

MESTRADO
GESTÃO DE SISTEMAS DE INFORMAÇÃO

TRABALHO FINAL DE MESTRADO
PROJETO

MAPEAMENTO DAS ATIVIDADES QUE CONTRIBUEM PARA OS ODS
COM O APOIO DO *BUSINESS INTELLIGENCE* – APLICAÇÃO PRÁTICA
NUMA INSTITUIÇÃO DE ENSINO SUPERIOR

ANA CATARINA LOURENÇO MARCELINO

NOVEMBRO - 2020

MESTRADO
GESTÃO DE SISTEMAS DE INFORMAÇÃO

TRABALHO FINAL DE MESTRADO
PROJETO

MAPEAMENTO DAS ATIVIDADES QUE CONTRIBUEM PARA OS ODS
COM O APOIO DO *BUSINESS INTELLIGENCE* – APLICAÇÃO PRÁTICA
NUMA INSTITUIÇÃO DE ENSINO SUPERIOR

ANA CATARINA LOURENÇO MARCELINO

ORIENTAÇÃO:

PROFESSORA DOUTORA WINNIE NG PICOTO

NOVEMBRO - 2020

AGRADECIMENTOS

A concretização deste trabalho final de mestrado sem a presença de um conjunto de pessoas não seria possível, e a elas expresseo o meu agradecimento. Quero agradecer a todos os que contribuíram, de forma direta e indiretamente para a conclusão do meu mestrado, sem deixar de mencionar a minha família, os meus amigos e colegas de trabalho pelo apoio prestado.

Inicialmente, expresseo o meu agradecimento à professora Winnie Ng Picoto, que me orientou, aconselhou e transmitiu conhecimentos base para a execução do projeto, e que sempre se mostrou disponível para me ajudar.

Um enorme obrigada aos meus colegas de trabalho e amigos, por todo o apoio e troca de ideias no desenvolvimento prático do projeto e que me acompanharam durante a realização do mesmo.

Por fim, quero agradecer à minha família pelas experiências que me proporcionaram para alcançar os meus objetivos e pelo apoio incondicional durante todo o meu percurso académico, e aos meus amigos pelos momentos de distração e força que me deram para acreditar no meu potencial e finalizar o projeto.

RESUMO

A Organização das Nações Unidas identificou um conjunto de 17 Objetivos de Desenvolvimento Sustentável (ODS) na Agenda 2030, como um esforço para abordar os desafios sociais, económicos e ambientais que o nosso mundo enfrenta. Os ODS são uma visão para alcançar um futuro sustentável e as Instituições de Ensino Superior (IES), sendo órgãos privilegiados na disseminação de conhecimento, são agentes cruciais que devem ter um papel ativo na promoção e na divulgação dos ODS.

As IES podem contribuir para os ODS através de *know-how* e de boas práticas para apoiar a implementação e integração da sustentabilidade nas suas atividades de investigação, educação, operações no campus e de interação com a sociedade. Muitas IES dos países signatários da Agenda 2030, ainda não possuem uma visão geral da forma como as suas atividades abordam os ODS. No entanto, é importante a identificação e a monitorização das atividades que divulgam a sustentabilidade para medir o progresso e as contribuições realizadas com o objetivo de contribuir para atingir os ODS.

O Instituto Superior de Economia e Gestão (ISEG), entidade para o qual foi concebido o projeto, pretende reforçar o seu posicionamento enquanto universidade que contribui para os ODS. Deste modo, o projeto tem como principal objetivo classificar e mapear as atividades académicas realizadas pelo ISEG de acordo com os ODS, na vertente de investigação e de ensino, e no desenvolvimento de uma solução de *Business Intelligence* onde é possível a visualização e a análise de informação para suporte à tomada de decisão. Esta solução permite a centralização de informação proveniente de diferentes fontes de dados e a transformação das mesmas em conhecimentos úteis para a instituição, estando alinhada com o objetivo de alcançar os ODS. Os resultados revelam que o projeto concebido, de uma forma geral, responde às necessidades do ISEG, possibilitando o suporte à tomada de decisão, a comunicação e a identificação quantitativa de contribuições.

Palavras-Chave: Objetivos de Desenvolvimento Sustentável; *Business Intelligence*; Mapeamento; *Dashboards*; Instituições do Ensino Superior; *Keywords*.

ABSTRACT

The United Nations has identified a set of 17 Sustainable Development Goals (SDGs) in the 2030 Agenda in an effort to address the social, economic and environmental challenges facing the world. The SDGs are guidelines for achieving a sustainable future. As privileged agents in the dissemination of knowledge, the Higher Education Institutions (HEIs) play a crucial and active role in the promotion and communication of the SDGs.

Through their know-how and good practices, HEIs can be useful to the SDGs through the implementation and integration of sustainability in their research, education, campus operations, and interaction with society. Many HEIs from the signatory countries of the 2030 Agenda do not yet have an overview of how their activities address the SDGs. However, it is important to identify and monitor the activities that promote sustainability in order to measure progress and contributions made in order to achieve the SDGs.

ISEG, the entity for which this project was planned, seeks to strengthen its position as a university that contributes to the SDGs. As such, the main objective of this project is to classify and map the academic activities carried out by ISEG that follow the SDGs, in terms of research and teaching, as well as develop a Business Intelligence solution that will make it possible to visualize and analyze data to support decision-making. This solution allows for a centralization and transformation of data from different data sources into useful knowledge for the institution, aligned with the objective of reaching the SDGs. The results reveal that, in general, the project conceived responds to the needs of ISEG, enabling support to decision-making, communication, and to the quantitative identification of contributions.

Keywords: Sustainable Development Goals; Business Intelligence; Mapping; Dashboards; Higher Education Institutions; Keywords.

ÍNDICE

Índice de Figuras	vi
Índice de Tabelas	vii
Lista de Abreviaturas.....	vii
1. Introdução.....	1
1.1. Enquadramento e Motivação	1
1.2. Objetivos do Projeto	4
1.3. Relevância dos Objetivos	5
2. Revisão da Literatura	6
2.1. Objetivos de Desenvolvimento Sustentável	6
2.2. ODS e as Instituições de Ensino Superior	8
2.2.1. Integração dos ODS.....	9
2.2.1.1. Mapeamento das atividades das IES	11
2.3. Business Intelligence	12
2.3.1. Arquitetura de BI.....	12
2.3.2. Dashboards	15
2.4. Trabalhos relacionados em IES	16
3. Metodologia.....	17
3.1. Enquadramento do Projeto	18
3.2. Identificação das Fontes de Dados	18
3.3. Conjunto de Keywords	19
3.4. Ambiente de Desenvolvimento	21
3.5. Tratamento dos Dados	23
3.6. Algoritmo de Classificação	24
3.7. Modelo Dimensional	25

3.8. Visualização dos Dados.....	27
4. Conclusões.....	32
5. Limitações e Recomendações para Trabalhos Futuros.....	33
6. Referências Bibliográficas.....	35
7. Anexos.....	40

ÍNDICE DE FIGURAS

FIGURA 1 - Objetivos de Desenvolvimento Sustentável.....	7
FIGURA 2 - Quatro pilares das contribuições das IES.....	8
FIGURA 3 - Passos para a integração dos ODS	10
FIGURA 4 - Arquitetura de <i>Business Intelligence</i>	13
FIGURA 5 - Fluxograma do projeto.....	17
FIGURA 6 - Comparação do tempo de processamento entre Regex e o FlashText	22
FIGURA 7 - <i>Magic Quadrant</i>	23
FIGURA 8 - Modelo dimensional de dados	26
FIGURA 9 - <i>Dashboard: SDG Summary</i>	28
FIGURA 10 - <i>Dashboard: SDG Contributions</i>	28
FIGURA 11 - <i>Dashboard: SDG Contributions Over Time</i>	29
FIGURA 12 - <i>Dashboard: SDG Contributions Over Time</i>	29
FIGURA 13 - <i>Dashboard: SDG Group Contributions</i>	30
FIGURA 14 - <i>Dashboard: Master and PhD Thesis</i>	31
FIGURA 15 - <i>Dashboard: SDG Contributions Indicators</i>	31
FIGURA 16 - <i>Dashboard SDG Analysis: Bibliometrics of Relevance</i>	40
FIGURA 17 - <i>Script em Python do Algoritmo de Classificação</i>	41
FIGURA 18 - <i>Dashboard: SDG Profile</i>	42
FIGURA 19 - <i>Dashboard: SDG Contributions</i>	42
FIGURA 20 - <i>Dashboard: SDG Group</i>	43
FIGURA 21 - <i>Dashboard: Articles - Scopus</i>	43
FIGURA 22 - <i>Dashboard: Course Units</i>	43

ÍNDICE DE TABELAS

TABELA I – Exemplos de <i>keywords</i> provenientes do conjunto utilizado.....	20
TABELA II – Variáveis utilizadas no algoritmo de classificação	24
TABELA III - 17 Objetivos de Desenvolvimento Sustentável.....	40
TABELA IV - Comparação entre os três conjuntos de <i>keywords</i>	40
TABELA V - Relação entre as tabelas de fato e de dimensão.....	41
TABELA VI - Métricas realizadas no Power BI	42

LISTA DE ABREVIATURAS

BI – *Business Intelligence*

DW – *Data Warehouse*

ETL – *Extract Transform Load*

HEI - *Higher Education Institutions*

ISEG – Instituição Superior de Economia e Gestão

ODS – Objetivos de Desenvolvimento Sustentável

ONU – Organização das Nações Unidas

SDG - *Sustainable Development Goals*

SI – Sistema de Informação

SSD - Sistemas de Suporte à Decisão

1. INTRODUÇÃO

1.1. ENQUADRAMENTO E MOTIVAÇÃO

O mundo atual é caracterizado pelo aumento constante do impacto da atividade humana em todos os ecossistemas que resulta em constantes desafios ambientais, sociais e económicos (Körffgen et al., 2018). O estilo de vida da população atual leva a que o conceito de Desenvolvimento Sustentável, que segundo o Bruntland Commission Report (1987), é a capacidade de satisfazer a geração presente sem comprometer as gerações futuras de satisfazerem as suas próprias necessidades (UNESCO, 2012), seja um tema cada vez mais relevante nos dias de hoje. Para lidar com estes desafios, não é suficiente haver acordos políticos, incentivos financeiros e soluções tecnológicas para enfrentar os desafios do desenvolvimento sustentável, mas sim, uma mudança completa na maneira como pensamos e interagimos com os ecossistemas. Para isso, todos os indivíduos e sociedades devem estar equipados e capacitados com conhecimentos, competências e valores, para impulsionar essa mudança (UNESCO, 2014).

Neste sentido, a Organização das Nações Unidas (ONU) criou em 2015 a Agenda 2030, assente em 17 Objetivos de Desenvolvimento Sustentável (ODS) (Anexo I) e 169 metas a serem atingidas até 2030 (Relatório Voluntário Nacional, 2017). Estes objetivos caracterizam-se por serem os pilares basilares de uma sociedade moderna e equilibrada, que requerem uma ação a nível mundial, por parte dos governos, empresas e sociedade civil (BCSD Portugal, 2019), para responder aos desafios que a sociedade enfrenta e criar uma vida com dignidade e oportunidades para todos (ODS, 2020). Para a concretização destes objetivos por parte das organizações é crucial a integração destas metas nas suas tomadas de decisão de modo a contribuírem com o seu poder de inovação para um futuro mais sustentável e inclusivo (BCSD Portugal, 2019).

Esta preocupação com a sustentabilidade não é diferente no setor do ensino, visto ser um dos poucos setores que pode apoiar, promover e contribuir para todos os ODS (Junior et al., 2019), como também tem um papel fundamental em produzir e amplificar a ação em todos os níveis e áreas de ensino, de forma a acelerar o progresso para o desenvolvimento sustentável (UNESCO, 2014). Em particular, as Instituições do Ensino Superior (IES) são cruciais nessa contribuição (Junior et al., 2019), uma vez que são entidades privilegiadas na produção e na comunicação de conhecimento através do ensino

e da investigação, e por isso devem desempenhar um papel preponderante na promoção do conceito de sustentabilidade, atuar como impulsionador da mudança em direção a um mundo sustentável e estar na vanguarda dos esforços globais de sustentabilidade (Haupt et al., 2015). Deste modo, as IES devem assumir um papel proativo e de liderança na contribuição e na implementação intersectorial dos ODS, proporcionando uma fonte valiosa de conhecimento em investigação e educação, abrangendo todas as áreas dos ODS (El-Jardali et al., 2018).

A partir da década de 1990, começaram as Declarações de Sustentabilidade no Ensino Superior, desde então as IES têm feito, ao longo do tempo, esforços significativos para integrar a sustentabilidade na reorientação e no reforço das suas atividades de ensino e de aprendizagem. As IES para serem líderes na promoção dos ODS, têm de incluir competências essenciais do desenvolvimento sustentável nas suas atividades; identificar os conteúdos e atividades realizadas pela instituição que estão vinculados aos ODS; e providenciar formação aos docentes, de forma a reforçar os seus recursos didáticos (Albareda-Tiana et al., 2018).

De forma a contribuírem para os ODS, as IES necessitam de identificar a relevância e o impacto social, a qualidade e o impacto académico das suas atividades relacionadas com a sustentabilidade (Aurora, 2018). Estas atividades correspondem às principais funções das universidades e englobam as seguintes vertentes: o campus, a investigação, o ensino e a sociedade. As IES podem contribuir para os ODS, de diversas formas, através das suas quatro atividades gerais (SDSN Australia/Pacific, 2017).

Conhecer qual a situação atual da IES relativamente aos ODS através do mapeamento e classificação das suas atividades de acordo com os ODS (SDSN Australia/Pacific, 2017), é essencial para que possa conduzir a uma análise eficaz das informações de sustentabilidade, permitindo uma melhor gestão e medição dos seus esforços, bem como uma melhor gestão dos riscos e identificação de oportunidades a longo prazo (Scholtz et al., 2018). Para a gestão e reporte eficaz das informações de sustentabilidade por parte das IES, é necessária uma ferramenta que permita analisar os dados de forma eficaz; aceder a informação atualizada e disponível quando requisitada; apoiar no processo de tomada de decisão e que possibilite a transmissão de conhecimento (Scholtz et al., 2018; Haupt et al., 2015). Neste sentido, *Business Intelligence* (BI) surge como a solução

indicada, tendo um papel relevante a desempenhar para ajudar as instituições a implementar, monitorizar e comunicar as práticas sustentáveis (Petrini & Pozzebon, 2009).

