Name: Kfir Bar   **Date:** 11.11.2020

Computer Science, College of Management, Academic Studies

CURRICULUM VITAE

1. Personal Details

Address: 8 Hadera St. Tel Aviv

Cellular Phone: +972-54-4457771

Electronic Address: barkfir@yahoo.com

2. Higher Education

A. Undergraduate and Graduate Studies

BA, Computer Science, The College of Management, Israel   1999-2002
Major topic and final project: Speech Recognition. Graduated summa cum laude.

MSc, Computer Science, Tel Aviv University   2002-2006
Research Title: Arabic to English Example Based Machine Translation.
Supervisors: Prof. Nachum Dershowitz and Prof. Yaacov Chouka (from Bar Ilan University, Israel). Graduated summa cum laude.

B. Doctoral Degree and Post-Doctoral Studies

PhD, Computer Science, Tel Aviv University 2008-2013
Research Title: Deriving Paraphrases for Highly Inflected Languages, with a Focus on Machine Translation.
Supervisor: Prof. Nachum Dershowitz.
Postdoc, Digital Humanities, The Hebrew University   2014-2015
Investigating semantic techniques for indexing motif titles Translation.
Supervisor: Prof. Alexander Kulik.

3. Participation in Scholarly Conferences

A. Active Participation

Date	Name of Conference	Place	Subject of Lecture/ Discussion	Role
04/2007	ICITIS	Morocco	Example-Based Arabic-to-English Translation	
06/2009	IEEE international conference on Intelligence and security informatics	Dallas, Texas	Automatically Classifying Documents by Ideological and Organizational Affiliation	

06/2010	Israeli Seminar on Computational Linguistics (ISCOL)	Tel Aviv, Israel	Using Synonyms for Arabic-to-English Example-Based Translation	Organizing Committee
08/2009	Open Machine Translation 2009 Evaluation	Ottawa, Canada	Tel Aviv University's System Description for NIST EVAL 2009	
11/2010	The Ninth Conference of the Association for Machine Translation in the Americas	Denver, Colorado	Using Synonyms for Arabic-to-English Example-Based Translation	
12/2010	International Workshop on Spoken Language Translation	Paris, France	Tel Aviv University's System Description for IWSLT 2010	
01/2011	International Workshop on Machine Translation and Morphologically-rich Languages	Haifa, Israel	Using Verb Paraphrases for Arabic-to-English Example-Based Translation	
07/2012	ACL 2012	Jeju, Korea	Building an Arabic Multiword Expressions Repository	
12/2012	The 24th International Conference on Computational Linguistics (COLING)	Mumbai, India	Deriving Paraphrases for Highly-Inflected Languages from Comparable Documents	
06/2013	Israeli Seminar on Computational Linguistics (ISCOL)	Beer Sheba, Israel	Extracting Multiword Expressions for Arabic	
06/2013	Bar-Ilan Symposium on Foundations of Artificial Intelligence	Bar Ilan University, Israel	Deriving Paraphrases for Highly Inflected Languages from Comparable Documents	
04/2014	Conference on Intelligent Text Processing and Computational Linguistics (CICLing)	Kathmandu, Nepal	Inferring Paraphrases for a Highly Inflected Language from a Monolingual Corpus	
09/2014	Israeli Seminar on Computational Linguistics (ISCOL)	Haifa, Israel	Automatic Transliteration of Judeo-Arabic Texts into Arabic Script	
10/2014	Conference on Empirical Methods on Natural Language Processing (EMNLP)	Doha, Qatar	The Tel Aviv University System for the Code-Switching Workshop Shared Task	
10/2015	Conference on Intelligent Text Processing and Computational Linguistics (CICLing)	Egypt	Processing Judeo-Arabic Texts	

03/2016	SemEval-2016	San Diego, California	SLS at SemEval-2016 Task 3: Neural-based Approaches for Ranking in Community Question Answering	

B. **Organization of Conferences or Sessions**

Date	Name of Conference	Place	Subject/Role of Conference, Comments	Role
06/2010	Israeli Seminar on Computational Linguistics (ISCOL)	Tel Aviv University, Israel	The most important Israeli conference on Natural Language Processing	Organizing Committee

