The Potential for a Data Science and Natural Language Processing Approach

- A meta-evaluation case study within the field of International Development Cooperation

Background

The potential as well as the value of receiving insights from alternative evaluation methods - besides the more traditional such as interviews, surveys and focus groups - are believed to increase with the intensification of digitalisation as well as increased social engagement on the internet. The amount of traffic and hence available data from alternative data sources, such as online and mobile usage, have reached close to unimaginable proportions and are expected to grow exponentially in the foreseeable future. Recent estimates suggest that the global annual internet traffic will increase with more than 200 percent between 2017 (1,5 Zettabytes) and 2020 (4,8).[1] The described process or phenomenon has a direct effect on the practice of evaluations, not least due to its heavy dependence on data to get insights. Questions have been raised relating to the necessity of a broadened analytical toolbox for evaluations if increases in volume, velocity and variety of data should be handled and taken full advantage of. We believe this to be the case, particularly within the field of international development cooperation, where paucity of data often hampers the quality and scope of evaluations.

In this light, we would like to design a study that aims to probe the possibility for using data science applications for meta evaluations, and put it's potential to the test when it comes to processing and analysing narrative written evaluations. Our proposed approach will involve a wide range of analytical methods with a heavy emphasis on computer-based processing of human language, or so called Natural Language Processing (NLP)[2].

NLP methods, developed within the field of computational linguistics, have grown increasingly popular during recent years due to their applicability to a variety of labor intensive and analytical tasks ranging from document summarization to sentiment classification. These methods are also being applied to other areas of research. For example, in finance, text from financial news, social media, and company filings have been used to predict asset price movements and study the causal impact of new information (Tetlock, 2007). In macroeconomics, text is used to

[1] Cisco 2019, White paper, Cisco visual networking index: forecast and trends, 2017-2022.

[2] Natural Language Processing or NLP is a field of Artificial Intelligence that gives computers the ability to read, parse and derive meaning from human languages (**See appendix 2 for an overview of NLP**)

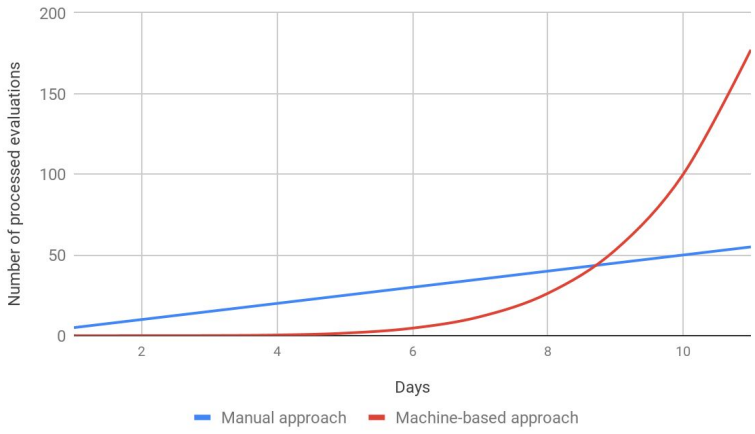
forecast variation in inflation and unemployment (Scott and Varian, 2015), and estimate the effects of policy uncertainty. In media economics, text from news and social media is used to study the drivers and effects of political slant (Gentzkow and Shapiro, 2010). Recently, Lauderdale and Herzog (2016) have also used these methods to quantify political polarization by extracting features from speeches given in the US Congress. A recent state-of-the-art review of the existing and future applications for economic and political research can be found in Gentzkow et al. (2019). These methods have also been used for conducting systematic reviews. The idea here is that technologies and methods for natural language processing have the potential to speed up the production of systematic reviews by reducing the amount of manual labour needed and hence to an extent partially automate the process. Marshall and Wallace (2019) provide an overview of current machine learning methods that have been proposed to expedite evidence synthesis, including their strengths and weaknesses, and how a systematic review team might go about using them in practice.

The field of international development cooperation also has its fair share of cases where new ways to conduct research and evaluation is being tested. OECD/DAC presented a working paper in 2019 that outlines that both the OECD and the World Bank are applying machine learning on a range of areas, such as topic modeling for classification of reports, tracking migration flows, and applying poverty prediction models. The working paper per se focused on using unsupervised machine learning to predict how international donors target the sustainable development goals (SDGs) with their projects (Pincet et al., 2019). Another central actor in this field is the UN Global Pulse, which is an UN initiative that works with and supports projects with focus on big data and artificial intelligence for development, humanitarian action and peace. Projects such as making Ugandan community radio machines readable using speech recognition which is a collaboration project between UN global pulse, the Makerere University and Stellenbosch University constitutes a good example of where novel methods are utilized. There are also other endeavours that uses high tech installations for data collection, for example using remote sensing and satellite imagery to improve response to humanitarian situations (Logar et al., 2020)

Project theme/question(s) and motivation

The ambition with this study will be to test how a data science and NLP approach for collection and analysis can enrich meta-evaluations within the field of international development cooperation. If successful, a data science approach could be used to generate high quality descriptive statistics, give quicker analytical turn around, and bring clear advantages in how meta evaluations are compiled. In particular, an advantage of this approach, once the methodology has been robustly implemented, would be that the number of evaluations included in the meta-analysis are of marginal importance for the overall effort required to conduct the analysis (see figure 1 for illustration).

Figure 1. Estimated time for Manual approach against Machine-based approach



Based on comments and inputs from EBA representatives we deem that the best structure for the proposed study involves the following steps:

1. We will make use of past evaluations included in an already published EBA meta-evaluation[3] with focus on OECD/DAC's evaluation criteria[4] sustainability to verify and calibrate various data science methods . In particular, we intend to use the analytical framework (see appendix 1 for details) from the mentioned study in order to evaluate the extent to which a data science approach is able to replicate the published results. Once this step is completed, we intend to:
2. Use a subset of the analytical framework from step 1, where the data science/NLP approach was successful in replicating the EBA results, on a larger dataset and a different OECD/DAC's evaluation criteria, namely that of relevance.[5]
3. Have a senior evaluator manually assess a sample of the evaluations that our model has processed for the evaluation criteria of relevance. The evaluator will be unaware of the model's assessment.
4. Comparison between the output from step 2 with the output from step 3. In short, to review the robustness of the results found in step 2.

[3] EBA has given tentative permission to access and use the data from the report *Livslängd och livskraft: Vad säger utvärderingar om svenska biståndsinsatsers hållbarhet?*

[4] Evaluation criteria *Relevance* - extent to which the evaluated project/programme is aligned with the needs, priorities and policies of the target group, beneficiaries donors, and other stakeholders. Note that the evaluation criterias have been amended over the course of the last 30 years. The most recent update was approved and implemented in 2019. The full set of OECD/DACs evaluation criteria are henceforth: Relevance, Coherence; Effectiveness; Efficiency; Impact and Sustainability (OECD/DAC 2019).

[5] Our approach thus relies on the available EBA results being accurate. However, this is not necessarily the case - which has been pointed out by EBA representatives. We therefore intend to manually assess cases where our analytical output deviates strongly from that of the available EBA results. We have also included the steps 3 and 4 as additional quality measures.

Research questions

In short, the study will address the following questions:

1) *Can a data science and NLP approach produce reliable assessments of what past evaluations have concluded about aid projects and programmes relating to OECD/DAC's evaluation criteria relevance?*
2) *What are the strengths and weaknesses of these methods compared to approaches relying on manual techniques?*

Method/s and literature/data to be used

This applied science study should be viewed as an attempt to test what is possible to accomplish with a data science and NLP approach when assessing available narrative evaluations. A brief overview of NLP methods and techniques can be found in appendix 2. Our approach will rely heavily on the high level and general purpose programming language Python. Examples of third party open source python dependencies and packages we will rely on include:

- Scrapy, Selenium and Beautifulsoup for scraping data and web crawling.
- Jupyter, Pandas, Numpy for data manipulation and analysis.
- PDFminer for data extracting and parsing of narrative texts.
- Spacy, Gensim and Fuzzywuzzy for natural language processing.
- Scikit-learn for machine learning and modeling.
- Plotly Dash and Matplotlib for data visualisation.

The implementation of the study is expected to involve a multi-pronged approach in accordance with the steps below, and we intend to utilize a range of novel analytical methods to process the past evaluations' sections with bearing on content of relevance for the questions spelled out in appendix 1.