Apesar do interesse e preocupação com as práticas sustentáveis, a sua implementação apresenta obstáculos (Petrini & Pozzebon, 2009). Por exemplo, a informação encontra-se dispersa por diferentes fontes de dados, a qualidade dos dados é reduzida e os relatórios são ineficazes (Scholtz et al., 2018). Existe uma separação entre a temática da sustentabilidade e as considerações estratégicas e o desempenho do negócio (Petrini & Pozzebon, 2009). Verifica-se também, dificuldade em definir e aplicar indicadores que permitam analisar problemas, apoiar no processo de decisão e na transmissão de conhecimento (Scipioni et al., 2009). Para além destas dificuldades, a Agenda de 2030, para a maioria dos países signatários, não apresenta uma visão geral das contribuições dos conteúdos produzidos pelas IES que abordam os ODS (Körffgen et al., 2018). Deste modo, considerou-se importante estudar a aplicabilidade das ferramentas de BI como instrumento relevante na identificação e na comunicação das contribuições realizadas para os ODS.

O Instituto Superior de Economia e Gestão (ISEG), sendo uma IES, tem como orientação estratégica reforçar o seu posicionamento enquanto universidade sustentável, que contribui para a implementação dos ODS. Como IES, o ISEG tem a responsabilidade de atuar como impulsionadora na implementação dos ODS (Junior et al., 2019). Esta responsabilidade foi reforçada em 2020, quando o ISEG se associou à Aliança ODS Portugal, que consiste na parceria entre várias entidades públicas e privadas que se comprometem no âmbito da Agenda 2030 e dos ODS (ISEG, 2020). O compromisso, por parte da instituição, tem o intuito de aumentar a consciencialização, medir o impacto e comunicar as contribuições efetuadas pela instituição para alcançar os ODS.

Atendendo às necessidades da instituição e tendo em consideração que não existe nenhuma plataforma que identifique e comunique o seu posicionamento relativamente ao desenvolvimento sustentável, este projeto sugere a classificação e o mapeamento das suas atividades de investigação e de ensino que contribuem para os ODS, e a apresentação dos resultados através de *dashboards*. Deste modo, vai ser possível identificar a situação atual, bem como a sua evolução ao longo dos anos, relativamente às contribuições

realizadas pela instituição. Uma parte dos *dashboards* vão ser divulgados no *website* da instituição, de forma a mostrar as suas contribuições ao público, enquanto que outros vão ser direcionados para a Presidência e para o grupo de trabalho dos ODS do ISEG. Estes *dashboards* vão facilitar a gestão e a monitorização das atividades de investigação e de ensino que contribuem para os ODS, bem como apoiar a tomada de decisão para a definição de estratégias futuras com base nos *insights* dos *dashboards* e perceber o seu grau de compromisso com a Agenda 2030.

O documento encontra-se repartido em cinco capítulos, este primeiro capítulo diz respeito ao enquadramento e à identificação dos objetivos do projeto. No capítulo 2, é apresentada a revisão da literatura, que aborda o conhecimento teórico necessário para a edificação e compreensão do projeto. No capítulo 3, são identificadas e descritas todas as fases realizadas no desenvolvimento do projeto. No capítulo 4, é apresentada as principais conclusões com base nos resultados obtidos. Por fim o capítulo 5, aborda as principais limitações e recomendações para trabalhos futuros.

1.2. OBJETIVOS DO PROJETO

O projeto tem como objetivo principal classificar e comunicar as atividades de ensino e de investigação realizadas pelo ISEG entre 2011 e 2020, de acordo com a matriz dos ODS definida na Agenda 2030. Deste modo, o projeto tem o intuito de realizar o mapeamento das suas atividades de ensino e de investigação de acordo com os ODS, e a edificação de *dashboards* com base na aplicação de boas práticas de desenho para comunicar e monitorizar as contribuições.

De forma a concretizar o objetivo principal mencionado anteriormente, deverão ser realizados os seguintes objetivos secundários:

- a) Seleção das bases de dados e do conjunto de *keywords* a usar na identificação dos ODS presentes nas atividades;
- b) Definição do processo de tratamento dos dados, para garantir a qualidade dos mesmos;
- c) Criação do algoritmo de classificação de artigos, teses e programas de unidades curriculares por ODS;
- d) Criação de *dashboards* para a visualização da informação;
- e) Análise dos resultados obtidos;

f) Identificação de oportunidades de melhoria.

Investigar plataformas e processos semelhantes ao artefacto do projeto será importante para identificar as abordagens utilizadas por outras organizações de ensino no mapeamento e na visualização dos dados, este será um elemento crucial para o desenvolvimento do projeto.

1.3. RELEVÂNCIA DOS OBJETIVOS

O projeto vai permitir ao ISEG reforçar o seu posicionamento enquanto universidade que contribui ativamente para a temática e para a implementação dos ODS. Adicionalmente, vai proporcionar uma maior capacidade de análise crítica e de tomada de decisão, com base no mapeamento da situação atual e assim permitir identificar os ODS prioritários para a instituição.

Este trabalho permite automatizar o processo de mapeamento das atividades realizadas pela instituição, permitindo uma redução do tempo, uma maior eficácia na realização dos processos e na obtenção de informação, assim como uma maior qualidade de informação. Adicionalmente permite, a apresentação da informação através de *dashboards* simples e intuitivos que demonstram o compromisso e a contribuição académica do ISEG. Estes possibilitam a análise e o cruzamento de dados de diferentes fontes; a análise da evolução ao longo do tempo; a apresentação de um conjunto de métricas e indicadores significativos para a compreensão da informação; e um melhor acesso à informação. Deste modo, este trabalho inicial é fundamental para ajudar na definição dos próximos passos em relação à estratégia da instituição.

A nível da revisão da literatura, este estudo pretende contribuir para o enriquecimento das temáticas sobre os ODS nas universidades e BI, assim como reconhecer a importância da correlação existente entre os temas, uma vez que são temáticas bastante significativas e debatidas nos dias de hoje. Simultaneamente, a implementação do projeto no ISEG pode ser relevante para ajudar outras instituições, com um *use case* semelhante, na introdução e na implementação dos ODS, uma vez que se trata de uma abordagem simples e flexível, sem a necessidade de elevados conhecimentos técnicos.

2. REVISÃO DA LITERATURA

Neste capítulo será realizada a revisão de literatura dos temas e tópicos análogos ao tema do projeto. Vão ser apresentadas, sucintamente, as principais análises e as considerações feitas por alguns autores, como também trabalhos relacionados. Deste modo, vão ser abordados os temas relacionados com os ODS, a sustentabilidade como parte integrante da estratégia nas IES e o BI.

2.1. OBJETIVOS DE DESENVOLVIMENTO SUSTENTÁVEL

A 25 de setembro de 2015, a Assembleia Geral das Nações Unidas adotou por unanimidade na Cimeira de Chefes de Estado e de Governo a iniciativa intitulada de “Transforma o nosso mundo: a Agenda 2030 para o Desenvolvimento Sustentável”, que entrou em vigor em 1 de janeiro de 2016, assinada por mais de 190 países (Relatório Voluntário Nacional, 2017). A Agenda de 2030 é um “plano de ação centrado nas pessoas, no planeta, na prosperidade, na paz e nas parcerias (5P), tendo como objetivo final a erradicação da pobreza e o desenvolvimento sustentável, no âmbito do qual todos os Estados e outras partes interessadas assumem responsabilidades próprias no que diz respeito à sua implementação, enfatizando-se que ninguém deve ser deixado para trás” (Relatório Voluntário Nacional, 2017, p.5), que pretende guiar o mundo no caminho de um futuro melhor para todos até 2030.

A Agenda de 2030 assenta em 17 Objetivos de Desenvolvimento Sustentável (ODS) (Figura 1) e em 169 metas a serem implementadas em todos os países (desenvolvidos e em desenvolvimento), surgiram dos Objetivos de Desenvolvimento do Milénio, estabelecidos entre 2000 e 2015, que tinham como intuito apoiar apenas os países em desenvolvimento. Por sua vez, a Agenda de 2030 “pressupõe a integração dos ODS nas políticas, processos e ações desenvolvidas nos planos nacional, regional e global” (Relatório Voluntário Nacional, 2017, p.5).

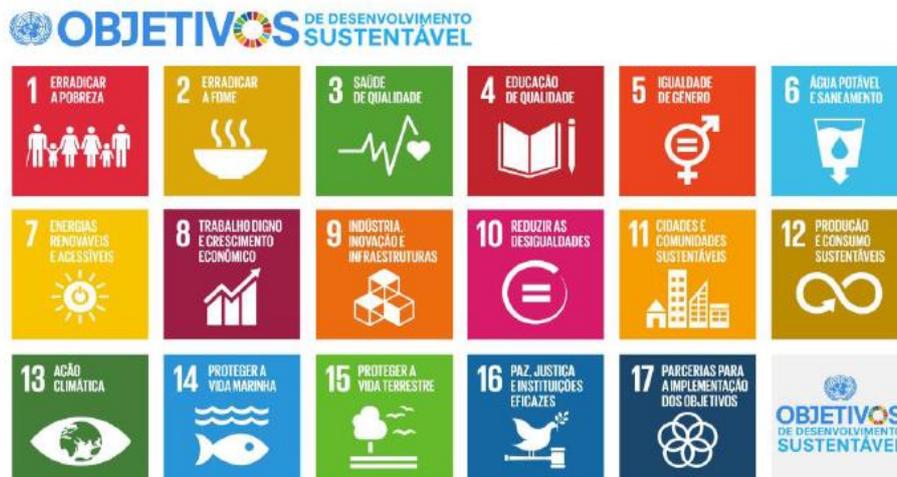


FIGURA 1 - Objetivos de Desenvolvimento Sustentável (BCSD Portugal, 2019)

Para o Secretário Geral da ONU Ban Ki-Moon, “os 17 ODS são a nossa visão comum para a Humanidade e um contrato social entre os líderes mundiais e os povos” , acrescentando que os mesmos “são uma lista das coisas a fazer em nome dos povos e do planeta, e um plano para o sucesso” (UNRIC, 2018, p.1).

Os ODS abordam três dimensões do desenvolvimento sustentável nos seus objetivos, nomeadamente, a dimensão social, económica e ambiental, promovendo a justiça, a paz e as instituições eficazes (UNRIC, 2018), e atua sobre cinco áreas que são identificadas como os 5P: **peessoas** - erradicação da pobreza e da fome, respeito da dignidade e da igualdade (ODS 1, 2, 3, 4, 5 e 6); **planeta** - produção e consumo sustentáveis, combater as mudanças climáticas e gestão de recursos naturais (ODS 7, 8, 9 e 10); **prosperidade** - garantir que todas as pessoas usufruem de prosperidade, saúde e progresso económico e social (ODS 11, 12, 13, 14 e 15); **paz** - promover sociedades pacíficas, justas e inclusivas que sejam livres de medo e violência (ODS 16); **parcerias** - integração transversal, interconexão e mobilização conjunta em prol dos mais vulneráveis (ODS 17) (SDSN Australia/Pacific, 2017; Relatório Voluntário Nacional, 2017).

Enfrentar os desafios dos ODS é responsabilidade de todos, das empresas, da sociedade civil, do setor terciário e do setor do ensino, entre outros, que deverão colaborar para conquistar os ODS (SDSN Australia/Pacific, 2017) de forma singular ou em parcerias (Junior et al., 2019).

2.2. ODS E AS INSTITUIÇÕES DE ENSINO SUPERIOR

Os ODS devem ser usados pelas IES como um instrumento crítico para introduzir e incorporar o desenvolvimento sustentável nas suas atividades e estratégias (Junior et al., 2019). Estas desempenham um papel preponderante na criação e na disseminação de conhecimento, sendo impulsionadoras da inovação global, nacional e local, assim como do desenvolvimento econômico e bem-estar social (SDSN Australia/Pacific, 2017). Deste modo, as universidades têm um elevado impacto positivo como impulsionadoras na implementação dos ODS (Körfggen et al., 2018). Segundo o SDSN Australia/Pacific (2017), o compromisso das instituições para com os ODS reflete-se nas suas atividades de ensino; na produção de investigação; na interação com a sociedade; e nas operações e na gestão da instituição, como se pode verificar na Figura 2.



FIGURA 2 - Quatro pilares das contribuições das IES (SDSD Australia/Pacific, 2017)

As IES, são elementos cruciais para alcançar os ODS (Junior et al., 2019). Têm a capacidade e os meios de prover a geração atual e futura com conhecimentos, capacidades e motivação (SDSN Australia/Pacific, 2017) para enfrentar os desafios e as oportunidades que advém da temática de sustentabilidade, como também, a capacidade de incorporar o desenvolvimento sustentável em investigações que impulsionem a inovação, soluções e conhecimentos, que contribuam para a Agenda (Junior et al., 2019; SDSN Australia/Pacific, 2017). Podem reorientar os programas lecionados para uma educação orientada para as questões sociais, económicas e ambientais (Tilbury, 2012; Filho, 2011),

e adicionalmente fortalecer as capacidades dos professores e funcionários para ajudar a incorporar o desenvolvimento sustentável na instituição (Vaughter, 2018).

A nível organizacional, podem apoiar, adotar e implementar boas práticas dentro das respetivas instituições (SDSN Australia/Pacific, 2017), por meio da sua cultura, das atividades realizadas no campus, da gestão da instituição, e da monitorizar e comunicar as suas contribuições para o desenvolvimento sustentável e para os ODS (Junior et al., 2019). Fora do campus, para impulsionar e fortificar as IES como impulsores importantes no desenvolvimento sustentável, é importante construir, fortalecer e institucionalizar ligações e parcerias internas e externas, com o governo, com comunidades científicas e com outras IES (Körfggen et al., 2018), para aumentar a conscientização e a comunicação das questões de desenvolvimento que são prioridade a nível local ou nacional e os desafios que estão a ser negligenciados ou desvalorizados (Vaughter, 2018).

