

A large grid of approximately 100 small portrait photographs of people from various ethnicities, ages, and genders, arranged in a grid pattern. The central part of the grid is overlaid with a dark grey rectangle containing text.

Oxford Global Languages

A defining project

Collaboration and Computing for
Under-Resourced Languages
(CCURL)

23rd May 2016

Jon French | Advocacy and Partnerships Manager, Oxford Dictionaries
Phil Reimann | Product Manager, General Academic Dictionaries

Oxford University Press

Mission Statement

OXFORD
UNIVERSITY PRESS



Oxford University Press (OUP) is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education through its worldwide publishing.

*Publishing in 40+
languages,
branches and
representation in
100+ countries ...*



- *Largest university press in the world*
- *500 year old, non-profit institution*
- *2014/2015 revenue: £767 million*
- *Proportion of surplus transfers to the University each year; rest is invested*
- *Employees: 6000+*
- *Digital products now account for more than half of academic turnover*

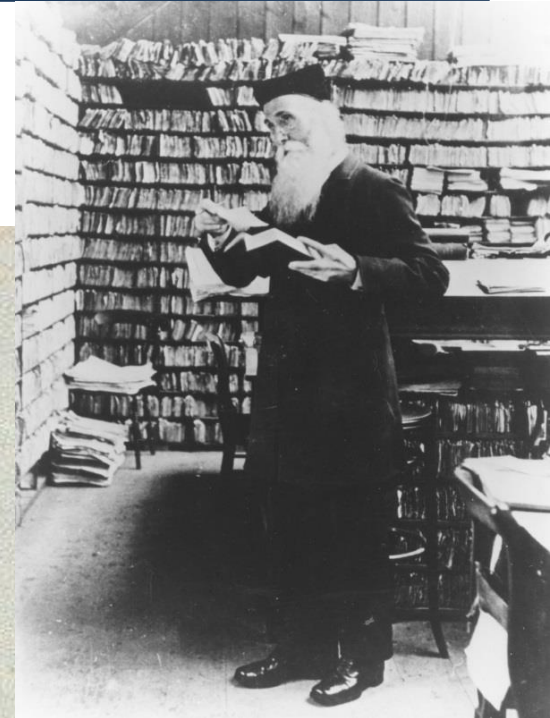
OUN dictionaries:

A tradition of public involvement

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY
PRESS

AN APPEAL
TO THE
ENGLISH-SPEAKING AND ENGLISH-READING PUBLIC
TO READ BOOKS AND MAKE EXTRACTS FOR
THE PHILOLOGICAL SOCIETY'S
NEW ENGLISH DICTIONARY.



Outline

Part 1 : Oxford Global Languages

- The challenge
- The response
- Phase 1: The beginning ...

Part 2 : Building the future

- Focus on usable outputs
- Working with communities
- Start small ...
- Partnerships– external and internal

Where we are and next steps

Lexical Engine and Platform and APIs

-



Oxford Global Languages

Further context

- Demand for multilingual resources
- Democratization of content
- Increase in mobile
- Global tech companies

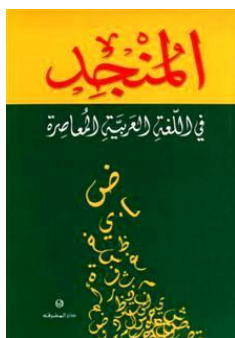


Before Oxford Global Languages

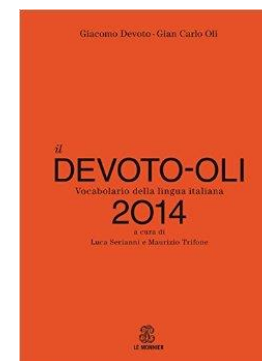
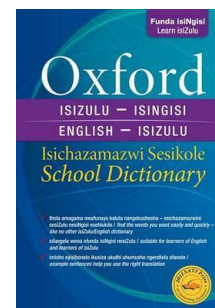
A global licensing programme

OXFORD
UNIVERSITY PRESS

47 datasets, 33 languages: monolingual dictionaries, bilingual dictionaries, and inflected forms



A word cloud of 33 languages: Dutch, Danish, French, Hebrew, Greek, Hindi, Turkish, Czech, Northern, Swedish, Malay, Vietnamese, Finnish, Japanese, Chinese, Hungarian, German, Zulu, Portuguese, Romanian, Latvian, Arabic, Norwegian, Bulgarian, Spanish, Croatian, Thai, English, Korean, Italian, Russian, Slovene, Polish, and Sotho.



Before Oxford Global Languages ...

Some of our licensing partners...