1. **Data collection**. An semi-automated approach for web scraping will be set up that can identify and download relevant evaluations from Sida's online archive. A preliminary assessment suggests that a large share of Sida's decentralised evaluations is structured in a fruitful way that will facilitate deployment of our suggested approach on a relatively large scale. *(estimated time: Jonas 5 days, Gustav, 5 days)*
2. **Parsing relevant data from collected data**. Algorithms making use of advanced parsing methods and NLP techniques will be developed with the goal of identifying, parsing and storing various excerpts of texts that are needed to answer the questions posed in the analytical framework. The success of this extraction and parsing process is a prerequisite for the below listed steps. For instance, identification of an "executive summary" in the evaluation reports is clearly necessary to respond to question 23 in the

analytical framework regarding whether the evaluation mentions contribution *sustainability* in the executive summary. *(estimated time: Jonas 3 days, Gustav, 5 days)*

3. **Train and evaluate an assessment model**. The results data/output from EBA's meta-evaluation on sustainability (mentioned above) will be used to train a machine learning model for responding to the selected questions from the assessment framework in the same study (see appendix 1 for details). A range of natural language processing techniques involving for example text-summarization, text classification, text similarity measures, sentiment analysis etc, are potentially useful for this task. The results from the EBA sustainability report will also be used to validate our model's accuracy and potential for use on other evaluation criteria *(estimated time: Jonas 10 days, Gustav, 15 days)*
4. **Utilization of the trained model.** The trained model from stage 3 will be used on text sections from stage 2 but with bearing on the OECD/DAC criteria relevance - the model and its trained statistical algorithms will thus be deployed on content with bearing on the criteria relevance rather than sustainability. For questions in the analytical framework where the accuracy of our trained model (from step 3) is deemed sufficiently high (approx. 90%), this step will also include a scale up of the number of evaluations that will be processed and analysed. *(estimated time: Jonas 7 days, Gustav, 10 days).*
5. **Compilation, analysis and review of findings**. This step includes compilation on a range of statistics, including descriptive statistics (i.e. data on thematic focus, implementation period, geography etc), accuracy statistics (i.e. data on the models estimated performance), as well as the evaluators' assessments of the projects relevance per se. *(estimated time: Jonas 10 days, Gustav, 5 days).*
6. **Qualitative validation of model accuracy for the OECD/DAC criteria relevance**. Our model's analytical output will be cross-referenced with the senior evaluator's assessments for a random sample of the processed evaluations involving a subset of the questions included in the analytical framework . *(estimated time: Cecilia Ljungman 4 days).*
7. **Draft report.** Based on the steps 1-6 write a report addressing our findings in relation to the research questions outlined above. The format of the draft report will comply with EBA's instructions - written in English, avoid technical terms to the extent possible, contain a summary, and the report is not expected to exceed 25 main pages (excluding annexes). The report is furthermore expected to hold section on the following parts:
    a. **Literature overview** - a minor overview of what has been done in the field.
    b. **Used methods** - contain a section that elaborates on limitations, strengths and weaknesses of the applied methods.
    c. **Results** - presentation of the study findings and how it might contribute with new knowledge and/or insights.
    d. **Requirements** - what, in terms of resources and competence, is needed to undertake a study of this sort.
    e. **Potential future areas of application** - discussion of potential usage of the developed approach.

    (estimated time: Jonas 5 days, Gustav, 5 days and Cecilia 2 days).

8. **Visualisation and results dashboard**. A dashboard can be adjusted to serve many different purposes. For example, the findings from the analysis can be presented and be used as a platform for policymakers or the general public to retrieve information. Examples of content may include data on the evaluations' conclusions grouped and displayed for sub categories of interest, such as thematic focus, geographical areas, as well as how if might have altered over time. *(estimated time: Jonas 5 days).*

Scope and limitations

Our primary intention is, as mentioned above, to put the presented methods to the test. For clarification and transparency purposes we have developed Appendix 1 that holds a thorough assessment of limitations for each question included in the analytical framework. There is little doubt that there will be challenges to set up the approach, as well as to fine tune the intrinsic models to achieve high accuracy. However, based on our initial assessment and previous experiences, we hypothesise that this approach has the potential to generate good results for a majority of the questions included in the analytical framework used in the EBA study on sustainability. The scope of the study is expected to include a relatively large number of past evaluations, that are structured in accordance with OECD/DAC's evaluation criterias, and are available on Sida's web domain.

Note that this study will thus not focus on the quality or craftsmanship of the evaluations per se.

Target group

We think there are two main groups of beneficiaries of this study. One group is deemed to be more interested in the methodological and technical aspects of the study. This group is believed to contain evaluators in a broader sense from both the public and private sectors that are engaged in data management and analysis of larger narrative data sets. Examples of actors in this group might be governmental agency units, academia as well as private companies.

A second group is deemed to be more interested in the actual results and insights that the study potentially will produce. This group includes practitioners and policymakers within the realm of international development cooperation that might benefit from an approach that produces frequent updated meta evaluations. Examples of entities can be Sida, the Swedish Ministry of foreign affairs, and the Swedish National Audit Office. In addition, the general public with an interest in international development cooperation might also benefit should our approach prove to be successful in generating aggregated insights on what past evaluations conclude about Swedish funded projects and programmes[6].

[6] Note that this study aims to give an image of the evaluations conclusions relating to the evaluation criteria relevance. However, a system with all of OECD/DAC criterias could be set up if there is interest and if the model is proven to be successful.

Tentative results

The expectation of this applied science study is to illuminate the strengths and weaknesses of a data science and NLP approach for conducting meta evaluations for international development cooperation. As mentioned above, we have made a preliminary assessment of the challenges associated with each specific enquiry put forth in the analytical framework used in EBA's meta evaluation for sustainability (see appendix 1 for details). This assessment is based on our own previous work in the field, as well as our reading of the state-of-the-art applied research literature using these methods. Our main hypothesis is that the selected methods have a strong potential for addressing many of the questions put forth in the analytical framework. In addition and as mentioned above, if this study is successful it will be able to generate a range of improvements for how to generate meta evaluations, regardless the field of operation. Given that the approach is robust it will have clear advantages in both speed and accuracy in the assessments being made. We furthermore believe that our study fits well and will supplement other conducted EBA studies that have similar structure and limitations[7], but has applied more qualitative methods.

Budget and timeplan

Jonas Norén and Gustav Engström will be fully engaged throughout the project period. The proposed senior advisor - Cecilia Ljungman - will be engaged during the second half of the project. The number of estimated working days are motivated in the method section above. The team is ready to start implementation of the project during the second quarter of 2020, and conclude the project by years end. For the sake of transparency, we feel it pertinent to disclose that all team members have conducted work for Niras International Consulting, who is a frequent service provider of the evaluations that will be analysed. However, we see no conflict of interest; the study will not be based on a specific sample (i.e. no cherry picking); and the study will not review the quality nor account for responsible entities.

Team member	Engagement	Work days (8h)	Fee	Total costs
Jonas Noren	Full time	45	6400 kr	288 000 kr
Gustav Engstöm	Full time	45	6400 kr	288 000 kr
Senior advisor	Part time	6	6400 kr	38 400 kr
				614 400 kr

[7] Referring to EBA meta-evaluations on the evaluation criterias for sustainability (completed) and effectiveness (work in progress).

Authors

Gustav Engström,
(Ph.D Economics)
10+ years of mathematics and statistical research and consulting
Data Science – Python, Javascript, SQL, Stata, Excel

Gustav has a Ph.D. in economics from Stockholm University and has worked as a researcher for the past 6 years. His research has mainly focused on areas related to the economics of climate change, tipping points, development and urban economics. He has co-authored 15+ published articles in top level academic journals including Science and Nature. Currently he is engaged as a data scientist at Davcon where he conducts data science services with bearing on data driven approaches for collection, analysis and presentation of data. Gustav has experience with a wide variety of approaches for data collection and analysis. Apart from a high degree of knowledge in mathematics, statistics and inferential techniques his expertise covers areas such as machine learning, web-crawling, network analysis, text analysis/classification and web development. He thrives on open source solutions and the potential for automation of mundane processes using present state of the art technology. See attached CV for further details.

Jonas Norén
(M.Sc Political Science and Economics)
10+ years of evaluation, monitoring and strategy consulting
Data Science – Python, Excel, SQL

Jonas has worked with evaluations, strategy, capacity building, research, statistics and associated analysis for close to 15 years. He has mainly worked within the fields of private sector development and international development assistance. Currently he is engaged as a consultant at Davcon and conducts data science services with bearing on data driven approaches for collection, analysis and visualization of data. Jonas is experienced with different techniques and tools for data collection and analysis, which includes techniques for case studies, electronic survey systems, web crawling, network analysis, text analysis/mining and data storage/retrieval. Jonas is also experienced in launching and managing fully and/or semi-fully data-driven analytical IT-based systems that facilitates and drastically reduces time for data processing. See attached CV for further details.

Cecilia Ljungman (senior evaluation advisor)
(M.Sc Political Science, Economics, and Philosophy)
25+ years of evaluation, monitoring and strategy consulting

Cecilia is a senior evaluator with more than 25 years of international development cooperation experience. She has undertaken over 50 evaluations in over 30 developing countries for bilateral and UN clients; foundations; and CSOs. She has in-depth theoretical knowledge and practical experience of evaluation concepts, methodologies, and data collection.

