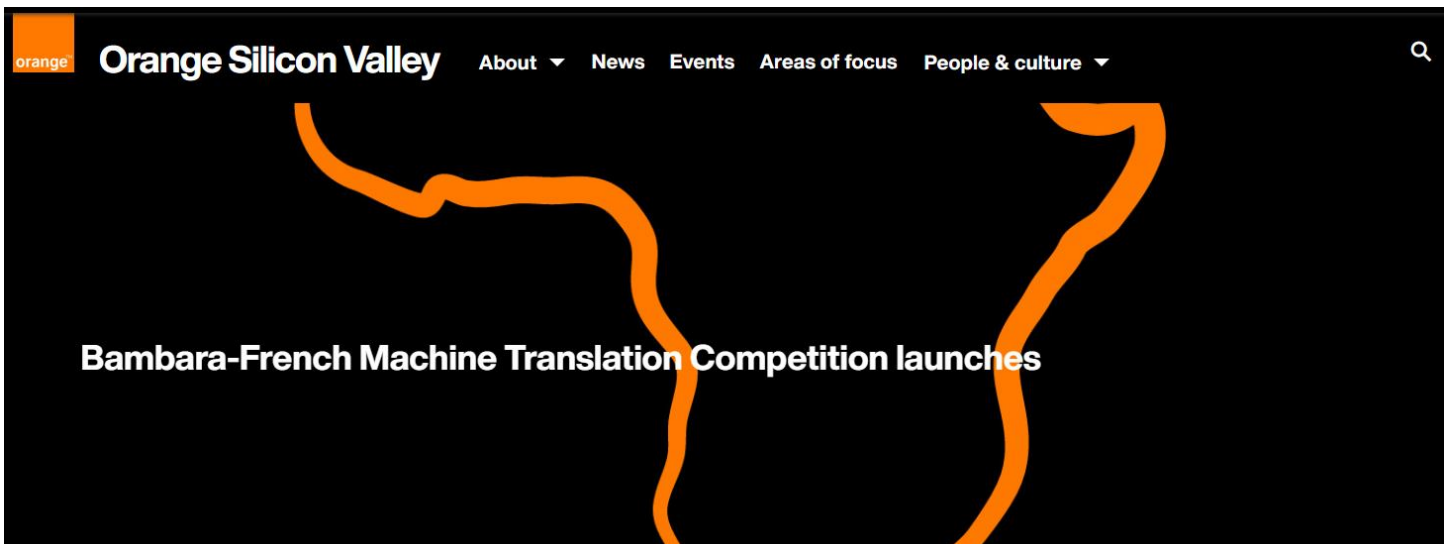# SAMPLE: Technology Company Business Incubator Competition Project

Subject: machine learning analysis, B2B global event content creation & website design

Douglas Cramer: douglascramer.com, douglascramer@live.com

orange **Orange Silicon Valley**   About ▾   News   Events   Areas of focus   People & culture ▾

Bambara-French Machine Translation Competition launches

Orange Silicon Valley ❯ news ❯ Bambara-French Machine Translation Competition launches   Listen 🔊

By Douglas Cramer | Aug 1, 2022

Dr. Sarah Luger, Technology Group Principal at Orange Silicon Valley, has launched our **Bambara-French Machine Translation Competition**. We invite our interested readers to participate and share with others this compelling opportunity to increase language representation in telco technologies.
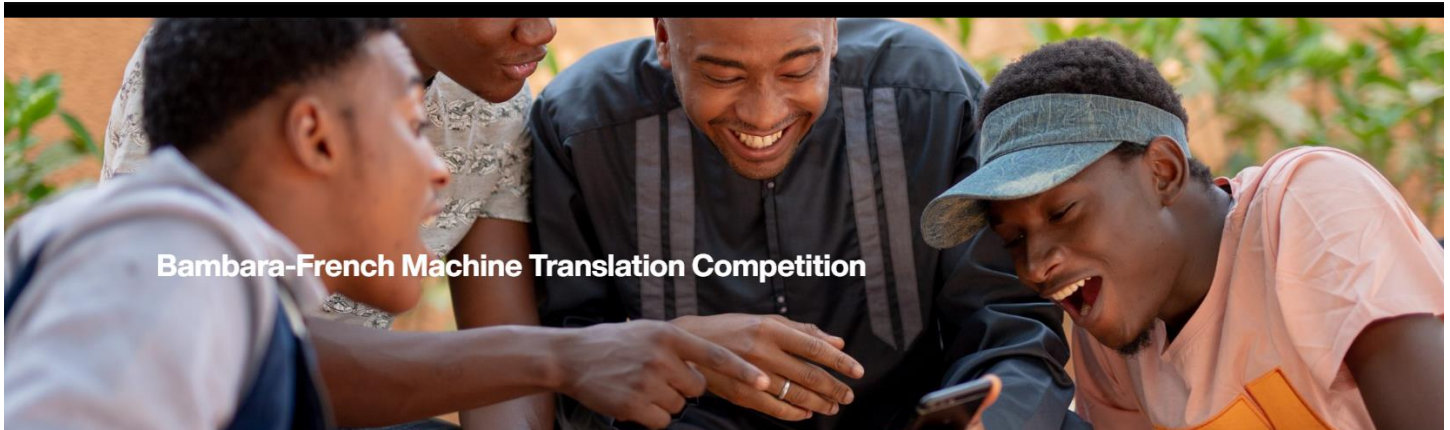
The competition focuses on the development of a machine translation model for Bambara, a low-resource language. It is predominantly spoken in Mali and is one of the country's 13 national languages, with 15 million speakers. Bambara is one of the closely related Manding languages, which are part of the larger Mandé language family.