OXFORD
UNIVERSITY PRESS

amazon

Google

SHARP

CASIO®

SONY



Microsoft



Oxford Global Languages

The response

OXFORD
UNIVERSITY PRESS

- Focus on digitally under-represented languages
- Develop content quickly
- Inspire language communities
- Make it social
- Reusable content





- **Localized websites and digital access for 100+ languages**



- **Language communities contributing content**



- **Supporting digitally under-represented languages**



- **Flexible data serving multiple needs**

Oxford Global Languages

What it is *not*

- *An educational programme for schools*
- *A research or academic unit*
- *Focused on endangered languages*
- *A government sponsored initiative*

***But we share a lot in common
and can work together!***



Oxford Global Languages

The beginning: Phase 1

OXFORD
UNIVERSITY PRESS



October 2014 - September 2016: Phase 1

Goals

- 10 language **websites** with 3-4 active **communities**
- Complete **workflow** trialled and tested
- **Data** models and platform able to store and update multilingual interlinked content
- First usable outputs for third parties

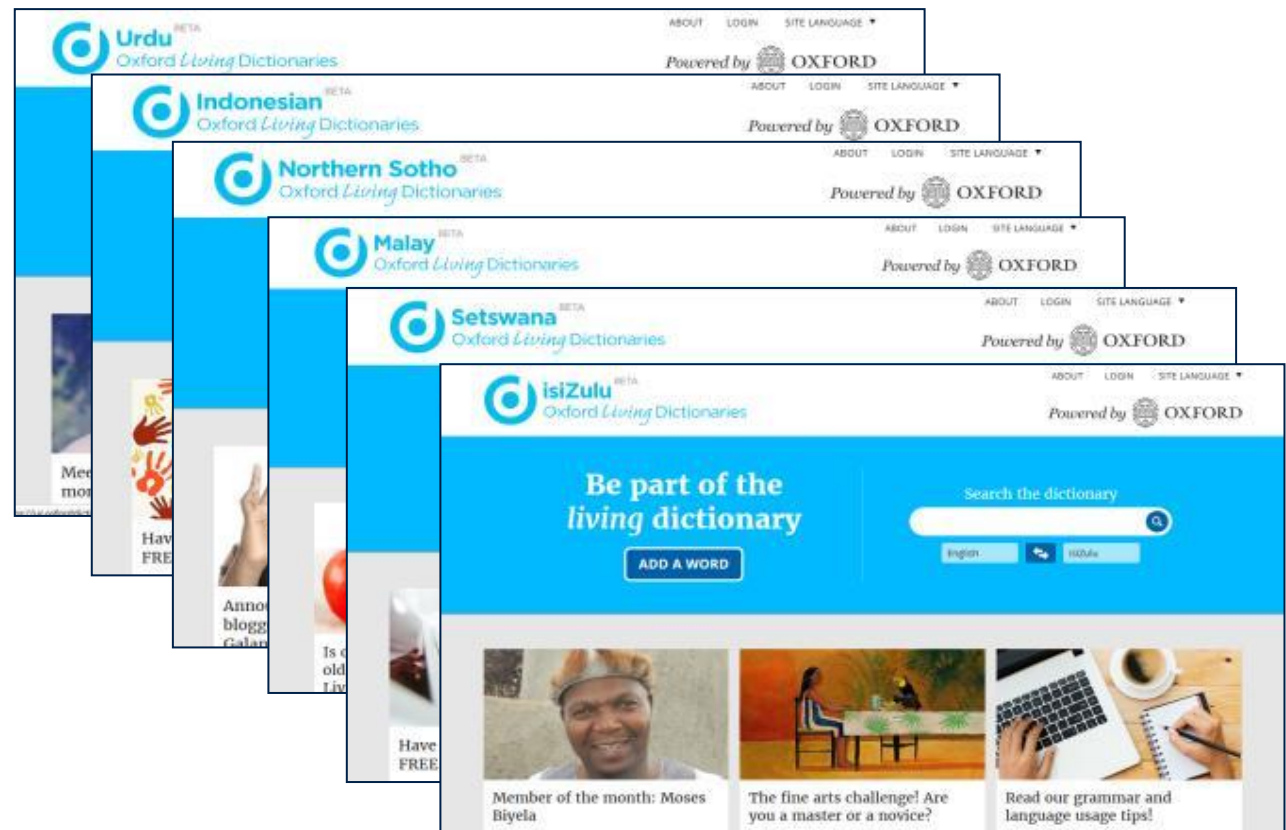
Oxford Global Languages

Phase 1

OXFORD
UNIVERSITY PRESS

**Urdu,
Indonesian,
Northern Sotho,
Malay,
Setswana,
isiZulu**
have been
launched...

Romanian, Latvian
Hindi, Kiswahili to
complete Phase 1 of
programme



Outline

Part 1 : Oxford Global Languages

- The challenge
- The response
- Phase 1: The beginning ...

Part 2 : Building the future

- Focus on usable outputs
- Working with communities
- Start small ...
- Partnerships– external and internal


Where we are and next steps

Lexical Engine and Platform and APIs


Building the future ...

Focus on usable outputs ...

OXFORD
UNIVERSITY PRESS

 **Northern Sotho** BETA
Oxford Living Dictionaries


MABAPI LE TSENA LELENE LA SAETE ▼

Powered by  OXFORD


E ba karolo ya Pukuntšu ye Phelago

Tsoma ka pukuntšung


Search bar with dropdown menu and search icon

Sesimane  Sesotho sa Leboa


TSENYA LENTŠU



Naa o keteka bjang Letšatši la Bašomi?



Tlhohlo ya bogabo bja pogo – naa o matwetwe goba o sa le matlogadibešong?




Kopana le setšhaba sa rena: Biteyo ye e nago le Victor Mojela

Lentšu la Letšatši

nyanyeng

ADVERB

HWETŠA TLHALOŠO



Lentšu la Letšatši Nyanyeng Lehlathi
Hwetša Tlhalošo <http://bit.ly/1UVt8FP>


Posted 7 hours ago

Have you signed up for your **FREE** account?