References

Cisco. 2019. White paper, Cisco visual networking index; forecast and trends, 2017-2022.

Gentzkow, M., Kelly, B. and Taddy, M., 2019. Text as data. Journal of Economic Literature, 57(3), pp.535-74.

Gentzkow, M. and Shapiro, J.M., 2010. What drives media slant? Evidence from US daily newspapers. Econometrica, 78(1), pp.35-71.

Lauderdale, B.E. and Herzog, A., 2016. Measuring political positions from legislative speech. Political Analysis, 24(3), pp.374-394.
Scott, S. L., and Varian, H. R., 2015, Bayesian Variable Selection for Nowcasting Economic Time Series. In Economic Analysis of the Digital Economy, edited by Avi Goldfarb, Shane M. Greenstein, and Catherine E. Tucker, 119–35. Chicago: University of Chicago Press.

Logar, Tomaz. Bullock, Joseph. Nemni, Edoardo. Bromely, Lars. Quinn, John A. Luengo-Oroz, Miguel. 2020. PulseSatellite: A tool using human-AI feedback loops for satellite image analysis in humanitarian contexts.

Marshall, I.J. and Wallace, B.C., 2019, Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst Rev* 8, 163.

OECD/DAC. 2019. Better Criteria for Better Evaluation Revised Evaluation Criteria Definitions and Principles for Use OECD/DAC Network on Development Evaluation

Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. The Journal of finance, 62(3), pp.1139-1168.

UN Global Pulse, Making Ugandan Community Radio MachineREadable Using Speech Recognition Technology, Tool Series, no.1, 2016.

Pincet, Arnaud. Okabe, Shu. Pawelczyk, Martin. 2019. Linking Aid to the Sustainable Development Goals - A Machine Learning Approach. OECD Development Co-Operation Working Papers 52.

Python dependencies and packages that are expected to be utilised in the study include:

- Beautifulsoup - https://www.crummy.com/
- FuzzyWuzzy - https://github.com/seatgeek/fuzzywuzzy
- Gensim - https://radimrehurek.com/gensim/
- Jupyter - https://jupyter.org/
- Matplotlib - https://matplotlib.org/
- MySQL - https://www.mysql.com/
- Numpy - https://numpy.org/
- Pandas - https://pandas.pydata.org/
- PDFminer - https://github.com/euske/pdfminer
- Plotly Dash - https://plotly.com/dash/
- Python 3 - https://www.python.org/
- Scikit-learn - https://scikit-learn.org/stable/
- Scrapy - https://scrapy.org/
- Selenium - https://www.selenium.dev/
- Spacy - https://spacy.io/

Appendix 1 - analytical framework

The translated questions in framework below are copies of the set from the analytical framework used in EBA's study "*Livslängd och livskraft: vad säger utvärderingar om Svenska biståndsinsatsers hållbarhet*? Besides the actual questions, the inserted framework holds our current suggestions for how to approach a majority of the questions. It also includes our estimation relating to the level of difficulty, as well as our current level of confidence of success for each question. Note also that a few questions have been excluded due expected lack of data or limited value for the study.

Analytical framework

Includes suggestions for analytical approach, estimation on difficulty, estimation on success, as well as narrative comments

No	Question	Included in study	Analytical approach*	Difficult level**	Confidence for success***	Comments
1	Title of evaluation	Yes	Web scraping and text parsing.	Low	Highly confident	A web scraper will be set up and parse the evaluation title on Sida's web domain.
2	Evaluation number	Yes	Web scraping and text parsing.	Low	Highly confident	A web scraper will be set up and parse the evaluation series number on Sida's web domain.
3	Country (include all countries that has been mentioned in the evaluation)	Yes	Setup rule-based model for identification of countries.	Moder ate	Confident	All countries mentioned in the evaluations will be identified and their occurrences will be counted.

4	Region (geographical)	Yes	A rule-based model for identification of pre-coded regions.	Moderate	Confident	All regions mentioned in the evaluations will be identified and their occurrences will be counted. If needed NLP analysis/Named Entity Recognition for geographical entities might be applied.
5	Evaluation's geographical focus (Country/local; Region; Global)	Yes	Setup rule-based model for identification of evaluations geographical focus.	High	Unconfident	This analysis will probably be restricted to text passages that relate to the evaluations paragraphs for the contribution's objectives. Rules for what entails land/local; Region or Global need to be developed and applied. This task might seem simple enough. However, it is a rather complex task for a computer to identify text passages that are related to a specific topic.
7	Number of years that the contribution has received financing	No	n/a	Not included	Not included	Based on preliminary assessment of available data this number wasn't made available in most evaluations.
6	Time period that is being evaluated	Yes	Mixed-method of web scraping and parsing evaluation text/dates	High	Fairly confident	Based on parsing of project date interval and frequency counts the likely time period could be extracted. Alternatively based on EBA's study on sustainability, combination of rule based matching of date intervals and word embedding techniques could also be used to find sentences in other evaluations that have similar contexts from which date intervals could be extracted.
8	Number of financed project periods in succession	No	n/a	Not included	Not included	Based on preliminary assessment of available evaluations this data wasn't made available in most cases. In addition, the value for the overall analysis is deemed limited.
9	Is Sida a sole financier?	Yes	Mixed approach with NLP model for text similarity and a rule-based model for identification of pre-coded international donors.	High	Fairly confident	Based on a small subset of evaluations we will manually extract text passages mentioning project financing. These passages will then be used to create a model for text similarity based on word embeddings which will be utilised to extract text passages from other evaluations with bearing on finance. This will be combined with a rules based approach in order to identify mentioned donors.
10	Has the contribution received support from Sida before the current period?	No	n/a	Not included	Not included	Based on preliminary assessment of available data this number wasn't made available in most evaluations. In addition, the value for the overall analysis is deemed limited.
11	Thematic area	Yes	Setup rule-based model for identification of predefined thematic areas.	High	Unconfident	All (expected) thematic areas in the evaluations will be identified and their occurrences will be counted. There might be reason to include similarity measures based on word embeddings as well as sorting mechanisms to generate a sense of how various thematic areas have been prioritised.

12	Object for evaluation (Project; Programme; Strategy; Organisation; Theme)	Maybe	Mixed approach with statistics-based model and cross referencing to Sida's database.	Very high	n/a	The "Object" categorisation is arbitrary and even a human would have trouble to distinguish between the different types - e.g. what's a project, a programme or a strategy. We have therefore concluded that the best solution is to rely on some concrete reference for this to produce reliable output. This could for instance be variables in Sida's contribution statistics. We have been in contact with Sida's evaluation unit as well as the statistical unit in an effort to conclude if there is a way to connect the conducted evaluations with Sida's contribution statistics. Unfortunately, preliminary research suggests that there is no direct or available data (e.g. id contribution number) that can connect conducted evaluations with Sida's statistics. However, we believe that it might be possible to join on certain variables in Sida's statistics that are expected to match with the analytical output from our analysis (e.g. on title from question 1 as well as other characteristics, such as region, volume etc.). This will thus be dependent on the quality and our availability to use Sida's contribution database.
13	Implementing partner	Maybe	Ibid	Very high	n/a	Same logic as described above but look-up and extraction of "type of implementing partner"
14	At what phase of the contribution is the evaluation being conducted?	Yes	Web scraping, cross reference of outputs and rules-based model.	High	Fairly confident	This question is conditional on the success of question 6. Apart from this, the publication date can be scraped and compared with the output for the evaluated period. In addition, a rules-based assessment of key words (e.g. midterm review, end of programme evaluation etc.) can be applied on difficult cases.
15	Is the sustainability criteria being specified to the same extent as other evaluation criteria in evaluation's terms of reference?	Maybe	Rules-based model for assessing word frequency/count for the various OECD/DAC evaluation criteria	High	Fairly confident	Conditional on terms of references being included in the evaluation reports.
16	Does the evaluation give information on when the contribution's objectives are to be met?	No	n/a	Not included	Not included	Based on preliminary assessment of available data this number wasn't made available in most evaluations.
17	Is the contribution (and/or its results) deemed to be sustainable?	Yes	Statistic based model/s with focus on text similarity, word embeddings, n-gram negation analysis.	High	Confident	A synthesis analysis of a parsing mechanism able to identify text passages with bearing on specific content (e.g. OECD/DAC criteria for sustainability and/or relevance) and application of n-gram analysis for negation as well as text similarity using word embeddings.