Recent advancements in Neural Machine Translation have made designing high-resource language translation models achievable — for big tech companies with access to large sets of parallel, translated data. However, low-resource languages like Bambara — ones that do not have extensive digital data sets and online communities, often because they are primarily spoken and not written — have benefited much less. Additionally, linguistic research is often lacking in low-resource languages, making conventional natural language processing (NLP) tasks more difficult.

After they have been approved and been given access to a common data set, competitors will create Bambara-to-French and French-to-Bambara translation models. Our **Competition page** includes additional information about evaluation criteria, prizes, workshop opportunities, and more. The competition begins on August 1st.

Dr. Luger's professional background blends NLP engineering and product architecture, responsible AI, conversational AI, and emerging machine translation technologies for under-resourced languages. She built several "first-of-a-kind" NLP solutions — including automated email and text response experiences that are mainstream today — and received both a Master of Science and PhD in Informatics from University of Edinburgh in Edinburgh, Scotland.

Bambara-French Machine Translation Competition
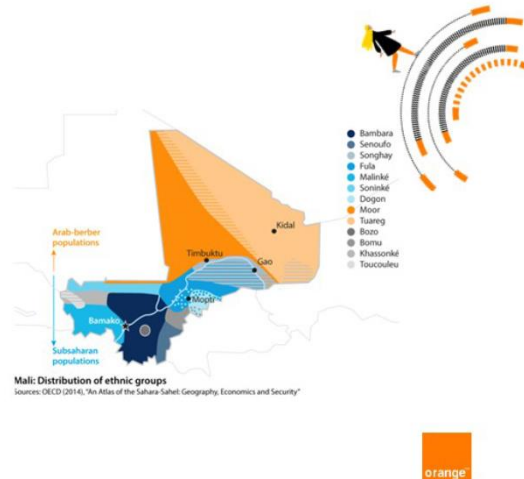
Listen 🔊

Introduction

Welcome to the **Orange Silicon Valley Bambara-French Machine Translation Competition**. This competition focuses on the development of a machine translation model for Bambara, a low-resource language, as part of an overall goal of improving global language representation across technologies.

Orange Silicon Valley

Bambara-French Machine Translation Competition

We provide the data, you provide the algorithms!

Cash prizes and research workshop open to all after the competition.

Mali: Distribution of ethnic groups
Sources: OECD (2014), "An Atlas of the Sahara-Sahel: Geography, Economics and Security"

Bambara
Senoufo
Songhay
Fula
Malinké
Soninké
Dogon
Moor
Tuareg
Bozo
Bomu
Khassonké
Toucouleu

Resources

- Competition introduction and registration (EN)
- Evaluation and competition details (EN)
- Competition rules (EN)

Back to top

- Concours de Traduction Automatique Bambara-Français (FR)
- Évaluation (FR)
- Règlement du Concours de Traduction Automatique Bambara-Français (FR)

Contact

Sarah Luger
Principal, Technology Group

Please contact Dr. Sarah Luger with any questions about the Bambara-French Machine Translation Competition.

Hosted by Orange Silicon Valley

Orange Silicon Valley is a US subsidiary of Orange, one of the world's leading telecommunications operators serving more than 253 million customers across 27 countries. Learn more about us here.

Bambara is a one of the closely related Manding languages, which are part of the larger Mandé language family. It is predominantly spoken in Mali and is one of the country's 13 national languages, with 15 million speakers and 5 million native speakers.

Low-resource language challenges

Recent advancements in Neural Machine Translation have made the development and design of high-performing translation models for high-resource languages an accessible task for big tech companies with access to large sets of parallel, translated data.

However, low-resource languages — ones that do not have extensive digital data sets and online communities, often because they are primarily spoken and not written — benefit much less from these recent advancements. Additionally, linguistic research is often lacking in low-resource languages, making conventional natural language processing tasks more difficult.

The competition

In this competition you will create Bambara-to-French and French-to-Bambara translation models. See our evaluation and competition details page and our rules page for additional specifics.

To register for this competition you must fill out the form at the bottom of this page. After we receive your email, you will be given full access to our training data.