Por outro lado, as IES também retiram um conjunto de benefícios resultantes do apoio e da integração dos ODS. Nomeadamente, a adoção de uma definição abrangente e globalmente aceite do que é considerado uma universidade responsável e comprometida; acesso a novas fontes de financiamento (órgãos governamentais, bancos, entre outros) para ajudar no cumprimento dos ODS; criação de parcerias e colaborações externas e internas, uma vez que a Agenda fornece uma estrutura comum partilhada por organizações de diferentes setores que se conectam e trabalham em conjunto; adoção de uma educação relacionada com os ODS; e uma estrutura para demonstrar o impacto e a relevância que as IES têm na contribuição para o bem-estar local e global (SDSN Australia/Pacific, 2017).

2.2.1. INTEGRAÇÃO DOS ODS

As IES que se comprometem com a sustentabilidade estão a lutar para contribuir significativamente para o tema, no entanto a dificuldade destas prende-se em iniciar projetos/iniciativas que influenciem globalmente a sua cultura e os seus intervenientes (Tilbury, 2012). Deste modo, o desenvolvimento sustentável é um processo complexo e de longo prazo que requer uma mudança do paradigma de educação, que englobe a transformação da responsabilidade institucional e a reorientação do programa curricular

de modo a atender as necessidades das gerações atuais e futuras (Albareda-Tiana et al., 2018).

Para acelerar as contribuições para os ODS, o SDSN Australia/Pacific (2017) fornece um guia “*Getting Started with the SDGs in Universities*” para trabalhar em direção a soluções estruturais. Este guia apresenta um conjunto de passos (Figura 3) que devem ser tidos em consideração para iniciar e aprofundar o envolvimento com os ODS, e cada IES de acordo com os seus objetivos, necessidades, contexto e características podem ter em consideração uma sequência de passos distintos. Assim sendo, não existe nenhuma maneira correta de as IES implementarem e apoiarem os ODS (SDSN Australia/Pacific, 2017). Os três níveis para integrar os ODS são os seguintes:

- Reconhecimento: Identificar o que a IES realiza para contribuir para os ODS, através do mapeamento das atividades, com o intuito de fornecer uma narrativa do seu impacto e proporcionar uma visão para futuras ações;
- Oportunidades: cada uma das áreas da IES reconhece as vantagens e a importância dos ODS e identifica as prioridades, oportunidades e os *gaps* da implementação dos ODS na estruturação e na orientação das atividades e programas de forma coerente e em comum, entre as partes interessadas;
- Princípio de Organização: integração e implementação de compromissos e ações relacionadas com os ODS, nas estratégias e nas políticas da IES, de modo a ter um impacto constante nos ODS. Assim como, a criação de relatórios e a comunicação do compromisso e das contribuições realizadas para os ODS.



FIGURA 3 - Passos para a integração dos ODS (SDSN Australia/Pacific, 2017)

2.2.1.1. MAPEAMENTO DAS ATIVIDADES DAS IES

As IES contribuem para os ODS de muitas maneiras diferentes. A realização do mapeamento é uma prática benéfica no âmbito universitário para iniciar, implementar e comunicar os ODS para os quais estão a contribuir através das suas atividades (SDSN Australia/Pacific, 2017). O mapeamento das contribuições permite identificar as áreas de foco e as áreas prioritárias (Albareda-Tiana et al., 2018); identificar as competências das IES e as atividades realizadas que contribuem para os ODS; proporcionar oportunidades de colaboração entre os principais grupos de interesse; incorporar os ODS nas estruturas das IES (Körfggen et al., 2018); identificar os pontos fortes e os *gaps* nas atividades realizadas e a comunicação das contribuições realizadas pelas IES (SDSN Australia/Pacific, 2017).

O mapeamento pode ser realizado tendo em consideração três abordagens: abordagem manual, consiste na avaliação e identificação manual das atividades relacionadas com os ODS, recomendado para pequenas bases de dados; auto-identificação, consiste na identificação manual dos ODS a realizar por parte das indivíduos que produziram a atividade; por ultimo, a utilização de *keywords*, que inclui a utilização de um conjunto de *keywords* para cada um dos ODS, com o intuito de identificar a que ODS cada atividade está associada através de um processo automatizado (SDSN Australia/Pacific, 2017).

As *keywords* representam o foco dos conteúdos/atividades e os seus respetivos tópicos de pesquisa descritiva dentro do seu domínio (Su & Lee, 2010). Esta última abordagem é amplamente adotada por IES no mapeamento das suas atividades (Junior et al., 2019). Tem como limitação a variedade de significados que uma *keywords* poder apresentar, isto pode levar à não identificação ou à identificação incorreta. Esta pode ser contrariada com a realização de testes nas bases de dados de forma a encontrar erros, comparação entre conjuntos de *keywords* e com a participação das partes interessadas no refinamento e na validação das mesmas (Junior et al., 2019; Körfggen et al., 2018)

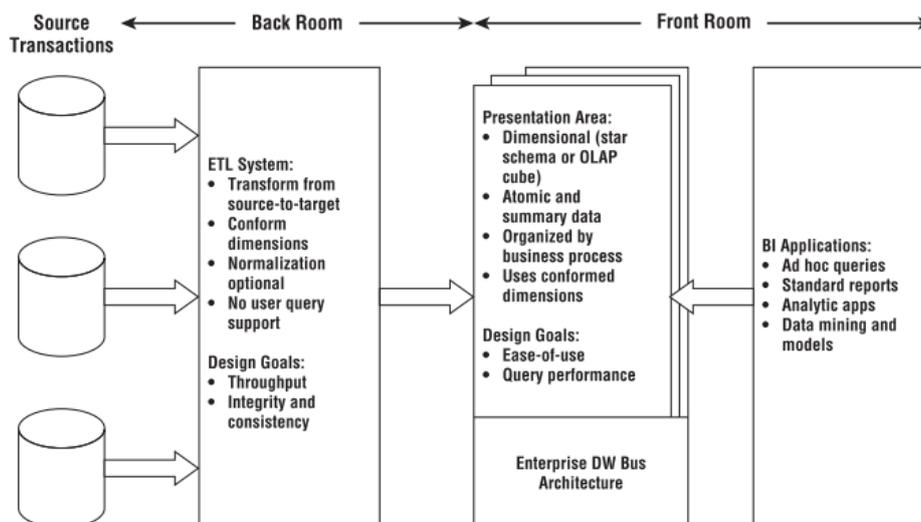
2.3. BUSINESS INTELLIGENCE

O termo *Business Intelligence* (BI), como é compreendido nos dias de hoje, foi introduzido por Luhn (1958), com um conceito semelhante ao usado atualmente (Chan et al., 2009). BI teve origem nos Sistemas de Suporte à Decisão (SSD) e resultou da criação de uma variedade de SSD que permitiram alargar o seu domínio (Watson & Wixom, 2007). Em 1989, Howard Dresner, analista do *Gartner Group*, introduziu o termo BI como sendo “um conjunto de softwares e soluções para agrupar, consolidar, analisar e fornecer acesso aos dados de forma a permitir aos utilizadores melhorar a tomada de decisão no negócio” (Chan et al., 2009, p.96).

Atualmente, apesar de não existir uma definição globalmente aceite, o seu significado pode apresentar diversas interpretações, conforme as áreas de intervenção e de estudo (Watson, 2009). Independentemente das diferentes definições, uma comparação entre as mesmas, mostrou que estas se baseiam em três categorias, sendo elas a gestão (processo), tecnologia e o produto (Chan et al., 2009). De forma abrangente, BI tem o objetivo transformar os dados em conhecimento (Niu et al., 2009). Segundo Watson (2009), BI refere-se a um conjunto de aplicações, tecnologias e processos com o intuito de recolher, armazenar, aceder e analisar dados de forma a ajudar no suporte à tomada de decisão. Tem o propósito de gerar *insights* e informações importantes sobre os objetivos estabelecidos pela organização, oportunidades e apoiar no processo de tomada de decisão (Wieder & Ossimitz, 2015). Sendo assim, BI é uma iniciativa estratégica que agora é conhecido como um instrumento para impulsionar a eficácia e a inovação do negócio (Watson & Wixom, 2007).

2.3.1. ARQUITETURA DE BI

Mediante a revisão de literatura, verifica-se que existem diversas arquiteturas de BI, que diferem nas suas camadas, componentes, processos e relacionamentos (Ong et al., 2011). Segundo Kimball & Ross (2013), a arquitetura de um sistema de BI apresenta quatro componentes. Estes colaboram entre si para facilitar as funções básicas de BI: extração de dados dos sistemas operacionais, sistemas de ETL, armazenamento dos dados num *Data Warehouse* (DW) e a apresentação dos dados nas várias aplicações de BI (Figura 4) (Kimball & Ross, 2013).

FIGURA 4 - Arquitetura de *Business Intelligence* (Kimball & Ross, 2013)

Sistemas Operacionais de Origem

A primeira parte consiste na identificação das diferentes fontes de dados (Ong et al., 2011). Os dados podem ser obtidos de diferentes Sistemas de Informação (SI) (Watson, 2009), provenientes de fontes internas ou externas à organização (Ong et al., 2011). Os SI têm como principais prioridades o processamento e a disponibilização dos dados, tendo pouco ou nenhum controlo sobre o conteúdo e o formato dos dados, ao contrário do DW (Kimball & Ross, 2013).

ETL (*Extract, Transform and Load*)

Esta componente consiste no pré-processamento dos dados e inclui três fases: extração, transformação e carregamento de dados (Niu et al., 2009). Primeiramente os dados são extraídos das diferentes fontes para o ambiente do DW (Kimball & Ross, 2013), esta etapa é importante para selecionar e compreender os dados que são cruciais para o suporte à tomada de decisão organizacional (Ong et al., 2011). Posteriormente sofrem um conjunto de transformações que consistem na limpeza, uniformização, agregação de dados, entre outras, com base num conjunto de regras de negócio, de forma a melhorar a qualidade dos dados (Watson, 2009). Por último, ocorre a estruturação física e o carregamento dos dados nos modelos dimensionais no DW (Kimball & Ross, 2013). Este componente é crucial, uma vez que a qualidade dos dados de todos os componentes seguintes dependem deste processo (Niu et al., 2009).

Data Warehouse

Segundo Inmon (2005, p.31), o DW é um dos componentes mais importantes que se caracteriza por ser um “conjunto de dados orientados por assunto, integrado, não volátil e variante na tempo para apoiar no processo de tomada de decisão”. Deste modo, o DW consiste num repositório central de dados que recolhe e armazena dados de diferentes fontes, internas ou externas, possibilitando análises e consultas para apoiar a tomada de decisão de toda a organização (Ong et al., 2011). Para facilitar a compreensão e simplificar a navegação em grandes volumes de dados, foram desenvolvidos os modelos dimensionais (Kimball & Ross, 2013). Segundo, Kimball & Ross (2013), a modelagem dimensional é uma técnica que tornar a DW mais compreensível para o utilizador, com uma navegação e um desempenho de resposta mais rápido e eficiente para ajudar os utilizadores na tomada de decisão.

Modelos dimensionais implementados em sistemas de gestão de base de dados relacionais, são referenciados como *star schema* (modelo em estrela) e *snowflake* (modelo de floco de neve), por norma estes são os esquemas mais utilizados (Niu et al., 2009). Cada um destes modelos é composto por tabelas de fato que estão associadas a tabelas de dimensão através de chaves primárias/estrangeiras, por um relacionamento de um para muitos (Moody & Kortink, 2000). As tabelas de fato guardam as informações de negócio de uma organização, consistem em tabelas centralizadas com todos os dados que não devem ser replicados em vários lugares e por várias funções da organização, sendo constituídos por dados fatuais e quantitativos, em função da sua granularidade. As tabelas de dimensão, por sua vez, contêm as descrições das informações de negócio presente nas tabelas de fato. Cada dimensão serve como fonte de restrição de *queries*, agrupamentos e *labels*. Tanto no *star* como no *snowflake schema*, as tabelas de fato estão associadas a tabelas de dimensões, a única diferença incida na normalização das tabelas de dimensão no modelo *snowflake* num conjunto de tabelas de dimensões de tamanho inferior, ou seja, o modelo *snowflake* é uma extensão do modelo *star* em que as tabelas dimensão podem representar o centro de outra estrela (Kimball & Ross, 2013).

Aplicações de Business Intelligence

As aplicações de BI compreendem a análise dos dados provenientes do DW através de um conjunto de ferramentas que exibem informações em diferentes formatos aos utilizadores para melhorar a tomada de decisão (Ong et al., 2011). Segundo Kimball & Ross (2013), as aplicações de BI podem ser: ferramenta de consulta Ad hoc, ferramentas de visualização, aplicações analíticas e *data mining*.

2.3.2. DASHBOARDS

Segundo, Few (2006, p.26), um *dashboard* é a “apresentação visual das informações mais importantes necessárias para atingir um ou mais objetivos, com a informação consolidada e organizada num único ecrã para que esta possa ser monitorizada de imediato”. O desenvolvimento do *dashboard* deve estar alinhado com o negócio para refletir de forma eficiente o propósito do mesmo. Assim, as visualizações devem seguir uma ordem que permita ao utilizador analisar e interpretar os dados (Few, 2007).

Para Eckerson (2011), um *dashboard* é um sistema de gestão de desempenho que fornece informações aos utilizadores para medir, monitorizar e gerir com eficácia esse mesmo desempenho, metas e objetivos, através de dados em tempo real. Apresenta os dados consolidados numa *interface* sob a forma de gráficos, tabelas, ícones, imagens, texto, indicadores e mecanismos de alerta (Malik, 2005).