18	Does the evaluation analyse the contribution's phasing-out strategy (incl. scaling-up / scaling-down)?	Maybe	Setup rule-based model for identification of paragraphs for recommendations.	High	Fairly confident	This analysis will probably be restricted to text passages that relates to the evaluations paragraphs for recommendations. Rules for what entails continued support; phasing out; termination of contribution. etc need to be developed and applied. This task might seem simple enough. However, it is rather complex task for a computer to identify text passages that is related to a specific topic.
19	Does the evaluation recommend whether the contribution should be phased out?	Maybe	Ibid	High	Fairly confident	Ibid
20	What kind of support is the evaluation recommending for the next phase of the contribution? (if yes in Q19)	Maybe	Ibid	High	Fairly confident	Ibid
21	Does the evaluation assess the importance of Sida's funding relating to the contributions sustainability/lack of sustainability?	Yes	Statistic based model/s with focus on text similarity, word embeddings, n-gram negation analysis.	Very high	Unconfident	A synthesis analysis of a parsing mechanism able to identify text passages with bearing on specific content (e.g. references to Sida and financial support) and application of n-gram analysis for negation as well as text similarity using word embeddings.
22	Does the evaluation analyse whether the contribution is dependent on funds from international donors?	Yes	Mixed approach with NLP model for text similarity and a rule-based model for identification of pre-coded international donors.	High	Fairly confident	See question 9.
23	Does the evaluation mention the contribution's sustainability in the evaluation's summary?	Yes	Text parsing and rule based identification	Moder ate	Confident	Involves extracting the executive summary and parsing text for identification of the relevant terms.
24	Does the evaluation mention the contribution's sustainability in the evaluation's recommendations?	Yes	Ibid	Moder ate	Confident	Analysis involves extracting the recommendation section (if available) and parsing text for identification of the relevant terms.
25	Does the evaluation give recommendations for how the contribution can improve its sustainability?	Yes	Mixed approach with NLP model for text similarity and a rule-based model	Very High	Unconfident	Analysis involves extracting the recommendation section (if available) and parsing text for identification of the relevant terms. Will also involve use of advanced methods such as parts of speech tagging and word embeddings.
26	What kind of improvements does the evaluation	Yes	Mixed approach with NLP model for text similarity	Very high	Unconfident	Analysis involves extracting the recommendation section (if available) and parsing text for identification of the relevant

	recommend?		and a rule-based model			terms. Will also involve use of advanced methods such as parts of speech tagging and word embeddings.

* A trained statistical model for identification of HTML structure in PDF documents are a prerequisite for the analysis of a majority of the selected questions. For example identify content of a specific nature (i.e. find paragraphs with bearing on method, limitations, results and/or specific OECD/DAC evaluation criteria), and/or characteristics of the texts (sentence length, text size, or format of text etc.). Despite a high level of difficulty, we are relatively confident that we can train a model that can reach a high accuracy level.

** The scale for difficult levels range from: Very low; Low; Moderate; High, Very high.

*** The scale for confidence for success ranges from: Highly confident; Confident; Fairly confident; Unconfident

Appendix 2 - Natural language processing

Natural language processing (NLP) is a subfield of linguistics, computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyze large amounts of natural language data.

Main advantages of NLP

Natural Language Processing plays a very important role in structuring data because it prepares text and speech for machines, so that they're able to interpret, process, and organize information. Some of the main advantages of NLP include:

- **Large-scale analysis**. Natural Language Processing can help machines perform language-based tasks such as reading text, identifying what's important, extracting sentiment, or hearing speech, on a large scale.
- **Structuring unstructured data.** Human language is complex, varied, and ambiguous, while machine language relies on logical and highly structured languages and information. NLP bridges the gap between the way we talk and how computers decipher information. By using grammatical rules, algorithms, and statistics, it can interpret natural language andhelp structure large quantities of text.

NLP techniques

Syntactic

Syntactic analysis — also known as parsing or syntax analysis — studies the grammatical rules in natural language with the purpose of uncovering the structure of a text. Identifying the syntactic structure of a text and the dependency relationships between words — which are represented on a diagram called a parse tree — also contribute to understanding the meaning of words.

Syntax analysis involves many different techniques, including:

Tokenization
This is the process of breaking up a string of words into semantically useful units called tokens. You can use sentence tokenization to split sentences within a text, or word tokenization to split words within a sentence. This NLP task works by defining boundaries, that is, a criterion of where a token begins or ends. Generally, word tokens can be separated by blank spaces, and sentence tokens by stops. However, you can perform high-level tokenization for more complex structures, like words that often go together, otherwise known as collocations (for example, New York).

Tokenization makes a text more simple and easy to handle, and it's the most basic task in text pre-processing.

Part-of-speech tagging
Part-of-speech tagging (abbreviated as PoS tagging) involves adding a part of speech category to each token within a text. Some common PoS tags are verb, adjective, noun, pronoun, conjunction, preposition, intersection, among others.

PoS tagging is useful for identifying relationships between words and, therefore, understand the meaning of sentences

Dependency Parsing
Dependency grammar refers to the way the words in a sentence are connected to each other. A dependency parser, therefore, analyzes how 'head words' are related and modified by other words in order to understand the syntactic structure of a sentence.

Lemmatization & Stemming
When we speak or write, we normally use inflected forms of a word (words that derive from others). Lemmatization and stemming are two similar NLP tasks that consist of reducing words to their base form so that they can be analyzed by their common root. The word as it appears in

the dictionary – its base form – is called lemma . For example, the words ‘are, is, am, were, and been’, are grouped under the lemma ‘be’. When we refer to stemming, on the other hand, the root form of a word is called stem. For example, using stemming for the words “consult”, “consultant”, “consulting”, and “consultants”, would result in the stem “consult”.

These are useful techniques for finding common word denominators.

Stopword Removal
This process consists of filtering out high frequency words that add little or no semantic value to a sentence. For example, which, to, at, for, is, etc, are all words that don’t help you understand a text. This is useful in NLP since these words often convey little meaning.

Semantic

Semantic analysis focuses on identifying the meaning of text. By analyzing the structure of sentences and the interactions between words in a given context, semantic analysis tries to find the proper meaning of words that might have different definitions. Combined with computer science, semantic analysis can help understand the topic of a text, as it can identify the presence of related concepts. That way, a news article containing the words investors, market, and recession would be labeled as “economics”.

Semantic analysis includes e.g.:

Word Sense Disambiguation
Depending on their context, words can have different meanings. Take the word “book”, for example in the sentences: “You should read this book!” , “You should book the flight.” and “You should do everything by the book to avoid conflicts.”. There are two main techniques that can be used for Word Sense Disambiguation (WSD): knowledge-based (or dictionary approach) and supervised approach. The first one tries to infer meaning by observing the dictionary definitions of ambiguous terms within a text; while the latter requires training data and is based on machine learning algorithms that can learn from examples. Identifying the meaning of a word based on context is still a major (and open) challenge faced by Natural Language Processing.

Relationship Extraction
This task consists of identifying semantic relationships between two or more entities in a text. Entities can be names, places, organizations, etc; and relationships can be established in a variety of ways. For example, in the phrase “Henry lives in Los Angeles”, a person (Henry) is related to a place (Los Angeles) by the semantic category “lives in”.

Methods

There are two main technical approaches to Natural Language Processing that create different types of systems: one is based on linguistic rules and the other on machine learning methods. In this section, we'll examine the advantages and disadvantages of each one, and the possibility of combining both (hybrid approach).

Rule-Based Approach

Rule-based systems are the earliest approach to NLP, and consist of applying hand-crafted linguistic rules to text. Each rule is formed by an antecedent and a prediction. So, when the system finds a matching pattern, it applies the predicted criteria.

Since the rules are determined by humans, this type of system is easy to understand and provides fairly accurate results with little effort. Another advantage of rule-based systems is that they don't require training data, which makes them a good option if you don't have much data and are just starting your analysis.

However, manually crafting and enhancing rules can be a difficult and cumbersome task, and often requires a linguist or a knowledge engineer. Also, adding too many rules can lead to complex systems with contradictory rules.

Machine Learning and statistic based models

Machine Learning consists of algorithms that can learn to understand language based on previous observations. The system uses statistical methods to build its own 'knowledge bank', and is trained to make associations between a particular input and its corresponding output.

The biggest advantage of machine learning models is their ability to learn on their own, with no need to define manual rules. All you'll need is a good set of training data, with several examples for each of the tags you'd like to analyze.

Machine learning models can have higher precision and recall than rule-based systems over time, and the more training data you feed them, the more accurate they are. However, you'll need enough training data relevant to the problem you want to solve in order to build an accurate system.

An example of successful applications of machine learning models involve text embeddings**,** which are real valued vector representations of text strings. This are structured so that each word has a dense vector, chosen so that it's similar to vectors of words that appear in similar contexts. Word embeddings are considered a great starting point for most deep NLP tasks. They allow deep learning to be effective on smaller datasets, as they are often the first inputs to a deep learning architecture and the most popular way of transfer learning in NLP. The most popular names in word embeddings are Word2vec by Google and GloVe by Stanford.

Hybrid Approaches

A third approach involves combining both rule-based systems and machine learning systems. That way, you can benefit from the advantages of each of them, and gain accuracy in your results.