START ADDING WORDS TO THE LIVING DICTIONARY TODAY

Mantšu a maswa a tsentswego

wait



Baabi ba bagolo

mmunhuru494

Imaseko895

 **isiZulu** BETA
Oxford Living Dictionaries

 **Setswana** BETA
Oxford Living Dictionaries

 **Indonesian** BETA
Oxford Living Dictionaries

 **Urdu** BETA
Oxford Living Dictionaries

 **Malay** BETA
Oxford Living Dictionaries

Building the future ...

Focus on usable outputs ...

OXFORD
UNIVERSITY PRESS



ABOUT LOGIN SITE LANGUAGE ▼

Powered by  OXFORD

Zu > Eng

Type word or phrase

Home > Translate > isiZulu to English > umfundi

Translation of *umfundi* in English

umfundi

NOUN (class 1, plural **abafundi**)

1. **learner, student**

Examples

Umfundi ngamunye makabukeze umsebenzi wakhe
Each learner must revise their work

2. **reader**

Examples

Umama umfundi wephephandaba iLanga
Mother is a reader of the Ilanga newspaper

Be part of the living dictionary

ADD A WORD

Newly added words

Umabonakude
kingMado

practitioner
nkosie

Parenthesis
Imvelase73

Embezzle
Imvelase73

ADD A WORD

Non-lexical content

Developing language specific resources

OXFORD
UNIVERSITY PRESS

Sertai kamus maya

TAMBAH PERKATAAN

Mencari kamus

Inggeris



Bahasa Melayu



Adakah anda memerlukan bantuan untuk bermula?



Dari manakah sumber makanan anda? Pilih dari senarai perkataan



Panduan tatabahasa, tip menulis dan banyak lagi

Perkataan hari ini

telur

NOUN

KETAHUI MAKSUDNYA



Ikuti kami di Facebook

Sudahkah anda mendaftar untuk akaun **PERCUMA**?

TAMBAHKAN PERKATAAN KE KAMUS MAYA SEKARANG

Be part of the living dictionary

ADD A WORD

Search the dictionary

English



Malay



Do you need help getting started?



Where does your food come from? A harvest wordlist



Grammar guides, writing tips and more

Word of the Day

telur

NOUN

FIND OUT WHAT IT MEANS



Ikuti kami di Facebook

Have you signed up for your **FREE** account?

START ADDING WORDS TO THE LIVING DICTIONARY TODAY


Short (500 word) articles on language usage

Building the future ...

A place for discussion

Forums

One per language
with aim to develop
the language
community

 **isiZulu** BETA
Oxford Living Dictionaries

NGENA BHALISA

Bhaka IYA

Ikhaya

**Indawo lapho kuxoxixwasana khona
yesichazamazwi sesiZulu sakwa-Oxford**

Wamukelekile Esichazamazwini Esiphilayo sesiZulu!
Isimemezelo Ukubuka okungu-281 Ukuphawula okungu-10 Okwakamuva kakhulu ngokuthi Sibusiso
April 26 Okwejawalekile

Imiyalelo yamalungu yezichazamazwi eziphilayo zesiZulu zakwa-Oxford
Isimemezelo Ukubuka okungu-11 Ukuphawula okungu-0 Iqaliswe ngu-KatherineSoroya March 15
Imiyalelo yamalungu yezichazamazwi eziphilayo zesiZulu zakwa-Oxford

Nomzamo Mzobe
Ukubuka okungu-15 Ukuphawula okungu-1 Okwakamuva kakhulu ngokuthi Sibusiso April 15 Igrama,
izimpawu zokubhala nokusetshenziswa kolimi

Umbhalomagama nokuphinyiswa kwamagama
Ukubuka okungu-5 Ukuphawula okungu-1 Okwakamuva kakhulu ngokuthi Sibusiso April
15 Umbhalomagama nokuphinyiswa kwamagama

Umhlangano oqondene nokuhumusha
Isimemezelo Ukubuka okungu-36 Ukuphawula okungu-3 Okwakamuva kakhulu ngokuthi Sibusiso
March 18 Izihumusho

Iziphakamiso
Ukubuka okungu-6 Ukuphawula okungu-0 Iqaliswe ngu-Sibusiso February 23 Okwejawalekile

Izigaba

Zonke Izigaba 9

**Oxford Living Dictionaries -
all languages** 1

**Imiyalelo yamalungu
yezichazamazwi eziphilayo
zesiZulu zakwa-Oxford** 1

Okwejawalekile 2

**Igrama, izimpawu zokubhala
nokusetshenziswa kolimi** 3

**Umbhalomagama
nokuphinyiswa kwamagama** 1

Izihumusho 1

Building the future

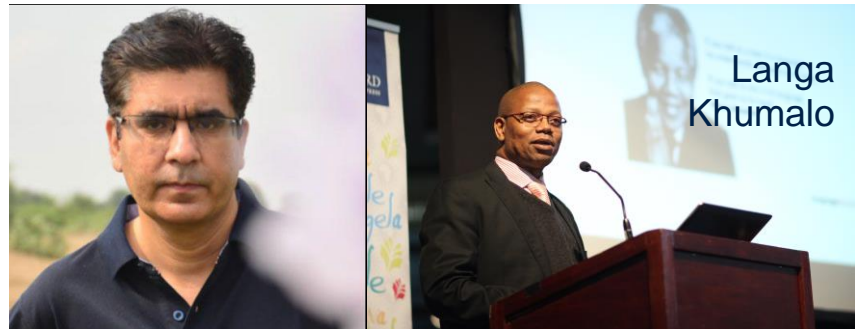
Communities and the 'living dictionary'

OXFORD
UNIVERSITY PRESS

**Managed by and for
the communities**

**Assisted by:
Language Managers,
who oversee, moderate,
and curate
Language Champions,
who promote and
advocate**

Zafar Syed



Dr Rafizah
Mohd Rawian

Building the future

Start small ...