4. Invited Lectures/Colloquium Talks

Date	Place of Lecture	Name of Forum	Presentation/Comments
01/2014	Menlo Park, California	SRI Artificial Intelligence Seminar	Invited by SRI, topic: Deriving Paraphrases for Highly Inflected Languages, with a Focus on Machine Translation
04/2014	IBM Tel Aviv	IBM Debating research seminar	Invited by IBM research group: Automatic Derivation of Arabic Paraphrases
03/2015	COMAS	School Colloquium	Processing Judeo-Arabic Texts
06/2016	Tel Aviv	The Israeli Natural Language Processing Meetup	Active Learning for NLP

5. Research Grants

Submission of Research Proposals – Pending

Role in Research	Co-Researchers	Topic	Funded by	Year
Co-Researcher	Prof. Shmuel Itzikovitz, Dr. Ido Ziv, Dr. Galit Haim, Dr. Igor Ruchlin	Trilateral Proposal for Early Identification and Intervention for Stress Related Disorders Using mHealth	DFG (German Research Foundation)	2017

6. Scholarships, Awards and Prizes

2001 – Wolf scholarship award for excellence (during BA)

2001 – College of Management award for excellence (during BA)

2005 – Tel Aviv University award for excellence (during MSc)

PhD – full scholarship by Tel Aviv University for 3 years.

Summer 2013 – visiting scholar at the Center for Computational Learning Systems at Columbia University, New York. Full scholarship by Tel Aviv University.

7. Teaching

Courses Taught in Recent Years

Year	Course Name	Type: Lecture/Seminar/ Workshop/ High Learn Course/ Introduction	Degree	No. of Students
2008-2012	**Java Programming** (Open University)	Lecture	BSc	50
2010, 2011, 2012	**Computer Networks** (COMAS)	Lecture	BSc	50-150
2009	**Algorithms 1** (COMAS)	Lecture	BSc	50-150
2013, 2014	**Natural Language Processing for graduate students** (Tel Aviv University)	Lecture	MSc, PhD	20-30
2012, 2014, 2015,2017,20 18,2019	**Natural Language Processing for undergrads** (Tel Aviv University, COMAS)	Lecture	BSc	20
2013, 2014, 2015	**Seminar of Natural Language Processing** (Tel Aviv University)	Seminar	MSc, PhD	15

2012-2017	**Machine Learning** (COMAS)	Lecture	BSc	100
2014-2017	**Network Lab** (COMAS)	Lab	BSc	100
2015	**Machine Translation** (Tel Aviv University)	Seminar	MSc, PhD	20
2016	**Digital Humanities for Computer Scientists** (Tel Aviv University)	Seminar	BSc	15
2019	**Applications for Machine Learning** (Tel Aviv University)	Lecture	BSc, MSc, PhD	80

8. <u>Professional Experience</u>

<u>2000-2002</u> **NLP Software Engineer, Baobab Inc.**
- Developing the natural language generation (NLG) engine.
- Developing tools for ontology management and integration.
- Development in C++.

<u>2002-2005</u> **Software Engineer, Itemfield Inc. (acquired by Informatica)**
- Member of the UI team.
- Designing and implementing front-end applications.
- Development in C++ (MFC) and Java (SWING, SWT).

<u>Summer 2011</u> **Visiting Scholar, Columbia University, NY**
- Invited by the Center for Computational Learning Systems (CCLS) @ Columbia University to participate in their research activities.
- Working under the supervision of Prof. Mona Diab.
- Performing research in the field of Multi Word Expressions (MWE) in Arabic. Part of the work has been recently published (see publications below).

<u>2005-2012</u> **CTO/VP R&D, IntuView Ltd.**
- Managing a team of 10 NLP developers and linguists. Some had years of experience in the field (PhD, MSc), and for some it was their first professional position.
- Responsibility for the design and implementation of a system, which analyses unstructured text written in different languages (Arabic, Urdu, English, French and Spanish) and generates immediate insights, such as named entities, concepts, sentiment, and document classification information.