Common NLP Algorithms

Natural Language Processing involves using all kinds of algorithms to identify linguistic rules, extract meaning, and uncover the structure of a text. Some of the most popular algorithms that can be used in NLP depending on the task you want to perform can be categorized as follows:

Text Classification
Text classification is the process of organizing unstructured text into predefined categories (tags). Text classification tasks include sentiment analysis, intent detection, topic modeling, and language detection. Popular algorithms for creating text classification include: Naive Bayes, Support Vector MAchines and Deep learning.

Text Extraction
Text extraction consists of extracting specific pieces of data from a text. You can use extraction models to pull out keywords, entities (such as company names or locations), or to summarize text. The most common algorithms for text extraction include: Tf-IDF, Regular Expressions, TextRank

Topic Modeling
Topic modeling is a method for clustering groups of words and similar expressions within a set of data. Unlike topic classification, topic modeling is an unsupervised method, which means that it infers patterns from data without needing to define categories or tag data beforehand. The main algorithms used for topic modeling include: Latent Semantic Analysis, Latent Dirichlet Allocation.

Appendix 3 - CV

dav|consulting
Data, Analytics & Visualisation

Gustav Engström
Date of birth: 17 April 1977
Adress: Björnssonsgatan 141
168 43
Phone: (+46)70-416 4244
E-mail: gustav@davsthlm.se

Current position

2018- Chief data scientist at dav consulting.

Education

2012 Doctor of Philosophy (Ph.D.), Department of Economics, Stockholm University, Supervisor: John Hassler
2009 Licenciate degree, Department of Economics, Lund University, Supervisor: Tommy Andersson
2006 Ph.D. candidate, Department of Economics, Lund University
2005 Master of Social Science, Department of Economics, Lund University

Consultancy work

2019 Automated natural language assessment for Ford Foundation's BUILD Programme. Worked on behalf of DAVCON which was procured by NIRAS to give analytical assistance and advisory service in setting up a data driven approach for mapping the foundation's objectives against grantee reports. The analysis enabled NIRAS to retrieve automated insights on grantees progress and how it corresponded to Ford's objectives/strategies. Acted as chief data scientist in project and worked mainly in the phases of data manipulation and analysis of the data.
2018 Worked as data scientist for a well-established Swedish private enterprise. Apart from standard data analysis work included, building prediction models of customer behaviour, data scraping and geocoding of data. This enabled them to improve customer related performance.
2018 Development of web-based platform for tracking financial institutions' performance relating to target prices on the Swedish stock exchange. Worked on behalf of DAVCON in developing an automated tracker system that collects, compiles and assesses the performance of all financial actors that are making their estimates on target prices publicly known. Acted as the chief data scientist in the project and worked in all parts of the project.

Pedagogical Merits

Teaching

2018-2016 Challenges of environmental decision making, Master course, Stockholm Resilience Center (Course leader & lecturer)
2016 Ecosystem Support of Humanity, Master course, Stockholm Resilience Center
2015 Challenges of environmental decision making, Master course, Stockholm Resilience Center (Course leader & lecturer)
2015 Ecosystem Support of Humanity, Master course, Stockholm Resilience Center
2014 Ecology and Economic management, Ph.D. Course given by the Beijer Institute - courtesy of Göteborg University
2014 Ecosystem Support of Humanity, Master course, Stockholm Resilience Center
2013 Ekologisk ekonomi, Master course given at the Department of Physical Geography and Quaternary Geology resilience centre (GE1004)
2013 Challenges of Environmental Decision-making - Master course given at the Stockholm resilience centre
2012 The Economics of the Environment Ph.D. Course given by the Beijer Institute – courtesy of Stockholm University
2012 Ecology and Economic management, Ph.D. Course given by the Beijer Institute - courtesy of Göteborg University
2011 Ecological Economics (GE1004), Stockholm University
2010 The Climate and the Economy, EZ7104 (Master course), Stockholm University
2010 Ecology and Economic management 2010 (Ph.D. Course given by the Beijer Institute)
2008 Teaching & Teaching Assistant in Mathematics, Microeconomic Theory, NEKA11 (introductory level), Lund University
2007-2008 Teaching Assistant, Financial Economics, NEKA11 (introductory level), Lund University

Supervision

2011 Strategies for pollination services as a productive input in Canola production : Fredrik Granath (Master thesis)

Scientific merits

Publications in international peer-reviewed journals

1. Walker, B., S. Barrett, S. Polasky, V. Galaz, C. Folke, G. Engström, F. Ackerman, K. Arrow, S. Car- penter, K. Chopra, G. Daily, P. Ehrlich, T. Hughes, N. Kautsky, S. Levin, and K.-g. Mäler (2009). Looming Global-Scale Failures and Missing Institutions. *Science* 325(September), 7–8.
2. Carpenter, S. R. et al. (2012). General Resilience to Cope with Extreme Events. *Sustainability* 4(12), 3248–3259.
3. Brock, W., G. Engström, D. Grass, and A. Xepapadeas (2013). Energy balance climate models and general equilibrium optimal mitigation policies. *Journal of Economic Dynamics and Control* 37(12), 2371–2396.
4. Norström, A., M. Metian, M. Schlu¨ter, L. Schultz, A. Dannenberg, G. McCarney, M. Milkoreit, F. Diekert, G. Engström, J. Gars, M. Sanctuary, R. Fishman, E. Kyriakopoulou, M. Sjöstedt, V. Manoussi, K. Meng, and M. Schoon (2013). Social change vital to sustainability goals. *Nature* 498(20 June), 299.
5. Brock, W., G. Engström, and A. Xepapadeas (2014). Spatial climate-economic models in the de- sign of optimal climate policies across locations. *European Economic Review* 69, 78–103.
6. Norström, A. V., A. Dannenberg, G. McCarney, M. Milkoreit, F. Diekert, G. Engström, R. Fishman, J. Gars, E. Kyriakopoolou, V. Manoussi, K. Meng, M. Metian, M. Sanctuary, M. Schlu¨ter, M. Schoon, L. Schultz, and M. Sjöstedt (2014). Three necessary conditions for establishing effective Sustainable Development Goals in the Anthropocene. *Ecology and Society* 19(3).

7. Engström, G. and J. Gars (2015). Optimal Taxation in the Macroeconomics of Climate Change. *Annual Review of Resource Economics* 7(1), 127–150.
8. Engström, G. (2016). Structural and climatic change. *Structural Change and Economic Dynamics* 37, 62–74.
9. Engström, G. and J. Gars (2016). Climatic Tipping Points and Optimal Fossil-Fuel Use. *Environmental and Resource Economics* 65, 541–571.
10. Ando, M., M. Dahlberg, and G. Engström (2017). The risks of nuclear disaster and its impact on housing prices. *Economics Letters* 154, 13–16.
11. Cre´pin, G., Å . Gren, G. Engström, and D. Ospina (2017). Operationalising a social-ecological system perspective on the Arctic Ocean. *Ambio* 6 (3).
12. Engström, G. and Å . Gren (2017). Capturing the value of green space in urban parks in a sustainable urban planning and design context: pros and cons of hedonic pricing. *Ecology and Society* 22 (2).
13. Lade, S. J., L. J. Haider, G. Engström, and M. Schlu¨ter (2017). Resilience offers escape from trapped thinking on poverty alleviation. *Science Advances* 3(5).
14. Blind, I., M. Dahlberg, G. Engström, and J. Ö sth (2018). Construction of Register-based Com- muting Measures. *CESifo Economic Studies* 64(2).

Book chapters

1. Brock, W., G. Engström, and A. Xepapadeas (2014). "Energy Balance Climate Models, Damage Reservoirs and the Time Profile of Climate Change Policy". In: *The Oxford Handbook of the Macroeconomics of Global Warming*. Ed. by L. Bernard and W. Semmler. Oxford University Press. Chap. 3.

Monographs

1. Engström, G. (2012). "Essays on Economic Modeling of Climate Change". PhD thesis. Stockholm University, Faculty of Social Sciences, Department of Economics.

Popular science and media articles

1. Troell, M., G. Engström, Å . Jansson, and A.-S. Cre´pin (2011). *Klimathotet består*. http://www.unt.se/asikt/debatt/klimathotet-bestar-1481424.aspx.
2. Blind, I., M. Dahlberg, and G. Engström (2016). Prisutvecklingen på bostäder i Sverige – en geografisk analys. *Ekonomisk Debatt* 4.
3. Cre´pin, A.-S., G. Finnveden, M. Hennlock, L. Neij, M. Nilsson, E. G., and B. Lars (2018). Möjligheter och begränsningar med samhällsekonomiska analyser. *VRHU rapport*.

Research grants

2015- "Global bio-physical processes in climate-economy modelling". Research program financed by "The Ragnar Söderberg Foundation". Amount: 5.4 MSEK. Co-applicants: Chandra Kiran and Johan Gars. Role: principal investigator.