OXFORD
UNIVERSITY PRESS

- Experimental approach
- Launch quickly
- Focus on users
- Partnerships
- Start small and build!



<https://ms.oxforddictionaries.com/> <https://ur.oxforddictionaries.com/>

Building the future

Partnerships – internal and external

- Pool and test ideas
- Share expertise
- More collective resource
- Move faster
- Access to new markets
- New business models



OGIL Partnerships Manager
jon.french@oup.com

Outline

Part 1 : Oxford Global Languages

- The challenge
- The response
- Phase 1: The beginning ...

Part 2 : Building the future

- Focus on usable outputs
- Working with communities
- Start small ...
- Partnerships– external and internal

Where we are and next steps

Lexical Engine and Platform and APIs

Where we are and next steps

Phase 1

We are on track

- Ten languages in progress
- Base technology of data platform in production
- First six languages launched
- Iterative website development
- First communities are live!



Where we are and next steps

OGL launched 2nd September 2015

OXFORD
UNIVERSITY PRESS

Oxford Global Languages: Northern Sotho and isiZulu

The screenshot shows the Oxford Dictionaries website interface. At the top, the logo 'Oxford Dictionaries BETA' is visible. Below it, a search bar contains the text 'Bhala igama noma umshwana'. The breadcrumb trail reads 'Ikhaya > Humusha > isiZulu kuya esiNgisini > ikhekhe'. The main entry for 'ikhekhe' is displayed, with the sub-header 'IBIZO' and the first definition '1. cake'. Below this, the word 'Izibonelo' is followed by a descriptive sentence in isiZulu: 'Unesitifiketi sokupheka, nokubhaka amakhekhe nokuthunga asizuza eNanda'. A quote in English follows: 'She has a certificate to cook, to bake cakes, and to sew, which she obtained at iNanda (College)'.

<https://zu.oxforddictionaries.com/>

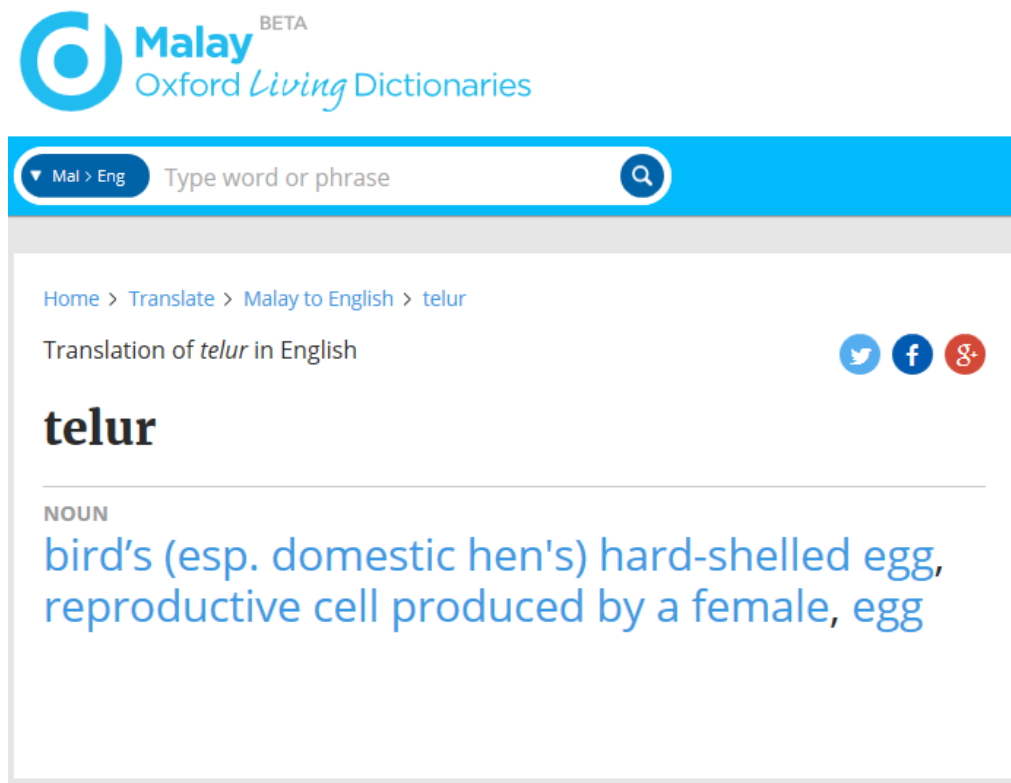
The screenshot shows the Oxford Dictionaries website interface. At the top, the logo 'Oxford Dictionaries BETA' is visible. Below it, a search bar contains the text 'Type a word or phrase'. The breadcrumb trail reads 'Home > Translate > Northern Sotho to English > apola'. The main entry for 'apola' is displayed, with the sub-header 'NOUN' and the first definition '1. apple'. Below this, the word 'Examples' is followed by two sentences: 'Legora la gagwe le ba le tletše ka mehlare ya diapola' and its English translation 'His yard was full of apple trees'.