- Designing algorithms that work on top of RDF store technologies, such as Oracle Semantic Network, Allegro graph, and Jena.
- Designing and developing a name-matching application, focusing on culture-based and naming convention models.
- Responsibility for designing and building large ontologies for the domain of homeland security and defense.

About IntuView: IntuView is dedicated to the development of software for security and defense applications. IntuView's core product performs analysis of textual documents, deducing explicit and implicit information on their authorship, political leanings, ideas, concepts, issues and sentiments, and extracting contextual information on entities mentioned in the texts.

2005-Present Lecturer at the School of Computer Science, College of Management Academic Studies

- Active member of the academic committee of the undergraduate program for computer science (2011-present).
- Teaching basic courses – Algorithms, Computer Networks.
- Teaching advanced courses – Machine Learning, Natural Language Processing.
- Leading a research on semantic processing of social-media texts.
- Developing a system for homework plagiarism, using sophisticated semantic technologies.

2014-Present External Lecturer at the School for Computer Science, Tel Aviv University
- Teaching a Natural Language Processing seminar for graduate students.
- Teaching, together with Prof. Nachum Dershowitz a Natural Language Processing course (http://www.cs.tau.ac.il/courses/0368-4341/2015b/).
- Teaching a Machine Translation seminar for graduate students (http://cs.tau.ac.il/~kfirbar/mt-seminar.html).
- Teaching a Digital Humanities seminar (with special focus on natural language processing), fall of 2016-217.

2013-2017 Co-founder and CTO, Comprendi Inc.
- Comprendi turns textual big data into actionable marketing insights. Our platform analyzes huge amounts of text from social media and helps marketers better understand their target audience and target the right users with the right offer at the right time. More about Comprendi – www.comprendi.net

- **NEW - September 2016**: Comprendi Wins $250,000 Grand Prize in Twitter #Promote Ads API Challenge (read more here: https://blog.twitter.com/2016/promote-ads-api-challenge-2016-announcing-the-regional-finalists-and-grand-prize-winner)

2018-present Chief Scientist, Basis Technology Inc.
- We're the leading provider of software solutions for extracting meaningful intelligence from multilingual text and digital devices. More about Basis – www.basistech.com

PUBLICATIONS

A. Ph.D. Dissertation
B. Articles or Chapters in Scientific Books

1. **Kfir Bar** and Nachum Dershowitz. Using Semantic Equivalents for Arabic-to-English Example-Based Translation. 2012. In: *Soudi, Abdelhadi, Ali Farghaly, Günter Neumann and Rabih Zbib (eds.), Challenges for Arabic Machine Translation*. viii, (pp. 49–72)

2. **Kfir Bar**, Yaacov Choueka, and Nachum Dershowitz, November 2014, "Matching Phrases for Arabic-to-English Example-Based Translation System", *Language, Culture, Computation: Computational Linguistics and Linguistics, Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday, Part III, Nachum Dershowitz and Ephraim Nissan, eds., Lecture Notes in Computer Science*, vol. 8003, Springer-Verlag, Berlin, pp. 54-63

3. **Kfir Bar**, Mona Diab and Abdelati Hawwari. "Arabic Multiword Expressions". November 2014. *Language, Culture, Computation: Computational Linguistics and Linguistics, Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday, Part III, Nachum Dershowitz and Ephraim Nissan, eds., Lecture Notes in Computer Science*, vol. 8003, Springer-Verlag, Berlin

C. Articles in Conference Proceedings

1. **Kfir Bar**, Nachum Dershowitz, and Yaacov Choueka, Apr. 2007, "Example-Based Arabic-to-English Translation", In *Proceedings of a Workshop on Arabic Natural Language Processing*, Fez, Morocco, pp. 325/1-4.

2. Moshe Koppel, Navot Akiva, Eli Alshech and **Kfir Bar**, June 2009, "Automatically Classifying Documents by Ideological and Organizational Affiliation", In *Proceedings of the 2009 IEEE international conference on Intelligence and security informatics (ISI'09)*.

3. **Kfir Bar**, Nachum Dershowitz, Nov. 2010, "Using Synonyms for Arabic-to-English Example-Based Translation", In *Proceedings of The Ninth Conference of the Association for Machine Translation in the Americas (AMTA-9)*, Denver, CO, USA.