Segundo Malik (2005), um *dashboard* deve apresentar um conjunto de características básicas, sendo elas caracterizadas por SMART. Este deve ser “*Synergetic*”, ou seja, visualmente eficazes de maneira a permitir ao utilizador criar sinergias relativamente às informação compiladas em cada ecrã; “*Monitor KPIs*”, exibir os KPIs cruciais para a tomada de decisão; “*Accurate*”, deve conter informação precisa de modo a transmitir confiança ao utilizador; “*Responsive*”, deve responder a limites predefinidas através da criação de alertas (alarmes, e-mails, entre outros) com o intuito de avisar o utilizador para assuntos importantes; e por ultimo “*Timely*”, deve apresentar informações o mais atualizadas possíveis para uma boa tomada de decisão (Malik, 2005)

Um *dashboard* deve ser funcional e perceptível, de forma a ir de encontro às necessidades dos utilizadores, como também deve ter em consideração o seu aspeto

estético e a qualidade da interação, de forma a ser apelativo (Wexler et al., 2017). Segundo Few (2006), existe um conjunto de erros comuns no desenho de *dashboards* que devem ser evitados, destaca-se os seguintes erros: a informação ultrapassar o tamanho de um ecrã, os dados sobre o mesmo assunto estarem em diferentes ecrãs, existir dados sem contexto, o uso excessivo de cores, detalhes excessivos sobre os dados, organização inadequada do conteúdo, seleção de gráficos inadequados, identificação ineficaz dos dados mais importantes e visualizações pouco apelativas.

2.4. TRABALHOS RELACIONADOS EM IES

A “*Alliance for Sustainable Universities*”, *network* Austríaca, iniciou o projeto “*Universities and Sustainable Development Goals*”. O projeto propõe uma ferramenta que realiza o mapeamento das publicações científicas e dos projetos de investigação com base na definição de desenvolvimento sustentável presente nos ODS, e contou com o mapeamento de treze universidade para encontrar as áreas de foco das suas atividades relacionadas com a sustentabilidade. Este protejo tem o intuito de fortalecer a cooperação entre universidades para a integração do desenvolvimento sustentável; fortalecer as interações entre a ciência, sociedade e política; e a identificação de opções para os tomadores de decisão políticos e sociais. O projeto utiliza no mapeamento um conjunto de *keywords*, que resultaram da combinação da lista fornecida pela SDSN Australia/Pacific juntamente com as *keywords* formuladas pelas universidades parceiras (Körfigen et al., 2018).

“*SDG Analysis: Bibliometrics of Relevance*”, é um *dashboard* interativo em Power BI que mostra a relevância e o impacto social das atividades de investigação das universidades que fazem parte do *network* Aurora, relativamente aos ODS (Anexo II). Tem o intuito de comunicar às universidades parceiras em quais ODS estas estão a contribuir, quantificar as contribuições em cada ODS, fortalecer o foco nas investigações e nos desafios sociais e identificar a importância das publicações com base nos seus indicadores bibliométricos. Adicionalmente, permitir a colaboração e a criação de alianças dentro e fora do *network* Aurora (Aurora, 2018).

3. METODOLOGIA

O presente capítulo consiste na realização do mapeamento e da criação de um sistema de BI com o objetivo de comunicar as contribuições realizadas e apoiar no processo de tomada de decisão, com base nos conhecimentos apreendidos no capítulo anterior sobre as temáticas do projeto.

O projeto apresenta seis fases relevantes para a criação do artefacto. Antes da realização de cada uma destas fases existiu primeiramente a definição do âmbito e dos objetivos do projeto, assim como o planeamento das principais atividades a realizar. A Figura 5, apresenta o fluxograma das fases que foram realizadas no presente projeto.

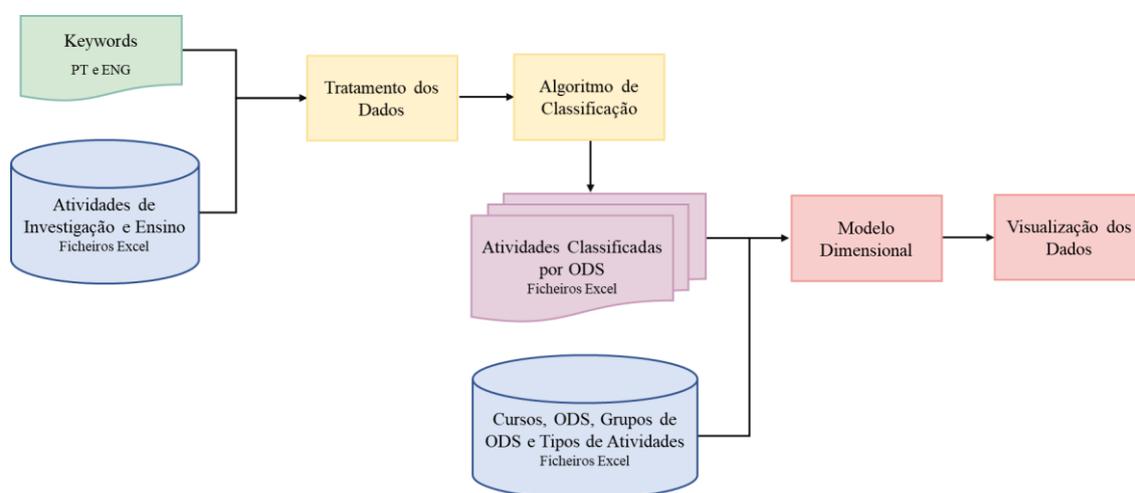


FIGURA 5 - Fluxograma do projeto

Primeiramente é realizada a identificação das fontes de dados que vão de encontro com as necessidades da instituição, e de seguida, a extração dos ficheiros dos seus respetivos SI. Seguidamente, é selecionado o conjunto de *keywords* mais robusto e fiável a utilizar de acordo com o número de correspondências e afinações a realizar.

A partir deste passo, dá-se início ao desenho e desenvolvimento do artefacto, no qual se vai realizar o tratamento e a limpeza dos dados. Esta fase vai ser acompanhada do desenvolvimento do algoritmo de classificação, que realiza o mapeamento das atividades de ensino e de investigação através de *keywords* alusivas à matriz dos ODS. A próxima fase diz respeito à análise dos dados proveniente do algoritmo através da criação de um modelo *snowflake*, e de seguida, a criação de *dashboards* que permitem analisar a informação de uma forma simples e objetiva. Estas duas últimas fases foram realizadas

na plataforma de BI, mais concretamente no Power BI. Para finalizar, vai-se proceder à avaliação dos resultados e à identificação das limitações do estudo.

3.1. ENQUADRAMENTO DO PROJETO

O projeto foi realizado sob a orientação do ISEG, com base nos seus valores, missão e visão. Segundo o SDSN Australia/Pacific (2017), as IES podem contribuir para os ODS em quatro vertentes diferentes, que são a investigação, o ensino, o campus e a sociedade. Apesar dessas quatro vertentes serem muito relevantes para o ISEG, o âmbito deste projeto tem em consideração apenas duas vertentes na realização do mapeamento, a investigação e o ensino. Ao nível das atividades de investigação considerou-se as teses de mestrado, teses de doutoramento e artigos científicos. A nível das atividades de ensino considerou-se os sílabos das unidades curriculares do ano letivo de 2019/2020.

O artefacto deste estudo permite conhecer a situação atual do ISEG nesses vários domínios, através da identificação quantitativa das contribuições realizadas e do seu compromisso com os ODS. Neste sentido, o projeto pretende criar *dashboards* para comunicar as contribuições realizadas aos *end-user*. Alguns *dashboards* são direcionados para o público através da divulgação dos mesmos no *website* da instituição, enquanto que outros são direcionados para a Presidência e para o grupo de trabalho dos ODS do ISEG. A atualização dos dados nos *dashboards* vai ser realizado de forma anual, uma vez que as atividades que se consideraram para o artefacto apresentam mudanças pontuais durante o decorrer de cada ano letivo.

3.2. IDENTIFICAÇÃO DAS FONTES DE DADOS

As informações no projeto provêm de diferentes SI, nomeadamente da plataforma Aquila (repositório interno) e Scopus (repositório externo). Todos os ficheiros utilizados apresentam um grande número de dados que não sofreram qualquer tipo de análise relacionada com os ODS. Nestes ficheiros são recolhidas informações sobre as atividades de ensino e de investigação produzidas pelos alunos e docentes do ISEG.

Foram selecionadas as seguintes fontes de dados:

- Ficheiro Excel, proveniente da plataforma Scopus, apresenta a listagem de artigos científicos realizados por docentes e estudantes– ficheiro denominado de “Scopus”;
- Ficheiro Excel, proveniente da plataforma Aquila, apresenta a listagem de trabalhos finais de mestrado realizados por estudantes – ficheiro denominado de “Master”;
- Ficheiro Excel, proveniente da plataforma Aquila, apresenta a listagem de trabalhos finais de doutoramento realizados por estudantes – ficheiro denominado de “PhD”;
- Ficheiro Excel, proveniente da plataforma Aquila, apresenta a listagem das unidades curriculares lecionadas no ano letivo de 2019/2020 – ficheiro denominado de “Course Units”.

Todos os ficheiros apresentam um conjunto de variáveis comuns, mais concretamente, o título, o tipo de atividade, o ano, o resumo/descrição, e caso disponível o curso, o grau académico, o/s autor/s, entre outras variáveis.

3.3. CONJUNTO DE KEYWORDS

Para realizar o mapeamento das atividades de ensino e de investigação relacionadas com os ODS, foi selecionado um conjunto de *keywords* alusivo a cada um dos ODS. Depois de investigadas as listas de *keywords* utilizadas para este tipo de trabalho em universidades estrangeiras, foram identificadas três listas. Foram analisadas e testadas essas três listas de *keywords* provenientes de diferentes fontes, nomeadamente:

- Monash – compilação realizado pela Universidade de Monash e pela SDSN Australia/Pacific, engloba os 17 ODS (Monash University and SDSN Australia/Pacific, 2017);
- Confraria – compilação ainda não finalizada e publicada, engloba os 16 primeiros ODS (Confraria, 2020);
- Siris – realizado pela consultora SIRIS Academic, engloba os 16 primeiros ODS (Duran-Silva et al., 2019).

Cada um destes conjuntos de *keywords* apresentam um grupo de palavras base, com os seus respetivos sinónimos e palavras derivadas, relacionados com cada um dos ODS.

Como termo de comparação, de forma a definir o conjunto definitivo a utilizar no artefacto os diferentes conjuntos de *keywords* foram aplicados no algoritmo de classificação (subcapítulo 3.6.) e analisados (Anexo III). Como consequência, foi selecionado o conjunto de *keywords* realizado pelo Monash, que se verificou ser mais abrangente e completo relativamente aos outros, visto ter em consideração os 17 ODS e apesar de ter um número inferior de *keywords*, e sendo estas mais generalistas relativamente aos outros conjuntos, é constituído por palavras simples e compostas que permitem em futuros processos iterativos ter uma melhor perceção para restringir e/ou generalizar as *keywords*. Para além disso, foi selecionado também por ser um conjunto já com credibilidade na sua utilização, dado ser uma lista utilizada por outras universidades (SDSN Australia/Pacific, 2017). Apenas foi selecionado este conjunto de *keywords* em vez de uma junção dos três, devido à pretensão de primeiramente ter um conhecimento robusto sobre a temática para depois fazer melhorias no conjunto de acordo com as necessidades da instituição.

Este conjunto de *keywords* do Monash sofreu um processo de natureza iterativa, com vista a apurar a sua fiabilidade. As alterações realizadas focaram-se em excluir e em restringir algumas *keywords* por serem muito generalistas (atividades não relacionadas com os ODS foram identificadas pelo algoritmo) e em garantir a tradução correta das *keywords* para português. O conjunto de *keywords* final utilizado no projeto apresenta 1702 palavras, expressas em português e em inglês.

TABELA I

Exemplos de *keywords* provenientes do conjunto utilizado

Keywords		ODS
<i>Sustainable agriculture</i>	Agricultura sustentável	2
<i>Global education</i>	Educação global	4
<i>Clean technologies</i>	Tecnologias limpas	9
<i>Marine biodiversity</i>	Biodiversidade marinha	14
<i>Entrepreneurship</i>	Empreendedorismo	8 e 17
<i>Human rights</i>	Direitos humanos	5, 10 e 16

3.4. AMBIENTE DE DESENVOLVIMENTO

A linguagem de programação utilizada no tratamento dos dados, assim como na edificação do algoritmo foi o Python. Python é uma linguagem de programação *high-level*, interpretada e orientada a objetos criada em 1989 por Guido Van Rossum (Python, 2020). O Python apresenta um conjunto de funcionalidades conhecidas por bibliotecas, que suportam uma diversidade de tarefas de programação (Lutz, 2007). A linguagem Python é simples e fácil de aprender e entender, o que levou a um aumento substancial da sua popularidade e da sua utilização ao longo dos anos (Malik, 2019a). As razões que levaram à seleção desta linguagem foram duas: foi a linguagem lecionada durante o mestrado em que este projeto se insere, pelo que existe familiaridade na sua utilização e, por ser fácil de perceber e estruturar algoritmos para analisar grandes conjunto de dados (Kruglyk & Lvov, 2012) .

O *software* utilizado foi a ferramenta Anaconda. Anaconda é uma plataforma de *open-source* mais focada para *data science* e *machine learning*, que utiliza a linguagem de programação R e Python, trazendo consigo um conjunto de pacotes e bibliotecas *open-source* (Anaconda, 2020). Esta decisão teve por base o facto de ser um *software* cuja instalação e licenciamento é simples e gratuito, por apresentar uma interface *user-friendly* e pela possibilidade de utilizar a linguagem de programação Python.

Neste contexto é utilizado a *framework* Jupyter da Anaconda e as bibliotecas utilizadas são: Pandas, Numpy e FlashText. Pandas é uma biblioteca de análise/manipulação de dados que cria *dataframes* (objeto com linhas e colunas) num formato semelhante a um *software* estatístico (Excel e SPSS) (Bronshstein, 2017; Pandas, 2020). NumPy é uma biblioteca para computação científica, que suporta *arrays* e matrizes multidimensionais e que contém um grande número de funções matemáticas (Malik, 2019b; NumPy, 2020). FashText é uma biblioteca para procurar e substituir *keywords* em textos, que se baseia no algoritmo de *Aho-Corasicik* e de *Trie Dictionary* (Singh, 2017).