2013-2016 "Utvärdering av samhällsförändringar och offentlig politik med hjälp av databasen 'Den urbana bostadsmarknaden'". Financed by Jan Wallanders och Tom Hedelius Stiftelse Tore Browaldhs Stiftelse (Handelsbanken). Amount: 3 MSEK. Co-applicants: Matz Dahlberg (principal investigator), Chuang-Zhong Li, Ina Blind. Role: researcher

2011-2013 "Databas: den urbana bostadsmarknaden". Financed by Jan Wallanders och Tom Hedelius stiftelse (Handelsbanken). Amount: 622 kSEK. Co-applicants: Matz Dahlberg (principal investigator), Chuang-Zhong Li. Role: researcher

International research collaborations

• Urban Biodiversity and Ecosystem Services, URBES, FP7-ERA-net project. Pan-european project aimed at improving knowledge surrounding urban biodiversity and eco-system services.

• Arctic Climate Change, Economy and Society (ACCESS), FP7-OCEAN-2010. Pan-european project aimed at evaluating the the consequence of climate change in the Arctic.
• Green Surge, FP7-ENV-2013-two-stage. Pan-european project with the purpose of identifying, developing and testing ways of linking urban green spaces, bio-diversity with the urban economy.

International conference & workshop participations

2016 On the use of Geocoded data in economic research, CESifo workshop, Mu¨nchen (organizer)
2016 Policies for the Anthropocene workshop, Handelshögskolan, University of Gothenburg (participant)
2016 The economics of the planetary boundaries, Ragnar Söderberg funded workshop (organizer)
2014 The Economics of Complex Systems, Stockholm (presenter)
2014 WCERE, The Fifth World Congress of Environmental and Resource Economists, Istanbul (presenter)
2013 20th Ulvön Conference on Environmental Economics, 18-20 Jun, Ulvön, Sweden (presenter)
2012 The Macroeconomics of Climate Change, CEPR London, 13 Dec. (presenter)
2012 EAERE, European Association of Environmental and Resource Economists, 19th Annual Conference, Prague (presenter)
2011 Resilience 2011, Phoenix, Arizona, USA (presenter)
2010 WCERE, The Fourth World Congress of Environmental and Resource Economists, Montreal (presenter)
2010 11th Workshop on Optimal Control, Dynamic Games and Nonlinear Dynamics, Amsterdam (presenter)
2010 20th International Climate Policy Workshop, Stockholm (organizer)
2009 EAERE, European Association of Environmental and Resource Economists, 17th Annual Conference, Amsterdam (presenter)
2009 ALEAR, IV Congreso de la Asociacio´n Latinoamericana y del Caribe de Economistas Ambientales y de Recursos Naturales, Costa Rica, (presenter)

Invited talks

2014 Department of economics, Gothenburg University

Referee work

Environment and Development Economics, Environmental Economics and Policy Studies, Economic Letters, Journal of European Economic Association, Macroeconomic dynamics

Editorial work

2016 Guest-editor CESIfo (special issue) - ”On the Use of Geocoded Data in Economic Research”

Other scientific merits

2016 Yttrande från Kungl. Vetenskapsakademien beträffande delbetänkandet från Miljömålsberedningen Ett klimatpolitiskt ramverk för Sverige (SOU 2016:21). Joint with: Henning Rodhe, Johan Kleman and Johan Gars
2015 Grading committee member: Licentiate defense, Jamila Haider, ”Understanding poverty traps in biocultural landscapes”

Technical skillset

Statistics: At least half of my academic work has involved statistical analysis. During recent years I have increasingly been engaged with what is known as quasi-experimental methods which can be summarized as a broad set of statistical methods for separating out causation from correlation in observational studies.

Mathematics: Mathematics is at the center of much of the academic work done in economics. My PhD and much work there after has involved optimal control and dynamic programming methods for infinite time horizon problems.

Python: Have been working in the Python programming environment since 2011. Mostly used for scientific computing and statistics in Jupyter notebook but also web scraping and web development.

SQL: Started working with SQL Server in 2005 handling high-frequency data. Recent work has been mostly in MySQL and PostgreSQL.

Stata: Most statistical work done by academic economists uses Stata. Hence almost all my collaborative work has relied on this software for statistical analysis.

Other non-academic merits

2014 Constructed a non-linear regression model in Stata, for housing value prediction for the Swedish housing market which was used as web based valuation tool for a Swedish startup company.

2017 Constructed an algorithmic trading bot in Python for the Stockholm stock exchange. The algorithm was based on the machine learning techniques Random Forest and XGBoost and built as standalone trading app with automatic trade execution on the Avanza webplatform.

2017-2018 Built a web-platform for an independent project for evaluating the performance of stock market rating firms.

dav|consulting
Data, Analytics & Visualisation

Jonas Norén
Date of birth: 26 May 1980
Adress: Björnssonsgatan 141
168 43
Phone: (+46)703-57 2656
E-mail: jonas@davsthlm.se

Educational record

2001 – 2006	**Lund University** – Master's degree (major in Political Science)
2005 (January – July)	**Loughborough University UK** – Erasmus scholarship (Institution of Politics, International Relations and European Studies.
1996 – 1999	**Falkenberg upper secondary school** (social science study programme with focus on economics and business economy

Languages

	Level of proficiency
Swedish	Mother tongue
English	Full professional proficiency (level 2)

Other

Professional proficiency in MS Office (Excel specifically); Python (including a range of modules and packages for statistical-, network- and data science analysis); SQL for data storage; Prezi for vivid presentations, and various survey and web scraping/crawling techniques for data collection.

Work related country experiences

Work related experience (**including field visits**) from: Serbia, Colombia, North Korea, Morocco, China, Cambodia, Zimbabwe, Tanzania, Belgium, Thailand, Lebanon, Uganda, Bolivia, and the UK.

Work related experience (**coordinated from Sweden**) from: Palestine, Iraq, Sri Lanka, Vietnam, Kenya, Jordan, Rwanda, Burundi, South Africa.

Professional experience

2018 – dav|consulting
(www.davcon.se)
Partner /Consultant
- DAVCON (Data Analysis Visualisation consulting) is a data science consultancy firm that provides services relating to data collection, data analysis and data visualisation.

2014 – Beyond Intent
(www.beyondintent.se)
Partner /Consultant
- Shift AB is a consultancy firm that provides services in four areas: sustainability, risk management, international business development.

2008 – norcon
(www.norcon.se)

Entrepreneur/Consultant

- Norcon AB is a consulting company that offers investigative and analytical services in the context of politics, economics and international development assistance.

2008 – 2010 — ARS Research AB (www.ars.se)

Consultant /analyst

- ARS Research is a research institute that specializes on surveys and opinion polls for a wide range of clients, in both the public and private sector.

2008 (March – October) — företagarna Swedish Federation of Business Owners (www.foretagarna.se)

Assistant at the Department for Opinion and Analysis

- Företagarna mission is to promote entrepreneurship, safeguard the rights of business owners, and contribute to conditions that make running a business easy and appealing.

2006 – 2008 — Sida Swedish International Development Cooperation Agency (www.sida.se)

Administrator at the Department EVU

- Sida is a government agency under the Ministry for Foreign Affairs that aim to contribute to making it possible for poor people in developing countries to improve their living conditions.

Selection of relevant assignments and studies

Evaluation and independent reviews of programmes and organisations

2019 — **Evaluation of Sida's Resource Base Programmes.** The purpose of the evaluation was to assess the relevance and results of eleven resource base programs – including MFS, JPO, BAE, SARC, UN and EU secondments – and give recommendations on how the programs' alignment with the new Swedish strategy for capacity development, partnership and methods for supporting the Agenda 2030 could be improved. Jonas was part of a NIRAS team commissioned to undertake the evaluation.

2019 — **Evaluation of Sida's Risk Management Contributions.** Jonas was procured to assist with data gathering and statistical analysis, in a NIRAS evaluation team, that was commissioned to undertake the evaluation by Sida.

2019 — **Evaluation of the International Disability Alliance (IDA) 2015-2018.** Jonas was part of an international team that conducted an independent evaluation of IDA. He was furthermore procured to oversee data management during the evaluation. Primary data was collected through a web-based survey and a tailor made web-crawling approach. Collected data underwent network analysis for deeper connectivity insights of the network per se. The evaluation was commissioned by Sida.

2019 — **Mid-term review of the Sida-funded programme SwedBio at Stockholm Resilience Centre.** Jonas was part of NIRAS evaluation team that was commissioned to undertake the review in accordance to OECD-DAC evaluation criteria and produce recommendations for the current (2016-2020) as well as for the potential next phase.