4. Abdelati Hawwari, **Kfir Bar** and Mona Diab, July 2012, "Building an Arabic Multiword Expressions Repository", In *Proceedings of the ACL 2012 Joint Workshop on Statistical Parsing and Semantic Processing of Morphologically Rich Languages*, Jeju, Korea.

5. **Kfir Bar** and Nachum Dershowitz, Dec. 2012, "Deriving Paraphrases for Highly-Inflected Languages from Comparable Documents", *In Proceedings of the 24th International Conference on Computational Linguistics (COLING)*, Mumbai, India.

6. Lior Wolf, Yair Hanani, **Kfir Bar**, and Nachum Dershowitz, April 2014, "Joint word2vec Networks for Bilingual Semantic Representations" (Poster), *Conference on Intelligent Text Processing and Computational Linguistics (CICLing)*, Kathmandu, Nepal.

7. **Kfir Bar** and Nachum Dershowitz, April 2014, "Inferring Paraphrases for a Highly Inflected Language from a Monolingual Corpus", In *Proceedings of the Conference on Intelligent Text Processing and Computational Linguistics (CICLing), Part 2*, Kathmandu, Nepal, *Lecture Notes in Computer Science, vol. 8404, Springer-Verlag*, Heidelberg, pp. 254-270.

8. **Kfir Bar** and Nachum Dershowitz, October 2014, "The Tel Aviv University System for the Code-Switching Workshop Shared Task", In *Proceedings of the First Workshop on Computational Approaches to Code-Switching, Thamar Solorio, Elizabeth Blair, Suraj Maharjan, Steve Bethard, Mona Diab, Mahmoud Gonheim, Abdelati Hawwari, Fahad AlGhamdi, Julia Hirshberg, Alison Chang, and Pascale Fung, eds., Conference on Empirical Methods on Natural Language Processing (EMNLP)*, Doha, Qatar.

9. **Kfir Bar**, Nachum Dershowitz, Lior Wolf, Yackov Lubarsky, and Yaacov Choueka, 2015, "Processing Judeo-Arabic Texts", In Proceedings of the *First Arabic Natural Language Processing Conference, IEEE, Cairo*

10. Mitra Mohtarami, Yonatan Belinkov, Wei-Ning Hsu, Yu Zhang Tao Lei, **Kfir Bar**, Scott Cyphers, James Glass, January 2016, "SLS at SemEval-2016 Task 3: Neural-based Approaches for Ranking in Community Question Answering", In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)* , San Diego

11. **Kfir Bar**, Nachum Dershowitz, Lena Dankin, March 2018, "Metaphor Interpretation Using Word Embeddings", In *Proceedings of the 19th International Conference on Computational Linguistics and Intelligent Text Processing (Cicling-2018)*, Hanoi, Vietnam

12. **Kfir Bar**, Nachum Dershowitz, Lena Dankin, March 2018, "Metaphor Interpretation Using Word Embeddings", In *Proceedings of the 19th International Conference on Computational Linguistics and Intelligent Text Processing (Cicling-2018)*, Hanoi, Vietnam

13. **Kfir Bar**, Vered Zilberstein, Ido Ziv, Heli Baram, Nachum Dershowitz, Samuel Itzikowitz, and Eiran Vadim Harel, June 2019, "Semantic Characteristics of Schizophrenic Speech", In *Proceedings of the Sixth Annual Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic (CLPsych), Kate Niederhoffer, Kristy Hollingshead, Philip Resnik, Rebecca Resnik, and Kate Loveys, eds., 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, Minneapolis, MN, pp. 84-93.

14. Ori Terner, **Kfir Bar**, and Nachum Dershowitz, December 2020, "Transliteration of Judeo-Arabic Texts into Arabic Script Using Recurrent Neural Networks", *The Fifth Arabic Natural Language Processing Workshop (WANLP 2020), co-located with COLING'2020*, Barcelona, Spain.