A biblioteca FashText apenas identifica *keywords* completas compreendidas entre os caracteres de limite, $\langle start \rangle keyword \langle eot \rangle$, em que o *start* e o *eot* representam um espaço e sinais de pontuação (Singh, 2017; NSS, 2017). A título exemplificativo, num dicionário, caso a *keyword* seja {Apple}, o algoritmo não vai identifica a *keyword* na frase “*I like Pieanapple*”, uma vez que a *keyword* está contida numa palavra e deste modo não há

correspondência relativamente ao *start*. Tem a particularidade de ser possível seleccionar o parâmetro *case sensitive*, deste modo como não foi seleccionado este parâmetro, a biblioteca vai tratar de igual modo as letras maiúsculas e minúsculas, {Apple} e {apple} são consideradas *keywords* iguais. Outra biblioteca bastante utilizada para o mesmo propósito é o Regex, que é bastante versátil e útil para identificar padrões como também para substituir palavras. Apesar de ambas as bibliotecas realizarem a função pretendida para a identificação das *keywords*, foi seleccionada a biblioteca FashText, visto ter uma melhor performance a nível de rapidez, como se pode verificar na Figura 6, com o aumento do número de *keyword* o tempo de processamento através do Regex aumenta de forma quase linear, enquanto o tempo de processamento do FashText mantém-se constante. Sendo assim, o FlashText é bastante vantajoso quando o número de *keywords* é elevado, dado que estas podem ser identificadas apenas com uma passagem pelo texto seleccionado (Singh, 2017).

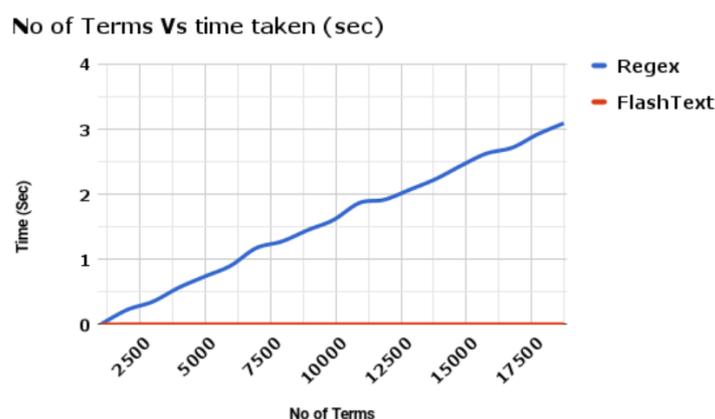


FIGURA 6 - Comparação do tempo de processamento entre Regex e o FlashText (Singh, 2017)

Para a implementação dos elementos visuais do projeto, foi utilizado o Microsoft Power BI. O *software* Power BI Desktop permite a integração de dados de diferentes fontes, a preparação, a análise e a visualização de dados, assim como a criação de relatórios de forma interativa que podem ser partilhados com outros utilizadores da organização (Microsoft, 2020).

A Gartner reconheceu, em 2020, a ferramenta Microsoft Power BI como líder de mercado no *Magic Quadrant* (Figura 7) no relatório de “Plataformas de *Business Intelligence* e Análise de Dados”. No relatório as principais vantagens identificadas incidem sobre a sua inclusão no Office 365, que permite aumentar o seu alcance no

possibilidade de os juntar numa fase posterior. Todos os ficheiros extraídos dos seus respetivos SI vão apresentar para este projeto apenas uma compilação de dados compreendidos entre 2011 até 2020, período de tempo estabelecido pelo grupo de trabalho dos ODS. O ano 2020 apesar de não estar finalizado, houve o interesse de utilizar os dados mais recentes disponíveis.

3.6. ALGORITMO DE CLASSIFICAÇÃO

Identificadas as *keywords* e realizado o tratamento dos dados, segue-se a classificação do/s ODS a que cada atividade presente nos ficheiros corresponde. Para facilitar este processo de identificação foi edificado um algoritmo, este vai ser denominado de algoritmo de classificação (Anexo IV). Um algoritmo engloba uma sequência de instruções explícitas com o intuito de resolver um problema e obter o *output* desejado para qualquer *input* válido (Levitin, 2012).

A utilização de um algoritmo no processo de identificação de *keywords* nas variáveis, revelou ser uma mais-valia, na medida em que permitiu rapidez no processo de identificação e na análise da informação. Cada ficheiro vai ter um conjunto específico de variáveis, onde vão ser procuradas as *keywords*, na Tabela 2 são identificadas as variáveis utilizadas no algoritmo. Esta seleção foi realizada em conjunto com o ISEG, foram selecionadas as variáveis que permitem ter uma noção transversal e extrair informação sobre o domínio a que corresponde cada atividade.

TABELA II

Variáveis utilizadas no algoritmo de classificação

	<i>Scopus</i>	<i>Master</i>	<i>PhD</i>	<i>Course Units</i>
<i>Title</i>	X	X	X	X
<i>Abstract</i>	X			
<i>Abstract_PT</i>		X	X	
<i>Abstract_EN</i>		X	X	
<i>Keywords</i>	X		X	
<i>Keywords_PT</i>		X		
<i>Keywords_EN</i>		X		
<i>Author_Keywords</i>	X			
<i>Objective_PT</i>				X
<i>Objective_EN</i>				X
<i>Pragmatic_Lines_PT</i>				X
<i>Pragmatic_Lines_EN</i>				X

Tendo em consideração o dicionário de *keywords* do Monash, o algoritmo vai percorrer o texto de todas as linhas das variáveis identificadas, e caso ocorra a identificação de uma *keyword* no texto, vai associar a atividade ao/s ODS a que correspondente a *keyword*. O algoritmo guarda essa informação e de seguida vai continuar a percorrer o resto das linhas das variáveis até à última linha da última variável considerada, repetindo o processo caso ocorra a correspondência de outras *keywords*, sendo possível durante esta iteração haver inúmeras ou nenhuma correspondência.

Cada atividade pode estar classificada a nenhum, a um ou a vários ODS. A título exemplificativo, um trabalho final de curso pode estar simultaneamente classificado no ODS 4 (Educação de Qualidade) e no ODS 9 (Indústria, Inovação e Infraestruturas), devido a conter *keywords* de ambos os ODS. Caso se dê a identificação de várias *keywords* para o mesmo ODS, este vai apenas contabilizar uma vez, independentemente do número de *keywords* identificadas para o mesmo ODS.

3.7. MODELO DIMENSIONAL

O *design* do DW utilizado no Power BI, no separador *Model View*, é o *snowflake schema*, que consiste num conjunto de tabelas de fato, associadas a uma ou várias tabelas de dimensão, que por sua vez estão associadas a uma ou várias tabelas de dimensão. Este modelo foi selecionado devido ao interesse de ter tabelas de dimensões dentro de tabelas de dimensões. Ou seja, neste projeto específico, existe o interesse de ter tabelas de dimensões a relacionarem-se entre si, que normalizam as tabelas de dimensões. Apesar deste modelo ter uma navegação mais complexa que o modelo em estrela (Kimball & Ross, 2013), foi dada prevalência à não repetição das informações ao logo das tabelas de dimensões principais ao contrário da rapidez de processamento, uma vez que os dados vão sofrer atualizações anuais. Adicionalmente foi selecionado pela facilidade que demonstra na criação de visualizações e filtros nos *dashboards*.

As fontes de dados identificadas no subcapítulo 3.2. correspondem às tabelas de fatos, a tabela “Master” e “PhD” vão ser agrupadas em apenas uma tabela fato “Thesis”, devido a terem a mesma granularidade, ambas consistem em trabalhos finais de curso. O modelo vai ter na sua totalidade três tabelas de fato, cada uma das tabelas refere-se a atividades diferentes, deste modo devem estar em tabelas de fato diferentes (Kimball & Ross, 2013).

As tabelas de fato são as seguintes:

- FactThesis – apresenta os trabalhos finais de mestrado e de doutoramento, por curso e por data;
- FactCourseUnits – apresenta as unidades curriculares lecionadas, por grau académico e por curso;
- FactScopus – apresenta os artigos científicos, por data.

Adicionalmente, foi criado quatro tabelas de dimensão que englobam atributos únicos que descrevem os dados provenientes das tabelas de fato, mais concretamente:

- DimSDG – apresenta os 17 ODS com as suas respetivas descrições;
- DimGroup – apresenta as cinco áreas, ao quais os ODS pertencem, com as suas respetivas descrições;
- DimCourse – identifica todos os cursos lecionados no ISEG, identificados pelo seu nome e pelas suas siglas, por grau académico;
- DimType – apresenta a lista de todas as atividades consideradas no mapeamento (artigos, teses, projetos, estágios e unidades curriculares);
- DimDate – apresenta uma lista de todos os dias compreendidos deste 1/1/2011 até à data atual, agrupados por trimestre, semestre, mês e ano.

A Figura 8 apresenta o esquema em *snowflake* utilizado, adicionalmente no Anexo V, está identificado a relação entre as tabelas de fato e de dimensão.

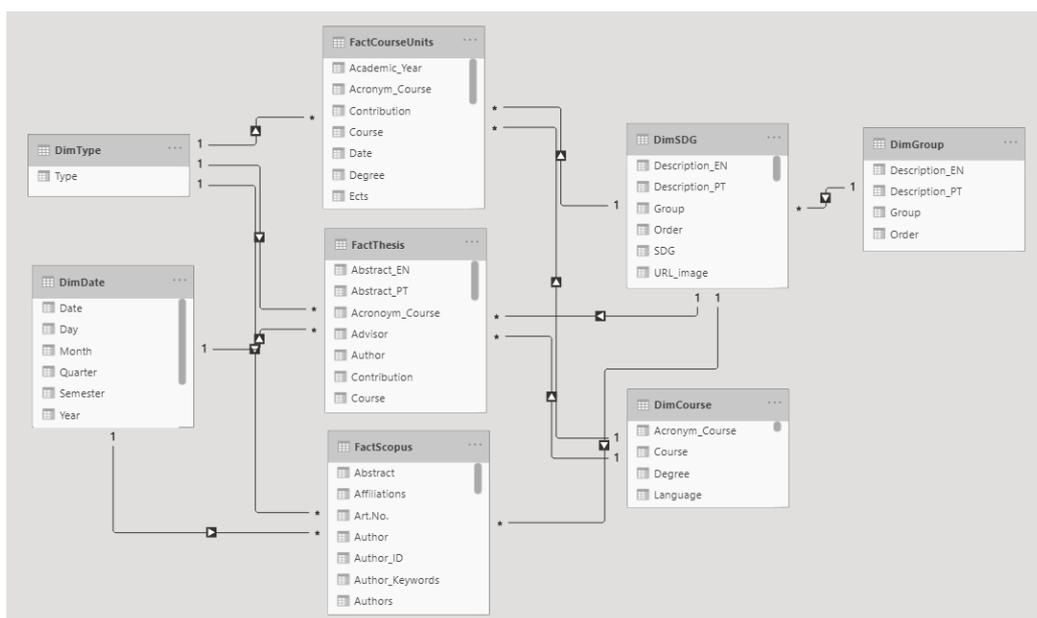


FIGURA 8 - Modelo dimensional de dados

3.8. VISUALIZAÇÃO DOS DADOS

A última etapa materializa-se na conceção dos *dashboards* no Power BI. Os *dashboards* foram criados tendo em conta as diretivas fornecidas pelo ISEG e os trabalhos realizados por outras IES. Foram edificados *dashboards* tendo em consideração as necessidades dos diferentes *end-users*. Estes vão apresentar uma visão global das contribuições e uma análise mais aprofundada sobre cada uma das atividades.

No desenvolvimento dos *dashboards*, foi realizado um conjunto de métricas utilizando a linguagem DAX (*Data Analysis Expressions*) do Power BI, que são armazenadas nas respetivas tabelas a que correspondem (Kimball & Ross, 2013). As métricas criadas permitem ter uma visão sobre os ODS para os quais a instituição mais contribui e as contribuições realizadas por cada atividade, a nível numérico e percentual, e ainda a média de contribuições por ano, a quantidade de contribuições do ano anterior e a sua variação relativamente ao período homólogo (Anexo VI). Foram aplicados vários tipos de filtros para a obtenção de informações mais detalhadas, que compreendem a identificação do intervalo temporal, seleção dos ODS, tipo de atividade, grupo de ODS, curso, grau académico e o autor/orientador. Estes foram colocados no lado esquerdo nos *dashboards* e variam consoante a informação de cada um.

Foram criados doze *dashboards*. Existiu a criação de alguns *dashboards* alternativos dada a não seleção definitiva por parte do ISEG dos *dashboards* que pretende usar. Importa referir que as unidades curriculares não apresentam uma data de realização, por se tratar de uma atividade que não varia com a mudança do ano letivo. Deste modo a atividade apenas é considerada nos *dashboards* das Figuras 9 e 10 e Anexos VII e XI.

A Figura 9 apresenta a proposta do *dashboard* da visão geral das contribuições de cada um dos ficheiros que se considerou no mapeamento, neste caso os valores numéricos e percentuais dos artigos científicos, das unidades curriculares e dos trabalhos finais de curso que contribuem para os ODS. O *dashboard* tem o intuito de mostrar o impacto que cada uma das atividades tem nos ODS e o respetivo número de contribuições. Desta forma, é possível observar que as atividades que mais contribuem para os ODS são os

trabalhos de final de curso, a nível numérico são os mestrados e a nível percentual são os doutoramentos, sendo que todas as atividades contribuem para os ODS em mais de 70%.