2018/2019 — **Evaluation of the IBON International and the CSO Partnership for Development Effectiveness (CPDE) Project.** Jonas was part of an international evaluation team assigned by Sida. The evaluation set out to evaluate the project and its ambition to foster effective development cooperation, as well as its ability to serve as an international platform for non-governmental organisations.

2018 **Evaluation of the Swedish Leadership for Sustainable Development network.** Jonas was part of an evaluation team assigned by Sida to evaluate the network's past performance in accordance with OECD/DAC standards, as well as to document conducted activities and the network's story per se. The network gathers some of the largest corporations in Sweden with a collective turnover that constitute a large part of Sweden's total GDP. The network was and is a novel collaboration modality for Sida with the ambition to find new and alternative ways to collaborate with the private sector and to find common strategies that can contribute to the Sustainable Development Goals and Agenda 2030.

2018 **Evaluation of Sida's International Training Programmes for Quality Infrastructure in support of World trade (ITP 304) and Food safety (ITP 305).** Jonas was part of an evaluation team, procured by Niras AB, that evaluated the ITPs which had been implemented by the Swedish Board for Accreditation and Conformity Assessment (Swedac). Jonas was assigned to reorganise available data from 10 years of operations, conduct statistical analysis as well as to compiled the portfolio analysis.

2017 **Independent Evaluation of the Swiss Agency for Development Cooperation's (SDC) System for Results-based Management.** Jonas was procured to structure and administrate an inception phase survey that went out to all SDC staff. Jonas also conducted the statistical analysis of the collected data and designed the results reports.

2017 **Team member of SMPs consultancy group for methodological advisory services to SCB/ICO international development cooperation.** The advisory group is selected to give methodological advice and perform evaluations on Statistics Sweden's (SCB/ICO) past and current development projects. Jonas is part of a small team that will deliver methodological services (results-based management, monitoring, evaluations, and equality) for the next two years.

2017 **Fact-finding assignment and facilitation of project design for SIS – IBNORCA collaboration standardization project in Bolivia.** The assignment set out to gather data and knowledge about the Bolivian context and stakeholders in an effort to design a standardisation project. Jonas was contracted to facilitate workshops, data gathering and to assist in the compilation of a project proposal.

2015 **Process evaluation of UNHCR Sweden and Clowns without Boarder's education project Back to School**. The project aimed to fill gaps in the education services for refugee children in Lebanon. Through the combination of Clowns without Boarder's methods and UNHCR Sweden's outreach, it was expected that the project would reach the most vulnerable children and give them education and trauma reduction. Jonas was contracted to evaluate the project per se and to assess the cooperation between the two project partners in accordance with OECD/DAC evaluation criteria for relevance, efficiency and effectiveness.

2014 **Review of IM's International Department.** IM is a Swedish aid organisation that fights and exposes poverty and exclusion. Jonas was procured to conduct a global perception-based survey in order to assess organisational strengths and weaknesses.

2014 **Evaluation of the challenge fund Innovation Against Poverty (IAP).** The IAP is a pilot challenge fund that was launched by Sida in 2011 and administrated by Swedish consultancy PricewaterhouseCoopers. The programme was a risk sharing mechanism for innovative business ventures with the potential to reduce poverty in developing countries.

2014 **Assessment with focus on programme effectiveness and efficiency during 2012-2014 for the Palestinian International Business Forum (PIBF)**. Jonas was commissioned to conduct the assessment and collect data in a mixed method approach in order to single out operational recommendations for optimization of future operations.

2013 **Needs assessment for the Institute of Standards of Cambodia (ISC).** Jonas was assigned to oversee the assessment and convey M&E advice in accordance with the RBM apparatus to local consultants. The overall purpose with the consultancy was to strengthen the institutional research capacity of ISC and reinforce their approach for a results-focused and evidence-based management culture.

2013 **Strategic review of Spider at Stockholm University**. Spider is a resource center for ICT for Development (ICT4D) at Stockholm University. Jonas was procured to collect evidence and compile a report on organisational strengths and weaknesses, as well as to identify areas for future improvements.

2012-2013 **Review of SymbioCity Academy pilot programme in Iraq.** Jonas was procured to provide an independent review of SymbioCity and its subcomponents within the sectors of electricity as well as water and sanitation. The programme was administrated by the Swedish Trade Council (STC) and the Swedish Energy Research Institute (IVL). The review set out to estimate programme performance; how the programme has promoted Swedish business; and ultimately how it can contribute to business development in Iraq.

2012 **Evaluation of the International Council of Swedish Industry's (NIR) development programme in North Korea.** NIR's programme has set out to increase knowledge and capacity for business management and international trade amongst North Korean stakeholders during the last ten years 2002-2012. Jonas was procured to evaluate the programme in a participatory process as well as to give recommendations for NIR's oncoming operations.

2012 **Evaluation of multi-stakeholder initiatives Better Aid and Open Forum.** The initiatives were designed to improve aid effectiveness by supporting Civil Society Organisations (CSO) working to enhance CSO effectiveness, donor engagement, and mutual CSO learning.

2012 **Study and draft of strategy for Colombian labour market programme**. Jonas was procured as an external analyst and evaluator by NIR, LO and TCO. The assignment set out to deliver recommendations for how the established Colombian network could become self-sustaining as well as to improve the network's future competitiveness on the Colombian labour market.

2011 **Midterm review of a programme for private sector cooperation in Zimbabwe PSC-SWESZIM).** The assignment included a results-oriented assessment with focus on programme outcomes on a beneficiary level. Jonas was procured as an independent contractor by NIR to conduct a web survey amongst central stakeholders within the programme and to analyze and present statistics and narrative findings.

2009 **Review of Swedish support for political party development and democratic governance through organizations affiliated to parties represented in the Swedish parliament (PAO).** Jonas worked as external analyst in affiliation with SPM consultants for Sida. The assignment included a thematic assessment and an evaluation in accordance with Results-based Management.

2009 **Evaluation of National Forum** – regional seminar program with purpose to gather regional- and national politicians on issues concerning regional growth. Procured as analyst to the Ministry of Enterprise. The assignment consisted research and stakeholder analysis in an inception phase for the Department for Regional Growth's broader evaluation of the National Forum (Co-author to inception report*: Uppföljning av Nationellt forum 2009*).

2009 **Evaluation on Green Public Procurement (GPP) for the Swedish Environmental Protection Agency**. Engaged as external consultant and analyst. A survey was carried out among the target group that included public actors – municipalities, governmental agencies etc. the main purpose was to evaluate the impact of GPP in the public sector (Report: *Miljöanpassad upphandling inom offentlig sektor, 2009).*

2008 **Evaluation of the Anti-discrimination regulation within the Public Procurement Act (LOU) for the Swedish Competition Authority**. Engaged as external consultant and analyst. A survey was carried

out and the target group included governmental agencies and private enterprises. The assignment was an initial step in an extensive appraisal of the regulations future relevance. (Report: *Undersökning av myndigheters och leverantörers upplevelser av Antidiskrimineringsförordningen, 2008*).

Research, data management, analyses and data synthesis

2019 **Automated natural language assessment for Ford Foundation's BUILD Programme.** DAVCON was procured by NIRAS to give analytical assistance and advisory service in setting up a data driven approach for mapping the foundation's objectives against grantee reports. The analysis enabled NIRAS to retrieve automated insights on grantees progress and how it corresponded to Ford's objectives/strategies. Jonas acted as a data scientist in the project and worked mainly in the phases of data extraction and data synthesis.

2019 **Development of web-based platform for tracking financial institutions' performance relating to target prices on the Swedish stock exchange.** DAVCON has developed an automated tracker system that collect, compile and assess the performance of all financial actors that are making their estimates on target prices publicly known. Jonas acted as a data scientist in the project and worked mainly in the assessment of financial institutions.

2017 **GAP analysis for the Swedish agency for accreditation (SWEDAC) with bearing on the results framework for Agenda 2030.** Jonas was contracted to conduct a gap analysis on SWEDACs operations and its potential contribution to Agenda 2030 and sustainable development. The objective was to map the agency's current operations to specific objectives and targets within the Agenda 2030 results framework. The assignment also included recommendations on appropriate actions, including limitations and (focus) strategy, that were deemed to strengthen Swedac's potential contribution to different target areas.

2017 **Chapter on usage of 'Big data' (or lack thereof) in evaluation within international development cooperation. Published in Routledge's book series covering comparative policy evaluation – Cyber Society, Big Data and Evaluation 2017.** The full title is: "Understanding and Utilizing the Dynamics in Data Ecosystems. The chapter was co-authored with Kim Forss and is available here.

2016 **Advisory services on survey design to the Swedish Civil Contingency Agency (MSB).** Jonas was part of a Niras/Indevelop team that gave advice and practical support in the development of a digital survey for MSB's program for sending resource persons to various places on the globe. Two different surveys were design based on a satisfied customer model. One targeted resource personnel and the other on institutional partner organisations.