<https://nso.oxforddictionaries.com/>

Where we are and next steps

Malay and Urdu Language sites launched March 2016
Indonesian and Setswana sites launched in May 2016

OXFORD
UNIVERSITY PRESS



<https://ms.oxforddictionaries.com/>

Where we are and next steps

Phase 2



Planning for Phase 2: (3 years)

- Many more languages and communities
- Content development
- Digital accessibility
- API programme for developers and licensees
- Many more strategic partnerships
- Internal OUP synergies

Where we are and next steps

Proposed phase 2 initial languages



| | | | | | |
|------------|------------------|----------|-----------|------------------|---------|
| Vietnamese | Amharic | Burmese | Bulgarian | Southern Quechua | Uzbek |
| Yoruba | Filipino/Tagalog | Hausa | Punjabi | Shona | Turkish |
| Slovak | Tamil | Gujarati | Estonian | Tok Pisin | Khosa |
| | | | | | Igbo |

Outline

Part 1 : Oxford Global Languages

- The challenge
- The response
- Phase 1: The beginning ...

Part 2 : Building the future

- Focus on usable outputs
- Working with communities
- Start small ...
- Partnerships– external and internal

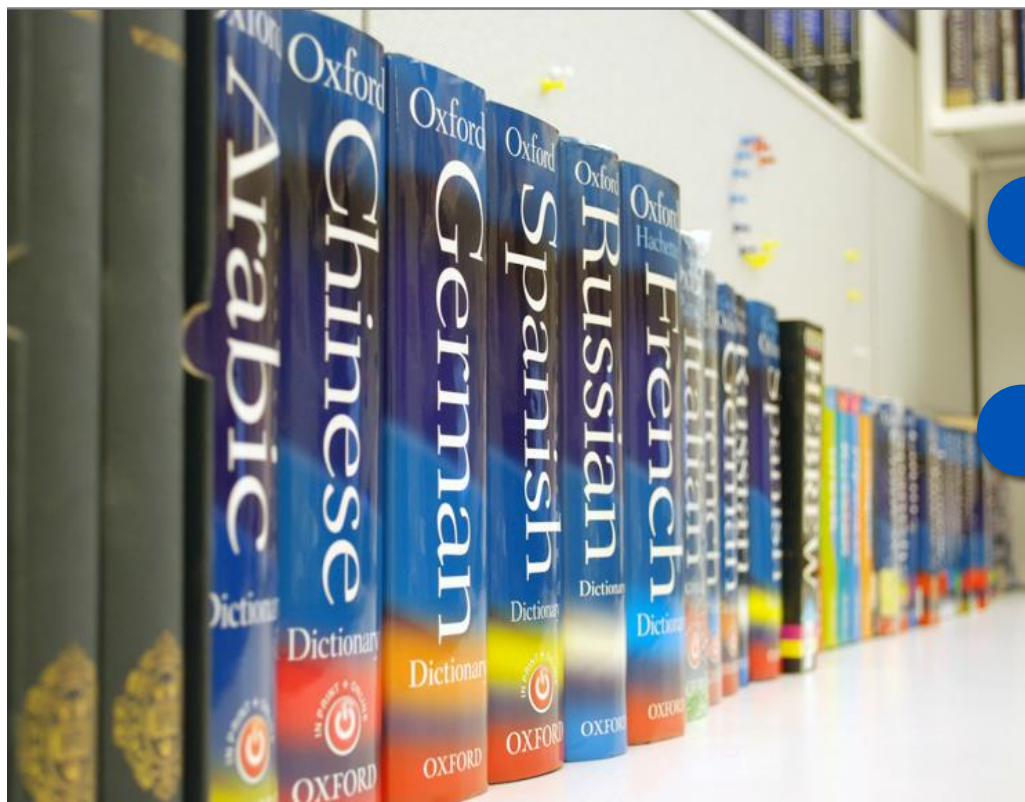
Where we are and next steps

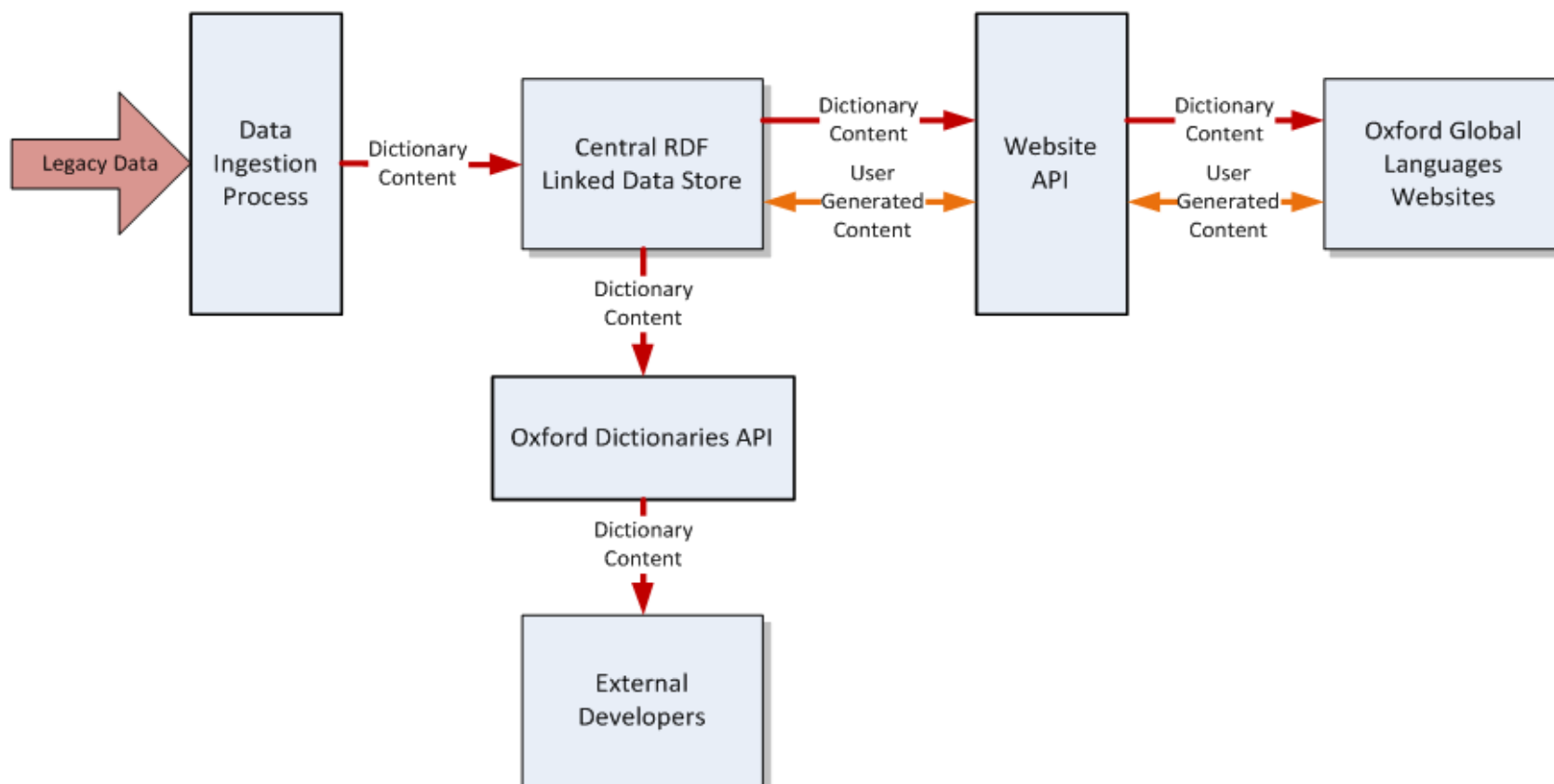
Lexical Engine and Platform and APIs

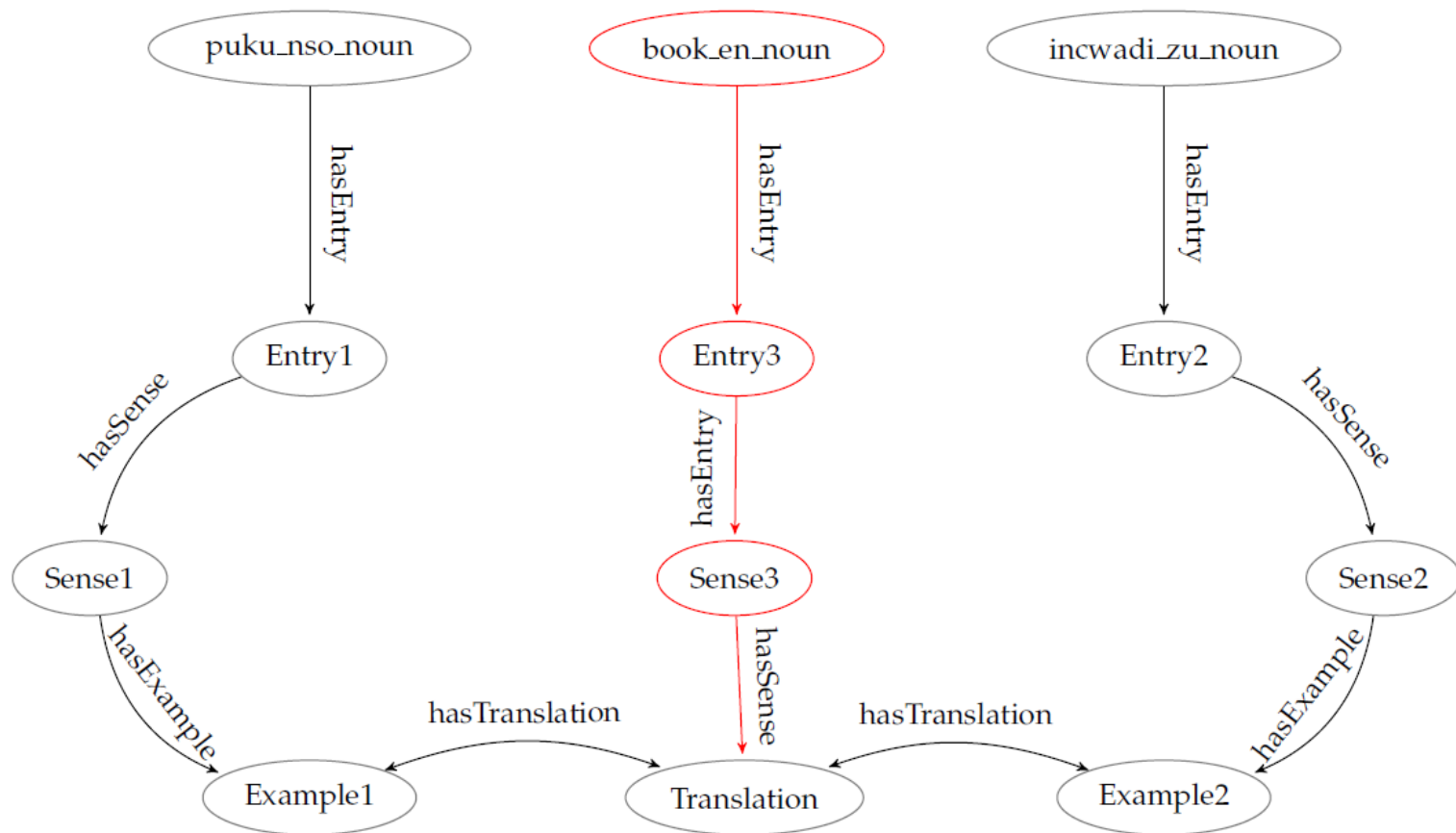