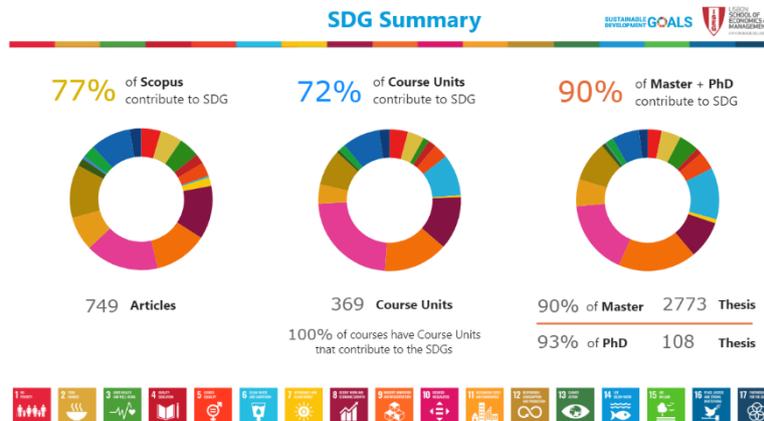


FIGURA 9 - Dashboard: SDG Summary

A Figura 10, consiste também numa visão geral das contribuições, mas com a possibilidade de aplicar filtros de forma a ser mais dinâmica a interação do utilizador. O *dashboard* tem o intuito de identificar os ODS para quais o ISEG mais contribui. Permite visualizar o número de contribuições realizadas em cada ODS, tendo em consideração todo o tipo de atividades realizadas pela instituição, juntamente com um conjunto de cartões que identificam a quantidade de contribuições realizadas em cada atividade e o respetivo total de contribuições. Verifica-se uma maior contribuição nos ODS 9 e 10. Numa mesma perspetiva, o *dashboard* do Anexo VII, apresenta uma visão geral do impacto que cada um dos ODS tem nas contribuições realizadas, em valores numéricos e percentuais, e a respetiva quantidade de contribuições realizadas por cada atividade em cada um dos ODS. Este tem o mesmo intuito que o *dashboard* da Figura 10, mas com informações mais detalhadas sobre cada um dos ODS.

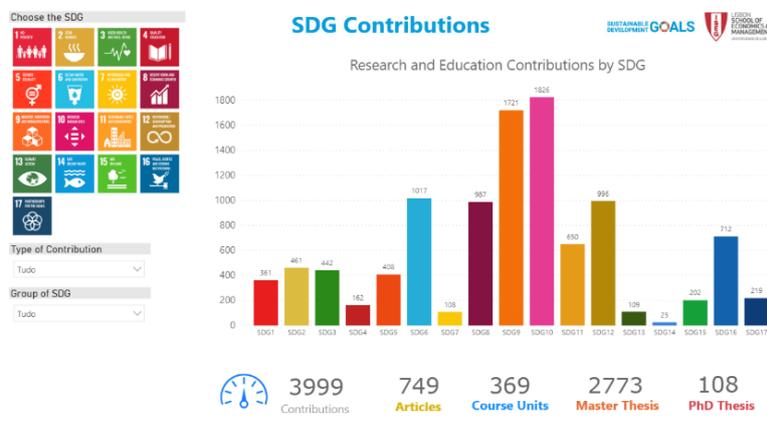


FIGURA 10 - Dashboard: SDG Contributions

O *dashboard* da Figura 11, apenas tem em consideração os trabalhos finais e os artigos Scopus, devido à intenção de utilizar filtros e indicadores temporais, sendo assim foram excluídas as unidades curriculares. O *dashboard* permite analisar a evolução do número de contribuições realizadas ao longo dos anos, por valor total e por cada ODS. Importa referir que o somatório das contribuições de cada ODS não vai corresponder ao número total de contribuições, uma vez que uma atividade pode estar classificada em um ou mais ODS. Como o ano 2020 não está finalizado, verifica-se um número inferior de contribuições relativamente aos anos anteriores. Como alternativa o Anexo VIII, apresenta um *dashboards* que apresenta a junção da Figura 10 e 11, tendo em consideração apenas os trabalhos finais e artigos Scopus.

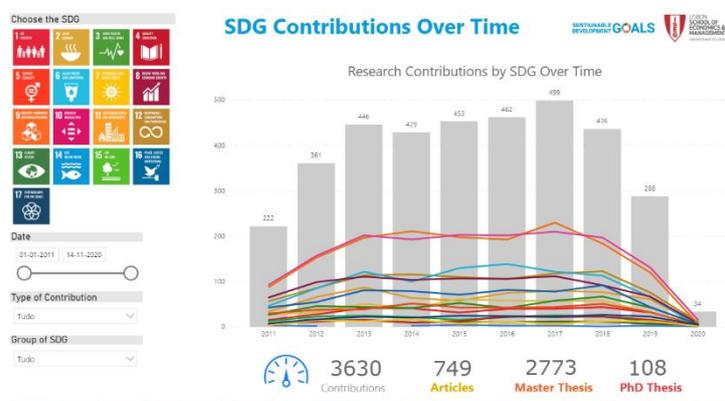


FIGURA 11 - *Dashboard: SDG Contributions Over Time*

Numa outra perspetiva, a Figura 12, apresenta o mesmo conteúdo que o *dashboard* da Figura 11, mas com os valores normalizados. Este *dashboard* tem o intuito de analisar as contribuições ao longo do tempo, e assim ter uma noção em termos percentuais dos ODS que contribuiriam mais ou menos em cada ano. Verifica-se que os valores são constantes ao longo do tempo.

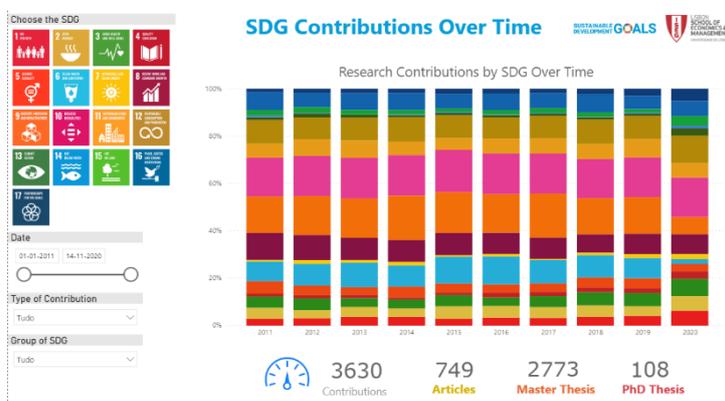


FIGURA 12 - *Dashboard: SDG Contributions Over Time*

A Figura 13 apresenta uma análise mais detalhada sobre cada uma das áreas a que pertencem os ODS. Apresenta um gráfico circular onde é possível analisar as proporções de cada uma das áreas e identificar o seu valor numérico e percentual. Adicionalmente apresenta um gráfico de barras com o número de contribuições realizadas em cada área dos ODS. Este *dashboard* tem como objetivo mostrar as áreas dos ODS para os quais o ISEG contribui e as suas proporções. Conclui-se que existe uma maior contribuição na área *Prosperity* (engloba os ODS com maior número de contribuições identificados na Figura 10). O Anexo IX, apresenta um *dashboard* alternativo com o mesmo conteúdo.

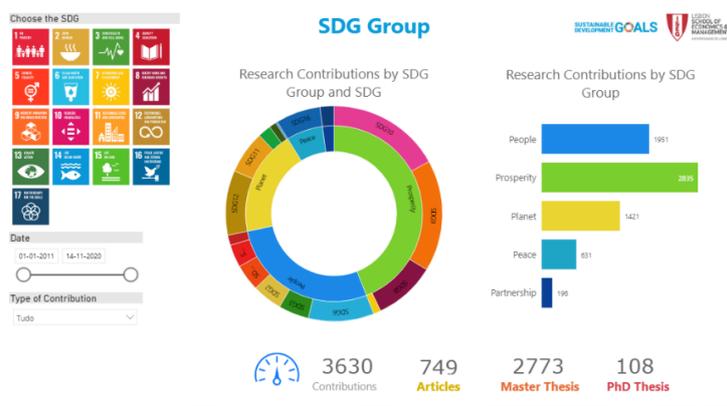


FIGURA 13 - *Dashboard: SDG Group Contributions*

O *dashboard* da Figura 14, é direcionado para os trabalhos finais de curso. Apresenta duas métricas posicionadas no centro, a percentagem de trabalhos finais que contribuem para os ODS e o número de contribuições realizadas. Apresenta também dois gráficos de barras, o *Course with SDG Theses* mostra os cursos das respetivas contribuições realizadas ao fazer o *drilldown* do grau académico e o gráfico *Top SDG Contributions* mostra os ODS ordenados por número de contribuições, e por último uma tabela que mostra informações mais detalhadas sobre os trabalhos finais. Com a análise da *dashoard* é possível ter uma perceção geral das contribuições realizadas pelos trabalhos finais, por cursos de mestrado e de doutoramento, orientador, data e por ODS. O Anexo X e XI, correspondem respetivamente aos *dashboards* direcionadas para os artigos Scopus e as unidades curriculares, que estão na mesma linha de construção e apresentação do *dashboard Master and PhD Thesis*.

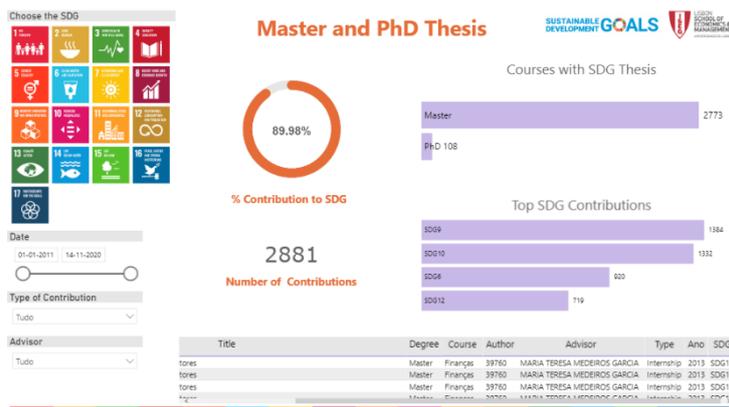


FIGURA 14 - *Dashboard: Master and PhD Thesis*

O *dashboard SDG Contributions Indicators*, na Figura 15, é projetado exclusivamente para a Presidência do ISEG, numa perspetiva mais estratégica. Apresenta cinco cartões, mais concretamente o número de contribuições por cada atividade, o respetivo somatório de todas as contribuições e a média de contribuições por ano. O *dashboard* apresentam também três gráficos de barras: *Research Contributions Over Time* - evolução temporal do número de contribuições face à média de contribuições; *Contributions by year-on-year* - qualidade de contribuições realizadas face ao ano homólogo; *Top SDG Contributions*- ODS ordenados de acordo com o número de contribuições. Este *dashboard* permite analisar a evolução temporal do número de contribuições e as suas respetivas quantidades totais por ODS e atividades, com o intuito de fornecer informações que permitam analisar o compromisso e a performance do ISEG com os ODS.

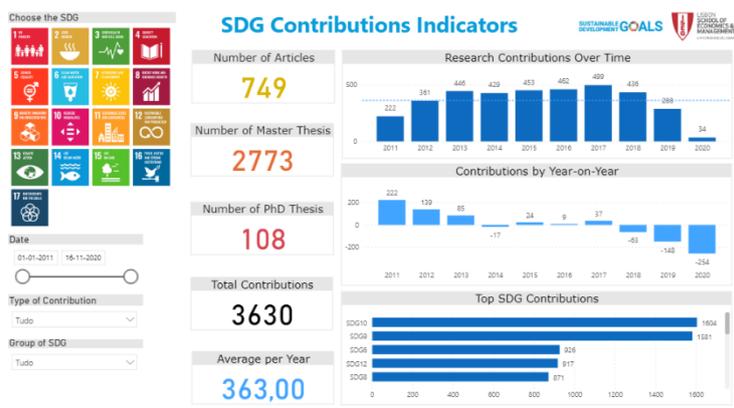


FIGURA 15 - *Dashboard: SDG Contributions Indicators*

4. CONCLUSÕES

O presente projeto surgiu com a necessidade de reforçar o posicionamento do ISEG enquanto universidade sustentável, que contribui para a implementação dos ODS. Assim, o projeto propôs-se a realizar o mapeamento e a edificação de um sistema de BI para a identificação da temática das atividades de ensino e de investigação realizadas pelo ISEG segundo a matriz dos ODS. Ambos os objetivos e as suas respetivas fases foram realizados de forma maioritariamente sequencial e contribuíram de forma significativa para a edificação do artefacto final. Com este projeto foi possível consolidar a formação académica e familiarizar-me com novas plataformas e linguagens.

O primeiro objetivo, o mapeamento das atividades de investigação e de ensino, foi crucial para transformar os dados dos ficheiros em informações relevantes a serem apresentados e analisados nos *dashboards*. Consistiu na seleção do conjunto de *keywords*, no tratamento dos dados e no desenvolvimento e implementação do algoritmo de classificação. Estas etapas foram as mais demoradas do projeto, uma vez que sofreram um conjunto de alterações ao longo do mesmo, dada a sua importância para a qualidade dos dados apresentados nos *dashboards*. O algoritmo de classificação proporciona uma identificação eficaz e eficiente das *keywords*, com base nas validações manuais realizadas. Chegou-se à conclusão que apesar de o conjunto de *keywords* Monash ser utilizado por inúmeras universidades o que lhe confere credibilidade, é passível de tratamento de modo a melhorar a identificação dos ODS, uma vez que algumas atividades de investigação e ensino apesar de não estarem relacionadas com os ODS são identificadas devido a *keywords* bastantes gerais.

O segundo objetivo, que consiste na visualização dos dados, responde às necessidades da instituição. Esta etapa abrange a criação do modelo dimensional e a apresentação dos resultados em *dashboards*, de forma a disponibilizar a informação pretendida de forma simples e dinâmica. Os *dashboards* sempre que possível pretenderam alinhar-se com as regras de boas práticas de desenho de *dashboards*, à exceção da apresentação de todo o conteúdo apenas numa página e da utilização de poucas cores, que são desrespeitadas em alguns *dashboards*. Estas foram idealizadas especialmente para os *end-users* que acedem ao *websites* da instituição, que por sua vez podem ou não ter conhecimento sobre os ODS, deste modo valorizou-se a simplicidade e a pouca quantidade de informação, de forma a

ser perceptível por todos. A utilização de diversas cores deve-se à intenção de ter os conteúdos dos gráficos a corresponderem às respetivas cores dos ODS. Este processo, pode ser melhorado com a integração de novas tabelas de fato/dimensão e com a reavaliação das necessidades de informação pretendidas.

