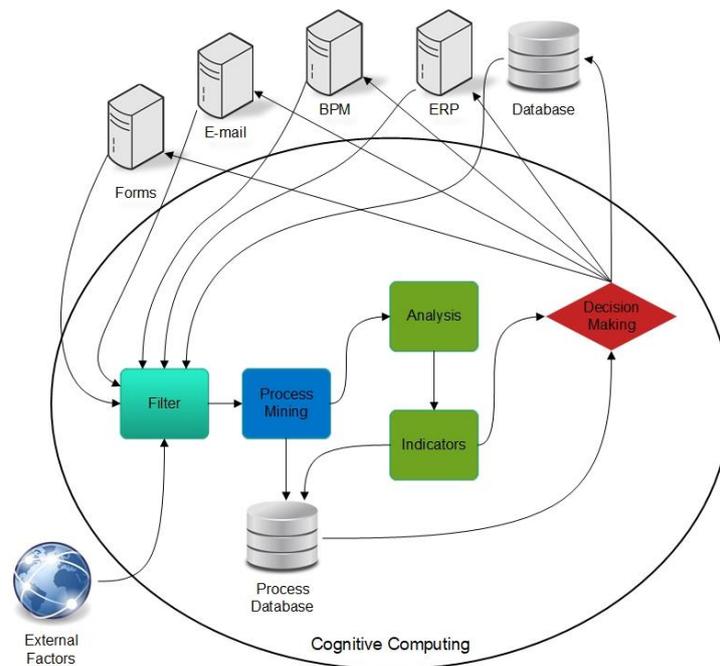


Figura 9 - Computação Cognitiva com Mineração de Processos

Por se tratar de um ambiente industrial, onde a exigência do controle sobre as diversas variáveis de processo é extremamente rígida, o uso de uma ferramenta de process mining é de fundamental importância para assegurar todo o conhecimento em cima do processo. Desta forma um impacto direto é percebido por toda a organização, seja em nível econômico, organizacional e também sobre o ambiente de trabalho, pois um ambiente controlado tende a ser muito mais fácil de se trabalhar do que em um ambiente suscetível as incertezas de processo. Os softwares de computação cognitiva ainda estão em estágios iniciais de seus ciclos de vida. Muitas empresas grandes no mercado estão investindo pesado na pesquisa deste segmento, como é o caso do Google e IBM. Existe uma grande oportunidade para o desenvolvimento de

novas ferramentas inteligentes que auxiliem nas atividades atuais. Se considerar a mescla de ferramentas como os softwares inteligentes e o process mining, o resultado é uma ferramenta muito poderosa e que traz muito mais valor do que a utilização de cada um de forma separada. Mesmo estando ainda pouco maduros estes softwares inteligentes, já existe pesquisas e estudos em cima da mescla de ferramentas, como é o caso do BPI da IBM. Este software é um protótipo ainda, mas se trata de uma tentativa de mesclar diversos conhecimentos com a computação cognitiva. Com esta mescla é possível utilizar uma quantidade muito maior de informações para alimentar as entradas de dados. Estes dados para um software de process mining são geralmente oriundos dos sistemas ERP e BPM, mas com a ferramenta inteligente é possível agregar novos dados em forma não estruturada, como dados oriundos de e-mails, formulários e até dados externos a organização para enriquecer cada vez mais o sistema. O resultado desta mescla é ter um software robusto e inteligente que avalia dados estruturados e não estruturados, de processos pelo menos semiestruturados, e retorna uma informação para auxiliar na tomada de decisão. Quando o software tiver aprendido bem com uma grande quantidade de dados analisada é possível permitir a tomada de decisão independente de interferência humana, assumindo assim que o programa é responsável por analisar todos os fatores internos e externos e ponderar qual é a melhor decisão a ser tomada. Nesta etapa, o software possui uma avaliação muito além daquilo que um colaborador conseguiria fazer de forma trivial, já que o volume de dados a ser analisado é extremamente grande.

Existem algumas limitações ainda em pontos chaves para o maior desenvolvimento do trabalho nesta área. Pode-se citar questões como a falta de um grande volume de profissionais capacitados para que seja estimulado cada vez mais o desenvolvimento deste conteúdo. Uma limitação em relação a mineração de processos foi o volume de dados analisados, o qual pode ser maior para que uma melhor visualização sobre o processo geral seja obtida. Existe a possibilidade de trabalho futuro com uma abordagem diferente também, onde ao invés de pegar diversas OPs sobre vários produtos, seja selecionado várias OPs, mas de apenas um produto para analisar caso a caso as possíveis otimizações.

## Referências

- BECHTOLD, J.; LAUENSTEIN, C; KERN A.; BERNHOFER L .** *Industry 4.0 – The Capgemini Consulting View*. Deutschland, p.01-36, 2014.
- AFONSO, I.; FONSECA, R.; CUNHA, S.** *Sondagem especial – Indústria 4.0*. Brasil: CNI, 2016.
- GOMES, J.; AFONSO, I.** Brasil pode criar a indústria 4.0 verde e amarela. Brasil, CNI, 2016.
- TURING, A. M.;** *Computing Machinery and Intelligence*. England: Oxford University Press, 1958.
- ROZSNYAI S., LAKSHMANAN G., MUTHUSAMY V.** *Business Process Insight: An Approach and Platform for the Discovery and Analysis of End-to-End Business Processes*. SRII, 2012.
- GUDIVADA, V.** *Cognitive Computing: Concepts, Architectures, Systems and Applications*. , p.3-38, 2016.
- FOUAD, M. M.; OWEIS, N. E.; GABER, T.; AHMED, M.; SNASEL, V.** *Data Mining and Fusion Techniques for WSNs as a Source of the Big Data*. Procedia Computer Science 65, p. 778-786, 2015.
- AALST, W. M. P.; REIJERS, H. A.; WEIJTERS, A. J. M. M.; DONGEN, B. F.; MEDEIROS, A. K.; SONG, M.; VERBEEK, H. M. W.;** *Business Process Mining: An industrial application*. Information Systems 32, p. 713-732, 2007.
- AALST, W.;** *Process Mining Manifesto*. BPM Workshops, part I, p. 169-194, 2012.
- VARGA, S.; ROTTA, T. C.;** *Afinal, o que é computação cognitiva*. IBM Academy of Technology Affiliated, Mini paper nº270, 2011.