2016 **Statistical and analytical services during an external review of AfDB's Rural Water Supply and Sanitation Initiative (RWSSI) and Trust Fund.** Jonas conducted a portfolio analysis of the initiatives various programmes and projects that had been or was planned to be implemented between 2003-2020. Jonas was procured as an external analyst in this assignment commissioned by the African Development Bank (AfDB).

2016 **Design and development of workload tracker tool for the Humanitarian Leadership Academy (HLA).** The assignment set out to create an interactive data driven workload tracker tool that enable HLA to accurately and proactively plan and track human resource allocations for on-going and upcoming projects.

2015 Research paper on '**Big data' in evaluations within international development cooperation** published in the European Evaluation Society's newsletter Connections June 2015. The full title is: *"Using 'Big Data' for equity-focused evaluation – understanding and utilizing the dynamics of data ecosystems"*. The paper was co-authored with Kim Forss as is available here.

2015 **National research survey on sustainability (i.e. CSR) among Swedish small- and medium-sized Swedish companies**. The survey was a joint venture with the Swedish Federation of Business Owners. The results and final report "*Hållbara företag bygger framtiden*" is available here.

2014 **Portfolio analysis of Sida's support to MENA 2010-2015**. The assignment's main purpose was to present the portfolio's thematic content and disbursements over time. The analysis was furthermore part of an inception phase of a larger strategic and forward-looking assessment for Sida support to the MENA region. Jonas was procured as an external analyst and private sector specialist by Indevelop AB in this assignment commissioned by Sida.

2013 **External review of Sida/UTV's project Indicators in Development Aid.** The assignment set out to review and estimate the quality of retrieved data from Sida's various units under the project "Indikatorer i biståndet". Jonas was commissioned to assist Sida with hands-on analysis of the compiled data and to draft recommendations for coming phases of the project.

2011-2012 **Design and development of a national risk index for NIR target economies**. The assignment included rigorous research, design and analysis of a wide range of statistics from valid producers of statistical data. The data was furthermore compiled in a web application for interactive usage. Jonas was procured by NIR (the International Council of Swedish Industry).

2011 **Assessment for the Swedish East African Chamber of Commerce (SWEACC).** Evidence-based data was collected through a survey carried out amongst SWEACC's members. Valid decision support documents were delivered with strategic information for future adjustments of SWEACC's operations.

2010 **Design and development of national risk assessments for NIR target economies.** The assignment included design and assembly of reports based on macro level statistics, with a comparative focus on business environment, political complexity and social context at a national level.

2010 **Portfolio analysis of Sida's support to democracy and human rights during 2006-2009**. The assignments main purpose was to present the portfolio's thematic content and significant changes over time. In addition, the evaluation included breakdowns of several sector codes and mapping of the composition of Sida's support to democracy, human right and public administration.

2010 **Evaluation on national public vaccination programs for the Ministry of Health and Social Health.** Procured as external analyst in affiliation with ARS Research. A national study within the Swedish health care system was conducted to conclude how public vaccination programs should be decided, organized and financed in the future. The evaluation was part of SOU 2010:39 (Ny ordning för nationella vaccinationsprogram – betänkande av Vaccinutredningen 2010).

2009 **Swedish Environmental Protection Agency's national survey on public perception on climate change**. Engaged as external consultant and analyst. Several surveys were carried out in a number of areas related to climate change – public perception on the issue, corporate responsibility, gender, etc. (Reports: *Allmänhetens kunskaper och attityder till klimatförändringen 2009; Allmänhetens syn på företagens klimatarbete 2009; Allmänhetens attityder och konsumtionsbeteenden kopplat till klimatförändringen 2009;* and *Genusperspektiv på allmänhetens kunskaper och attityder till klimatförändringen 2009*).

Results-based management, evidence-based approaches and advisory services

2018 (on-going) **Advisory services to the Swedish Standards Institute (SIS) and IBNORCA for standardization and capacity development project in Bolivia.** This project aims to strengthen the capacity and use of standards within the Bolivian Water sector. Jonas is procured to deliver strategic guidance and to deliver monitoring and evaluation services during the course of the implementation of the project.

2015 (on-going)	**Advisory services to the Swedish Board for Accreditation and Conformity Assessment (Swedac) and the Agadir Technical Unit (ATU)**. This programme is designed to support the establishment of technical infrastructure in the Agadir countries in an effort to bring down technical barriers to trade. Jonas is procured to convey strategic advice and deliver M&E services in accordance with RBM for the second phase of the programme. The consultancy includes components of operational research and strengthening of the programme's evidence-based governance approach. A systems-based M&E approach has been developed where data is collected analysed and synthesised in a systematic and semi-automated manner.
2016 -2018	**Monitoring and evaluation services to the ARSO-ECOWAS-SIS project.** Jonas was procured to facilitate the data collection process and operationalize the projects results framework in accordance with RBM and practices for good governance. The assignment included a regional wide survey to the 15 countries in the ECOWAS region, as well as guidance to the regional team in research-based operations.
2012 - 2017	**Advisory services to the Swedish Standards Institute (SIS)**. Jonas was procured to convey strategic advice and deliver M&E services in accordance with RBM for two development programmes in South and South East Asia as well as in East Africa. The consultancy included components of strategic planning and reporting; operational research; and strengthening of the programme's evidence-based governance approaches. The assignment also held trainings and M&E capacity building of local staff at both regional hubs. Programme coordinators have been trained in data collection technics and instructed to conduct annual semi-structured interviews and observations.
2017	**Advisory services to Forum Syd and the unit for Impact and Quality Assurance (IQA).** Forum Syd is a politically and religiously unaffiliated development cooperation organisation that work with human and civil rights around the globe. Forum Syd has around 160 member organisations from Swedish civil society. Jonas is procured to support the IQA and the consultancy includes a range of services within the areas of methods design, planning, monitoring and evaluation within Forum Syd's fields of operation.
2011-2014	**Advisory services to the International Council of Swedish Industry (NIR).** Jonas was procured to give advisory services and methodological consultancy for NIR's operations. Jonas gave technical support in the construction of a RBM monitoring system for NIR's international development programmes. The assignment included design of strategic indicators and implementation of qualitative and quantitative research techniques to evaluate programme outcomes in complex markets. The overall purpose was to improve the results-oriented approach and establish an evidence-based management culture in accordance with RBM and policies for good governance.
2013	**Strategic workshop for the Swedish Workplace HIV/AIDS Programme**. Jonas was procured to convey strategic advise for an upcoming programme application; and to facilitate a RBM workshop in Harare/Zimbabwe with SWHAP management and regional coordinators. The overall aim and purpose with the consultancy was to, through a participatory approach, establish a new programme design. The assignment also held a component and development of a tailor made process and evaluation tool that allowed for the full team to engage in programme planning, implementation and follow-up in a pragmatic way.
2011	**Advisory services to the Statistical Office in the Republic of Serbia (SORS).** Strategic advice was conveyed in an effort to strengthen SORS organisational capacity. Jonas was procured to facilitate workshops in Belgrade and advice on Result-Based Management and other OECD/DAC guidelines.
2011	**Advisory service on results frameworks for the Swedish Standards Institute (SIS).** The central aspects of the consultancy were improvements of the intervention logic (i.e. theory of change) and monitoring possibilities for a specific development project in the Eastern Africa region. All recommendations given were derived from the results-based management apparatus and other guidelines from OECD/DAC.

2010 **External advisor to the International Council of Swedish Industry (NIR) and the department of evaluation and methodology.** The overall purpose was to improve the results-oriented approach and establish an evidence-based management culture in accordance with Results-Based Management (RBM). The assignment had different phases with both theoretical and operative stages.

2008 **Assistant advisor to Sida on increasing results orientation (RBM)** in management, implementation and evaluation of the International Training Programmes. (Co-author to report: *Begreppsapparat för Resultatstyrning av ITP*, 2008).

Governance and implementation of development cooperation programmes

2015-2016 **Team leader for the East African Community (EAC) and the Swedish Standards Institute's (SIS) joint development programme.** Jonas was procured to lead a regional team and facilitate the programme, which aimed to strengthen standardization bodies as well as private sector institutions in the EAC region (Burundi, Kenya, Rwanda, Tanzania and Uganda). The programme ran between 2012 and 2016 and was financed by Sida.

Other experiences

2008 – 2012 **Member of the Civil Crisis Management Team at the Swedish Red Cross** – The team has a permanent stand by duty for international or national emergencies at Arlanda and Bromma airports. Educated in Rescue Psychology, First Aid, and Airport organization/Logistics.

1999 – 2006 **Positions as elected representative** – In the army I was elected to represent the platoon at gatherings with the regiment management. I was also elected foreman at the student union during the years at Lunds University.

1999 – 2000 **Army experiences** – 11 months of compulsory military service at the ranger regiment I22 in Kiruna (Mark 10-8-8)

References

On request