A seleção do conjunto de *keywords* e o desenho dos *dashboards* foi concebido com base nos objetivos e nas orientações da instituição. Relativamente ao artefacto final, este permitiu centralizar a informação e informar o ISEG sobre as suas contribuições para os ODS de uma forma simples e dinâmica. Existe a possibilidade de adicionar novos ficheiros para identificar os ODS e a liberdade de explorar os dados, obter informações e comparações, através da utilização de filtros. Os *dashboards* apresentam os indicadores mais relevantes de forma a permitir ao ISEG ver como tem de atuar a nível estratégico, assim como o número de contribuições realizadas para a Agenda 2030 e a sua evolução ao longo do tempo. Desde modo, o projeto representa uma robusta contribuição para as universidades que pretendem iniciar a sua contribuição para os ODS, este pode ser replicado por outras universidades ou organizações.

5. LIMITAÇÕES E RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Um projeto deste âmbito necessita de um trabalho contínuo de forma a ter um resultado final mais eficaz. Uma das limitações que este projeto teve foi a falta de um maior envolvimento por parte dos elementos da instituição na identificação das *keywords* a utilizar no processo e na reflexão robusta e clara sobre o que é pretendido apresentar nos *dashboards* finais do projeto, de modo a estar alinhado com a estratégia do ISEG. Outra limitação que o projeto tem e que pode comprometer a precisão das informações apresentadas nos *dashboards*, está relacionado com o conjunto de *keywords* utilizado. Este apresenta *keywords* bastante generalizadas, que precisam de sofrer um processo iterativo a realizar ao longo do tempo, deste modo o conjunto de *keywords* não está finalizado. No entanto, foi selecionado o conjunto de *keywords* com maior credibilidade. Para além disso, outra limitação prende-se no fato de os ficheiros terem alguns campos em falta e não haver coerência nas variáveis comuns aos diferentes ficheiros, deste modo a qualidade dos dados constituiu uma barreira.

Uma limitação encontrada no algoritmo de classificação, foi a identificação de *keywords* quando o <start> e o <eot> (caracteres de limite da *keyword*) apresentam um

acentos, esta questão ocorre devido à inexistência de acentuação gráfica no dicionário inglês. A título exemplificativo, a *keyword* {mar} é identificada em textos que contenham a palavra “maré”.

O projeto desenvolvido, embora totalmente funcional na identificação da temática de cada um dos conteúdos dos ficheiros e na apresentação dos *dashboards*, tem um conjunto de funcionalidades a melhorar e a acrescentar no futuro, para tornarem o processo e os *dashboards* mais eficazes. É de destacar as seguintes recomendações:

- Integração de novas fontes de dados, de forma a contemplar todas as atividades da instituição e a possibilidade de criar novas análises e conhecimentos relevantes para a instituição, como acrescentar a dimensão “autores” juntamente com a respetiva filiação, adicionar novos indicadores, KPI’s, entre outras;
- Apresentar uma solução integrada, visto ser importante ter um sistema de BI completo que integre todas as etapas. Atualmente o tratamento dos dados e o algoritmo é realizado para cada um dos ficheiros em separado de forma manual e não existe nenhum DW para colocar a informação. Deste modo é importante ter um sistema integrado entre si para que a atualização dos dados seja facilmente realizada, de forma a aprimorar a rapidez e a facilidade da análise;
- Conceber uma nova categoria de ODS no mapeamento, que esteja relacionado com as atividades que contribuam para todos os ODS num todo, como é o caso deste projeto em concreto. Este projeto enquadra-se no tema dos ODS, mas não existe nenhuma categoria específica para a sua classificação;
- Manutenção das *keywords*. Proceder ao tratamento, mais concretamente à alteração das *keywords* simples para compostas de forma a eliminar falsos positivos e na adição de novas *keywords* e as suas variantes, com o intuito de potenciar a identificação dos ODS.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Albareda-Tiana, S., Vidal-Raméntol, S. & Fernández-Morilla, M. (2018). Implementing the sustainable development goals at University level. *International Journal of Sustainability in Higher Education* 19 (3), 473–497. <https://doi.org/10.1108/IJSHE-05-2017-0069>
- Anaconda (2020). *Your data science toolkit* [Em linha]. Disponível em: <https://www.anaconda.com/products/individual> [Acesso em: 2020/08/22].
- Aurora (2018). *SDG analysis: Bibliometrics of Relevance* [Em linha]. Disponível em: https://aurora-network.global/project/sdg-analysis-bibliometrics-relevance/?fbclid=IwAR0Z52AwCOGj_vEGVia9XLcB_cQ-R5QqhjZuHA31NQIOmQdjKpXN4JUGZE [Acesso em: 2019/12/01].
- BCSD Portugal (2019). *Objetivos de Desenvolvimento Sustentável e o BCSD Portugal* [Em linha]. Disponível em: <https://www.ods.pt/> [Acesso em: 2019/12/01].
- Bronshstein, A. (2017). *A Quick Introduction to the “Pandas” Python Library*. Towards Data Science [Em linha]. Disponível em: <https://towardsdatascience.com/a-quick-introduction-to-the-pandas-python-library-f1b678f34673> [Acesso em: 2020/08/22].
- Chan, L., Chee, T., Chuah, M., Rahman, A., Tan, C., Wong, S. & Yeoh, W. (2009). Business Intelligence Systems: State-of-the-art review and contemporary applications. *Symposium on Progress in Information and Communication Technology* 2 (4), 96–101.
- Confraria, H. (2020). *Keywords Confraria* [Base de dados]. Documento não publicado.
- Duran-Silva, N., Fuster, E., Massucci, F. A. & Quinquillà, A. (2019). *A controlled vocabulary defining the semantic perimeter of Sustainable Development Goals* [Base de dados], dezembro 2019. Zenodo. Disponível em: <https://zenodo.org/record/3567769#.Xw3wEihKiUk>.
- El-Jardali, F., Ataya, N. & Fadlallah, R. (2018). Changing roles of universities in the era of SDGs: Rising up to the global challenge through institutionalising partnerships with governments and communities. *Health Research Policy and Systems* 16 (1), 1–5. <https://doi.org/10.1186/s12961-018-0318-9>
- Few, S. (2006). *Information dashboard design: The effective visual communication of data*. O'Reilly Media, Inc..
- Few, S. (2007). *Pervasive Hurdles to Effective Dashboard Design* [Em linha]. Disponível

em:

https://www.perceptualedge.com/articles/visual_business_intelligence/pervasive_hurdles_to_dd.pdf [Acesso em: 2020/09/12].

Filho, W. L. (2011). About the Role of Universities and their contribution to sustainable development. *Higher Education Policy* 24 (4), 427–438. <https://doi.org/10.1057/hep.2011.16>

Haupt, R., Scholtz, B. & Calitz, A. (2015). Using Business Intelligence to Support Strategic Sustainability Information Management. *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists*, 1–11. <http://dx.doi.org/10.1145/2815782.2815795>

Inmon, W. H. (2005). *Building the Data Warehouse*, 3ª ed. John Wiley & Sons, Inc..

ISEG (2020). *ISEG membro da Aliança ODS Portugal* [Em linha]. Disponível em: https://www.iseg.ulisboa.pt/aquila/publico/units/announcements.do?method=viewAnnouncement&unitID=102&announcementId=1843345&contentContextPath_PATH=/instituicao/ISEG/pagina-inicial&_request_checksum_=0468e18d599a0860c08f5b7593c9ebd47357fbf6 [Acesso em: 2020/08/21].

Junior, R. M., Fien, J. & Horne, R. (2019). Implementing the UN SDGs in Universities: Challenges, Opportunities, and Lessons Learned. *Sustainability: The Journal of Records* 12 (2), 129–133. <https://doi.org/10.1089/sus.2019.0004>

Kimball, R. & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3ª ed. Indianapolis: John Wiley & Sons, Inc..

Körffgen, A., Förster, K., Glatz, I., Maier, S., Becsi, B., Meyer, A., Kromp-Kolb, H. & Stötter, J. (2018). It's a Hit! Mapping Austrian research contributions to the sustainable development goals. *Sustainability* 10 (9), 1–13. <https://doi.org/10.3390/su10093295>

Kruglyk, V. & Lvov, M. (2012). Choosing the first educational programming language. *CEUR Workshop Proceedings*, 188–198.

Levitin, A. (2012). *Introduction to the design & analysis of algorithms*, 3ª ed. Pearson Education, Inc..

Lutz, M. (2007). *Learning Python*, 3ª ed. O'Reilly Media, Inc..

Malik, F. (2019a). *Everything About Python — Beginner To Advanced*. Medium [Em

- linha]. Disponível em: <https://medium.com/fintechexplained/everything-about-python-from-beginner-to-advance-level-227d52ef32d2> [Acesso em: 2020/08/23].
- Malik, F. (2019b). *Why Should We Use NumPy?*. Medium [Em linha]. Disponível em: <https://medium.com/fintechexplained/why-should-we-use-numpy-c14a4fb03ee9> [Acesso em: 2020/08/22].
- Malik, S. (2005). *Enterprise Dashboards*. Hoboken: John Wiley and Sons, Inc..
- Microsoft (2020). *What is Power BI Desktop?* [Em linha]. Disponível em: <https://docs.microsoft.com/pt-pt/power-bi/fundamentals/desktop-what-is-desktop> [Acesso em: 2020/08/15].
- Monash University and SDSN Australia/Pacific (2017). *Compiled Keywords for SDG Mapping* [Base de dados], maio 2017. Disponível em: http://ap-unsdsn.org/wp-content/uploads/2017/04/Compiled-Keywords-for-SDG-Mapping_Final_17-05-10.xlsx
- Moody, D. & Kortink, M. A. (2000). From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design. *Proceedings of the International Workshop on Design and Management of Data Warehouses*, 5–12.
- Niu, L., Lu, J. & Zhang, G. (2009). Business intelligence. In: *Cognition-Driven Decision Support for Business Intelligence*, Berlin: Springer, pp. 19–29.
- NSS (2017). *FlashText – A library faster than Regular Expressions for NLP tasks*. Analytics Vidhya [Em linha]. Disponível em: <https://www.analyticsvidhya.com/blog/2017/11/flashtext-a-library-faster-than-regular-expressions/> [Acesso em: 2020/09/19].
- NumPy (2020). *About Us* [Em linha]. Disponível em: <https://numpy.org/about/> [Acesso em: 2020/08/22].
- ODS (2020). *Os ODS representam as prioridades globais para a Agenda 2030 assinada por mais de 190 países* [Em linha]. Disponível em: <https://www.ods.pt/ods/> [Acesso em: 2020/08/15].
- Ong, I., Siew, P. & Wong, S. (2011). A Five-Layered Business Intelligence Architecture. *Communications of the IBIMA* 2011, 1–11. <https://doi.org/10.5171/2011.695619>
- Pandas (2020). *About Pandas* [Em linha]. Disponível em: <https://pandas.pydata.org/about/> [Acesso em: 2020/08/22].
- Petrini, M. & Pozzebon, M. (2009). Managing sustainability with the support of business

- intelligence: Integrating socio-environmental indicators and organisational context. *Journal of Strategic Information Systems* 18 (4), 178–191. <https://doi.org/10.1016/j.jsis.2009.06.001>
- Python (2020). *What is Python? Executive Summary* [Em linha]. Disponível em: <https://www.python.org/doc/essays/blurp/> [Acesso em: 2020/09/12].
- Relatório Voluntário Nacional (2017). *Relatório nacional sobre a implementação da Agenda 2030 para o Desenvolvimento Sustentável* [Em linha]. Disponível em: <https://www.cig.gov.pt/wp-content/uploads/2017/07/Portugal2017.pdf> [Acesso em: 2020/05/01]
- Richardson, J., Sallam, R., Schlegel, K., Kronz, A. & Sun, J. (2020). *Magic Quadrant for Analytics and Business Intelligence Platforms*. Gartner [Em linha]. Disponível em: https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb&ocid=mkto_eml_EM597235A1LA1[Acesso em: 2020/08/15].
- Scholtz, B., Calitz, A. & Haupt, R. (2018). A business intelligence framework for sustainability information management in higher education. *International Journal of Sustainability in Higher Education* 19 (2), 266–290. <https://doi.org/10.1108/IJSHE-06-2016-0118>
- Scipioni, A., Mazzi, A., Mason, M. & Manzardo, A. (2009). The Dashboard of Sustainability to measure the local urban sustainable development: The case study of Padua Municipality. *Ecological Indicators* 9 (2), 364–380. <https://doi.org/10.1016/j.ecolind.2008.05.002>
- SDSN Australia/Pacific (2017). *Getting started with the SDGS in Universities: A guide for universities, higher education institutions, and the academic sector*. [Em linha]. Disponível em: http://ap-unsdsn.org/wp-content/uploads/University-SDG-Guide_web.pdf [Acesso em: 2020/05/09]
- SDSN Australia/Pacific (2018). *Como Começar Com Os Ods Nas Universidades: Um guia para as universidades, os centros de educação superior e a academia* [Em linha]. Disponível em: http://ap-unsdsn.org/wp-content/uploads/Como-comecar-com-os-ODS-nas-Universidades_18-11-18.pdf [Acesso em: 2020/05/09]
- Singh, V. (2017). Replace or Retrieve Keywords In Documents at Scale. 1–10. <http://arxiv.org/abs/1711.00046>