LEAP

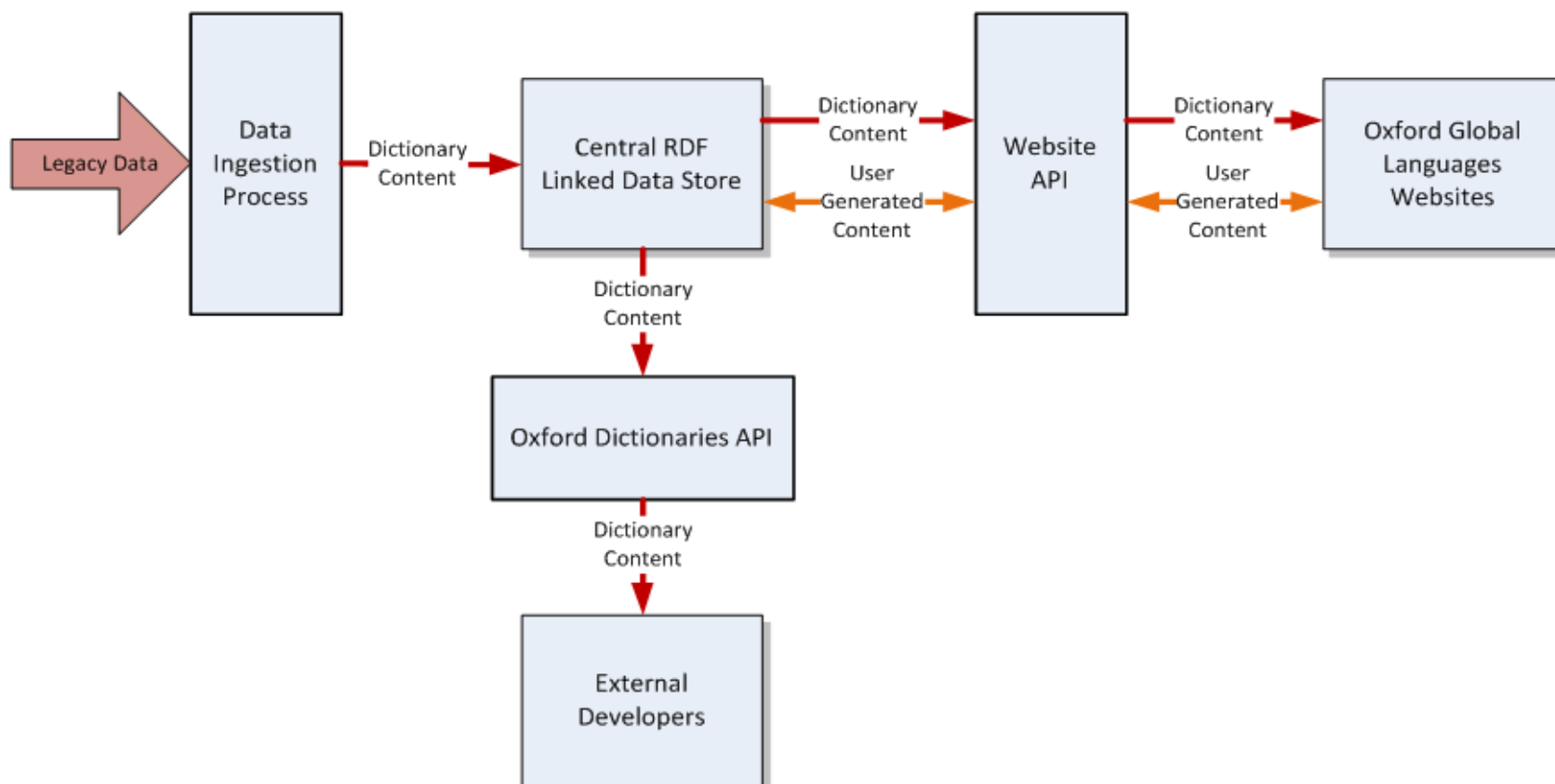
Lexical Engine and Platform

OXFORD
UNIVERSITY PRESS









APIs

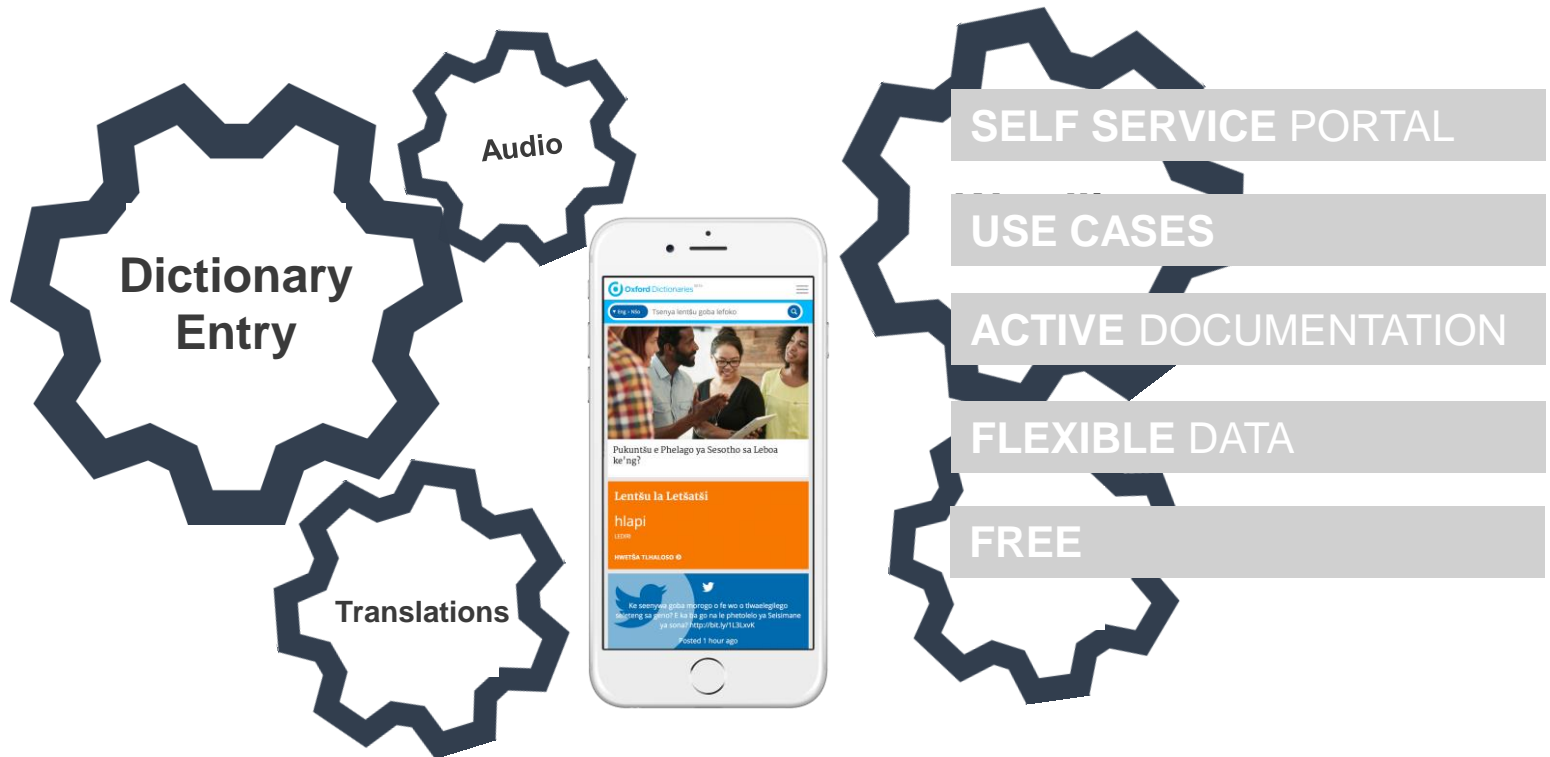
Oxford Dictionaries API



Oxford Dictionaries API



Oxford Dictionaries API



Oxford Dictionaries API

EARLY ADOPTER LAUNCH June 2016

- At least 6 months' free, unlimited access ✓
- New datasets, new functionality ✓
- Continuous improvement ✓

PUBLIC LAUNCH Late 2016

- Free for research and non-commercial use ✓

Get in touch

developer.oxforddictionaries.com



Don't Miss!

OXFORD
UNIVERSITY PRESS

Towards a Linguistic Ontology with an Emphasis on Reasoning and Knowledge Reuse

Artemis Parvizi

Oxford University Press

Session: O5 - LR Infrastructures and Architectures

Wednesday 3.05pm.

A large grid of approximately 60 small, square portrait photographs of people from various ethnicities, ages, and genders, arranged in a grid that frames the central text. The grid is composed of several rows and columns, with some portraits partially obscured by the text overlay.

Oxford Global Languages

A defining project

Collaboration and Computing for Under-Resourced Languages (CCURL)

23rd May 2016

Jon French | Advocacy and Partnerships Manager, Oxford Dictionaries
Phil Reimann | Product Manager, General Academic Dictionaries