D. Summary of My Research Activities and Future Plans

I have completed my PhD studies in Computer Science at Tel Aviv University in 2013, which I dedicated for studying data-driven approaches for machine translation, under the supervision of Prof. Nachum Dershowitz. After my graduation, I decided to focus more on the applied side of NLP, and started my own startup company, Comprendi, for improving marketing campaigns led by large enterprises, using online chatter analysis. After four years at Comprendi, I joined Basis Technology, an experienced American company that develops tools for processing human languages on different linguistic levels. Today, and for the last two years, I am Basis' Chief Scientist and head of research.

In parallel to my career in the industry, I have constantly been teaching courses in computer science, with a special focus on machine learning and NLP, in a number of academic institutes in Israel. Even before my graduation, I was teaching the Java Programming course at The Open University for four years. Since 2014, I have been taking a part-time position as a faculty member at the College of Management's Computer Science School, where I am teaching courses in NLP, and machine learning for undergrads. I developed a new practical course for machine learning, named Applications for Machine Learning for graduate students, which I was teaching during the Spring of 2018 at Tel Aviv University. The course was declared a success, and was invited to be given also in the following semester (Spring 2020) at Bar Ilan University. Prior to that, also at Tel Aviv University, I was teaching a seminar in machine translation in 2015, and helped to teach a few instances of the NLP and Digital Humanities course along with Prof. Nachum Dershowitz.

During my entire career I have always been interested and involved in research. Specifically, in the past 18 months I am officially co-advising a number of MSc students in Computer Science at Tel Aviv University, and a few others at IDC. My main topic of interest is NLP, with a special focus on solving real-life medical problems. Against this background, I am actively collaborating with a few medical institutes in Israel for leading research in related topics. To name a few, I am collaborating with a group of mental-health clinicians from the Beer Yaacov Mental Health Institute, and a number of computer scientists, to study a data-driven approach for automatically

detecting schizophrenia. We have recently published our first set of results produced by our innovative NLP algorithm for finding symptoms of thought disorders in transcribed speech of schizophrenia inpatients. As a follow-up work, we are now evaluating our findings on a larger collection of social-media posts. Using computational methods for diagnosis of mental illness is a relatively new interdisciplinary field, which combines expertise in psychiatry, psychology, sociology, neuroscience, linguistics, and computer science. Recently, I initiated and co-organized a successful event in Tel Aviv University for the first time in Israel, which brought together top experts from all the relevant fields, to discuss potential collaborations. I am planning to continue leading this initiative for further developing this important interdisciplinary conversation. I am currently collaborating with Prof. Doron Friedman from the School of Communications, and Prof. Yacov Hel-Or from the School of Computer Science, both from IDC. We co-advise a number of graduate students to work on the interface between human languages, facial expressions and emotional conditions.

In another project, I am collaborating with a group of oncologists led by Prof. Ido Wolf (MD), from the Tel Aviv Sourasky 'Ichilov' Medical Center. In this project we look at a large collection of disease-course summaries written by physicians in free Hebrew text, and use NLP algorithms to extract a timeline of medical events, including symptoms and treatments that the patient has gone through. As part of this project, one of the MSc students under my supervision, who works full time on this project, is building a unique collection of documents, taken from different domains, and manually annotate them for relevant events and temporal relations between them, which we will use for training different types of classifiers in order to arrange events into a timeline. Partially funded by the hospital, we currently focus on oncological patients, and we believe that the outcome of this NLP research will provide a platform for conducting data-driven oncological studies in the near future. I recently submitted a grant proposal to the Israeli Science Foundation (ISF), as a principal investigator in collaboration with Prof. Nachum Dershowitz, in order to fund this project.

In another project, based on a collaboration of the College of Management with Yitzhak Shamir Medical Center (also known as Assaf Harofe), I am working closely with Prof. Orna Tal (MD), to study risk factors in emergency-room (ER) reports, written in free Hebrew text, using NLP algorithms. This project is still in early stages.

On another front, I spent several years studying the Arabic language and its computational challenges, resulting in a few related publications, mostly about Arabic-to-English machine translation and detecting multi-word expressions in free running text. I am currently advising an MSc student who is working on building a partial vocalizer for Arabic, taking a deep-learning approach.