- Su, H. N. & Lee, P. C. (2010). Mapping knowledge structure by keyword co-occurrence: A first look at journal papers in Technology Foresight. *Scientometrics* 85 (1), 65–79. <https://doi.org/10.1007/s11192-010-0259-8>
- Tilbury, D. (2012). Higher education for sustainability: a global overview of commitment and progress. In: Global University Network for Innovation (Eds.) *Higher Education in the World 4: Higher Education's Commitment to Sustainability: From Understanding to Action*, Hampshire: Palgrave Macmillan, pp.18-28
- UNESCO (2012). *Education for sustainable development: Sourcebook* [Em linha]. Disponível em: <https://sustainabledevelopment.un.org/content/documents/926unesco9.pdf> [Acesso em: 2020/09/26]
- UNESCO (2014). *Roadmap for Implementing the Global Action Programme on Education for Sustainable Development* [Em linha]. Disponível em: <https://sustainabledevelopment.un.org/content/documents/1674unescoroadmap.pdf> [Acesso em: 2020/09/26]
- UNRIC (2018). *Guia sobre Desenvolvimento Sustentável* [Em linha]. Disponível em: https://unric.org/pt/wp-content/uploads/sites/9/2019/01/SDG_brochure_PT-web.pdf [Acesso em: 2020/05/01]
- Vaughter, P. (2018). *Beacons, not Towers: How Higher Education Can Help Achieve the Sustainable Development Goals* [Em linha]. Disponível em: <https://prospernet.ias.unu.edu/wp-content/uploads/2018/01/UNU-IAS-PB-No13-2018.pdf> [Acesso em: 2020/08/29]
- Watson, H. J. (2009). Tutorial: Business Intelligence – Past, Present, and Future. *Communications of the Association for Information Systems* 25 (39), 487–510. <https://doi.org/10.17705/1CAIS.02539>
- Watson, H. J. & Wixom, B. H. (2007). The current state of business intelligence. *Computer* 40 (9), 96–99. <https://doi.org/10.1109/MC.2007.331>
- Wexler, S., Shaffer, J. & Cotgreave, A. (2017). *The big book of dashboards : visualizing your data using real-world business scenarios*. Hoboken: John Wiley & Sons, Inc.
- Wieder, B. & Ossimitz, M. L. (2015). The Impact of Business Intelligence on the Quality of Decision Making - A Mediation Model. *Procedia Computer Science* 64, 1163–1171. <https://doi.org/10.1016/j.procs.2015.08.599>

7. ANEXOS

Anexo I:

TABELA III - 17 Objetivos de Desenvolvimento Sustentável

ODS 1	Erradicar a pobreza em todas as suas formas em todos os lugares
ODS 2	Erradicar a fome, alcançar a segurança alimentar, melhorar a nutrição e promover a agricultura sustentável
ODS 3	Garantir o acesso à saúde de qualidade e promover o bem-estar para todos, em todas as idades
ODS 4	Garantir o acesso à educação inclusiva, de qualidade e equitativa, e promover oportunidades de aprendizagem ao longo da vida para todos
ODS 5	Alcançar a igualdade de género e empoderar todas as mulheres e raparigas
ODS 6	Garantir a disponibilidade e a gestão sustentável da água potável e do saneamento para todos
ODS 7	Garantir o acesso a fontes de energia fiáveis, sustentáveis e modernas para todos
ODS 8	Promover o crescimento económico inclusivo e sustentável, o emprego pleno e produtivo e o trabalho digno para todos
ODS 9	Construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação
ODS 10	Reduzir as desigualdades no interior dos países e entre países
ODS 11	Tornar as cidades e comunidades inclusivas, seguras, resilientes e sustentáveis
ODS 12	Garantir padrões de consumo e de produção sustentáveis
ODS 13	Adotar medidas urgentes para combater as alterações climáticas e os seus impactos
ODS 14	Conservar e usar de forma sustentável os oceanos, mares e os recursos marinhos para o desenvolvimento sustentável
ODS 15	Proteger, restaurar e promover o uso sustentável dos ecossistemas terrestres, gerir de forma sustentável as florestas, combater a desertificação, travar e reverter a degradação dos solos e travar a perda de biodiversidade
ODS 16	Promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos a construir instituições eficazes, responsáveis e inclusivas a todos os níveis
ODS 17	Reforçar os meios de implementação e revitalizar a parceria global para o desenvolvimento sustentável

Fonte: Adaptado de UNRIC (2018)

Anexo II:



FIGURA 17 - *Dashboard SDG Analysis: Bibliometrics of Relevance* (Aurora, 2018)

Anexo III:

TABELA IV - Comparação entre os três conjuntos de *keywords*

		Monash	Confraria	Siris
Número de <i>Keywords</i>		1702	1966	6857
Número de Contribuições	Scopus	749	494	675
	Master	2773	2318	2661
	PhD	108	93	114
	Course Units	369	786	304

Anexo IV:

```

# Import modules
import numpy as np
import pandas as pd
from flashtext import KeywordProcessor

# Create an object of KeywordProcessor
keywordProcessor=KeywordProcessor()

# Keyword dictionary for each of the 17 SGD
keywordDict = {
    'Specify the SDG': ['Specify the Keywords']}

# Add the Keyword dictionary
keywordProcessor.add_keywords_from_dict(keywordDict)

# Function to create columns based on a number of SDG
'''Initializes those columns with an initialValue (default: False). The column names are based on the SGD number (17 columns)'''
def initSGDS(df, size, initialValue=False):
    for sgd in range(1, size+1):
        df['SDG{}'.format(sgd)] = initialValue

# Create SGD Columns
initSGDS(dataFrame, 'Number of SDG')

# Function to search the keyword in the columns
def classifyPaper(dataFrame, searchColumns, keywordProcessor):
    '''Make a Copy of the Original DataFrame, because iterrows() doesn't have direct connection with original DataFrame'''
    df = dataFrame.copy()
    for index, row in df.iterrows():
        paperSGDS = []
        for column in searchColumns:
            # Find SGDS for that Column at row
            # Check if that row column isn't NaN
            if not pd.isna(row[column]): paperSGDS.append(kp.extract_keywords(row[column]))
        # Explode & Remove Duplicates
        paperSGDS = list(dict.fromkeys(sum(paperSGDS, [])))
        # If Row has SGDS, set those columns to True
        if paperSGDS: df.loc[index, paperSGDS] = True
    return df

# Identify the columns to search the keywords
searchColumns = ['Specify the Columns']

# Search the Keywords in the Columns
dataFrame = classifyPaper(dataFrame, searchColumns, keywordProcessor)

```

FIGURA 18 - Script em Python do Algoritmo de Classificação

Anexo V:

TABELA V - Relação entre as tabelas de fato e de dimensão

Tabela de Fato	Tabela de Dimensão	Atributo em comum
FactTheses	DimSDG DimCourse DimDate	SDG Date Acronym_Course
FactCourseUnits	DimSDG DimCourse	SDG Acronym_Course
FactScopus	DimSDG DimDate	SDG Date
	DimSDG DimGroupSDG	Group

Anexo VI:

TABELA VI - Métricas realizadas no Power BI

Métricas	Descrição
%CourseUnits %Courses %Scopus %Thesis %Master %PhD	Porcentagem das respetivas atividades identificadas nas métricas que contribuem para os ODS
CountCourseUnits CountScopus CountThesis CountMaster CountPhD	Contagem das respetivas atividades identificadas nas métricas que contribuem para os ODS
CountResearch	Somatório do número de teses de mestrado, teses de doutoramento e artigos Scopus que contribuem para os ODS
CountAllActivities	Somatório do número de teses de mestrado e doutoramento, artigos Scopus e unidades curriculares que contribuem para os ODS
Average	Média de contribuições por ano, tendo em conta o número de número de teses de mestrado, teses de doutoramento e artigos Scopus
LastYear	Quantidade de contribuições do ano anterior
PeriodDifference	Quantidade de contribuições relativamente ao período homólogo
%SDG*	Porcentagem de contribuições realizadas em cada ODS, tendo em consideração todas as atividades de investigação e ensino *Nota: a métrica foi realizada para cada um dos 17 ODS
CountSDG*	Número de contribuições realizadas em cada ODS, tendo em consideração todas as atividades de investigação e ensino *Nota: a métrica foi realizada para cada um dos 17 ODS
CountRSDG*	Número de contribuições realizadas em cada ODS, tendo em consideração apenas as teses de mestrado, teses de doutoramento e artigos Scopus *Nota: a métrica foi realizada para cada um dos 17 ODS

Anexo VII:

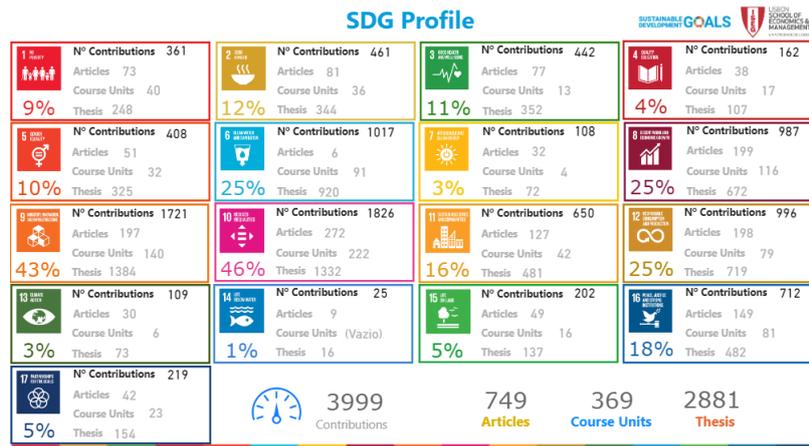


FIGURA 19 - Dashboard: SDG Profile

Anexo VIII:

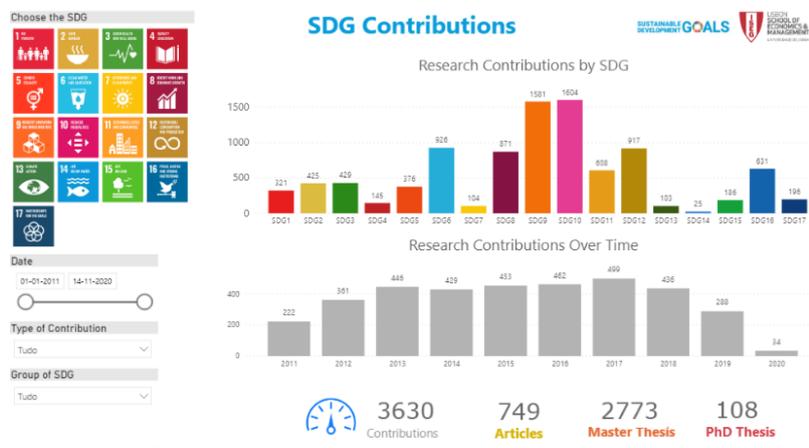


FIGURA 20 - Dashboard: SDG Contributions

Anexo IX:

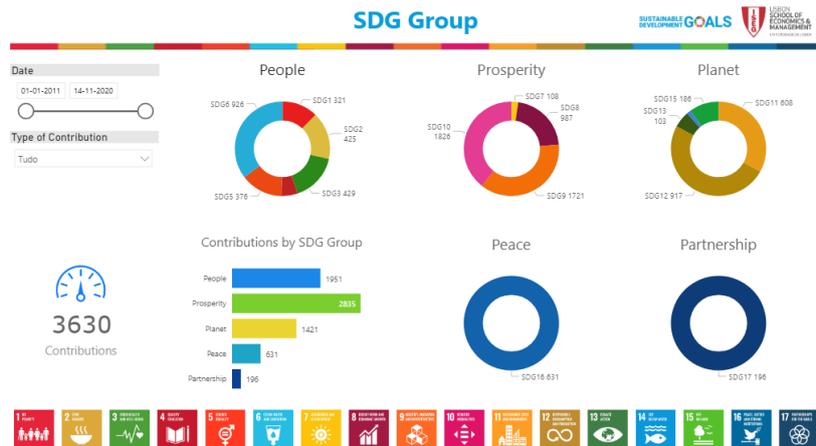


FIGURA 21 - Dashboard: SDG Group

Anexo X:

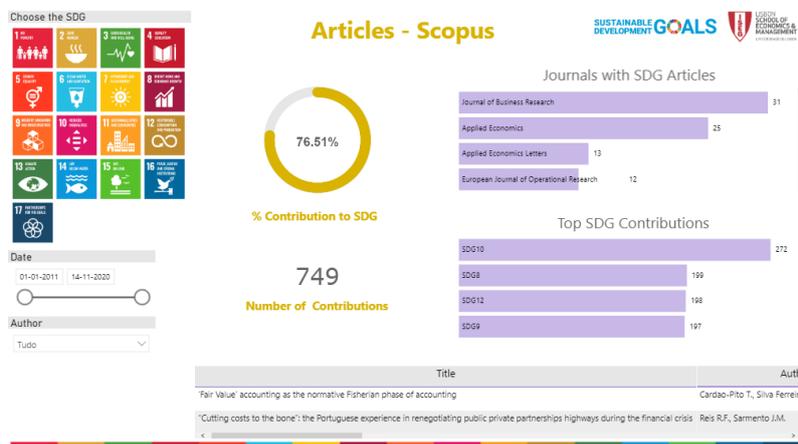


FIGURA 22 - Dashboard: Articles - Scopus

Anexo XI:

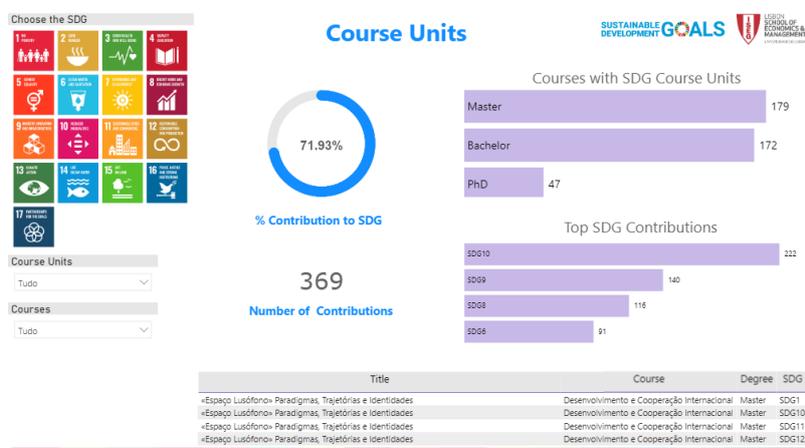


FIGURA 23 - Dashboard: Course Units